# The Effect of Animal Venom in the Treatment of Pancreatic and Colorectal Cancer

**by**

**Emily Knight**

**Canterbury Christ Church University**

**Thesis submitted**

**for the Degree of Doctor of Philosophy**

2021

# Table of Contents

# Table of Figures

# Table of Tables

# List of Abbreviations

**1D:** First dimension

**2D:** Second dimension

**3FTx:** Three-finger toxins

**AA:** Amino acid

**ACEIs:** ACE inhibitors

**ACN:** Acetonitrile

**Actin:** β-Actin

**ADME:** Absorption, distribution, metabolism and excretion

**AMI-MS:** Acoustic Mist Ionisation Mass Spectrometry

**ARBs:** Angiotensin II receptor blocker

**ATCC**: American Type Culture Collection

**Bad:** BCL-2 associated agonist of cell death

**Bax:** BCL-2 associated X/ BCL-2-like protein 4

**Bcl-2:** B-cell lymphoma 2

**BLASTp:** Protein Basic Local Alignment Search Tool

**Bp:** Base pairs

**CCBs:** Calcium channel blockers

**CCK:** Cholecystokinin

**cDNA:** Complementary DNA

**CITES:** The Convention on International Trade in Endangered Species of Wild Fauna and Flora

**CML:** Chronic myeloid leukaemia

**$CO_2$:** Carbon dioxide

**Cq:** quantification of cycles

**CT:** Cytotoxin (or cardiotoxin)

**Da:** Dalton

**DMSO:** Dimethyl sulfoxide

**DNA:** Deoxyribonucleic acid

***E.coli*:** *Escherichia coli*

**EDA:** Effect-directed analysis

**EGF(R):** Epidermal growth factor (receptor)

**EMBL:** European Molecular Biology Laboratory

**EMBL-EBI:** European Molecular Biology Laboratory - European Bioinformatics Institute

**EMBL EST:** European Molecular Biology Laboratory Expressed Sequence Tags

**ErbB/HER:** Human epidermal growth factor receptor

**ESI:** Electrospray ionisation

**FCS:** Foetal calf serum

**GAPDH:** Glyceraldehyde 3-phosphate dehydrogenase

**$H_2O$:** Water

**$H_2O_2$:** Hydrogen peroxide

**HER2:** Human epidermal growth factor receptor 2

**HER3:** Human epidermal growth factor receptor 3

**HER4:** Human epidermal growth factor receptor 4

**HPLC:** High performance liquid chromatography

**HTS:** High throughput screening

**$IC_{50}$:** Inhibitory concentration for 50% of test population

**ID:** Identification

**IE:** Ion exchange (chromatography)

**IGF:** Insulin-like growth factor

**KDa:** Kilodalton

**LAAO:** L-amino acid oxidase

**$LD_{50}$:** Lethal dose for 50% population

**MALDI:** Matrix-assisted laser desorption/ionisation

**MIQE:** The Minimum Information for Publication of Quantitative Real-Time PCR Experiments guidelines

**MMP:** Matrix metalloproteinase

**MMP:** Mitochondrial membrane potential

**MS:** Mass spectrometry

**MSA:** Multiple sequence alignment

**MS/MS:** Tandem mass spectrometry

**MTT:** 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

**MW:** Molecular weight

**MY(A):** Million years (ago)

**m/z:** Mass to charge ratio

**NanoDSF:** Nano differential scanning fluorimetry

**N.anc:** *Naja anchietae*

**N.aan:** *Naja annulifera*

**N.atr:** *Naja atra*

**N.haj:** *Naja haje*

**N.hle:** *Naja haje legionis*

**N.kao:** *Naja kaouthia*

**N.naj:** *Naja naja*

**N.nka:** *Naja naja karchachensis*

**N.nct:** *Naja nigricincta nigricincta*

**N.nig:** *Naja nigricollis*

**N.mel:** *Naja melanoleuca*

**N.mos:** *Naja mossambica*

**N.nct:** *Naja nigricincta*

**N.niv:** *Naja nivea*

**N.nub:** *Naja nubiae*

**N.pal:** *Naja pallida*

**N.sam:** *Naja samarensis*

**N.sia:** *Naja siamensis*

**N.spu:** *Naja sputatrix*

**NCBI:** National Centre for Biotechnology Information

**NCBInr:** National Centre for Biotechnology Information non-redundant proteins

**NTC:** No template control

**O.han:** *Ophiophagus Hannah*

**PARP:** Poly (ADP-ribose) polymerase

**PBS:** Phosphate Buffered Saline

**PI3K:** Phosphatidylinositol-3 kinase

**PLA2:** Phospholipase A2

**qPCR:** quantitative polymerase chain reaction

**RP:** Reverse phase (chromatography)

**RNA:** Ribonucleic acid

**SAR:** Structure activity relationship

**S/B:** Signal to background ratio

**SEC:** Size exclusion chromatography

**SMAD4:** Mothers against decapentaplegic homolog 4

**STAT:** Signal transducer and activator of transcription

**STR:** Short tandem repeat analysis

**TFA:** Trifluoroacetic acid

**TGF-β:** Transforming growth factor beta

**TP53:** Cellular tumour antigen p53

**ToF:** Time of flight

**Tubulin:** α-Tubulin

**VEGF(R):** Vascular endothelial growth factor (receptor)

**Z':** Z-factor measure of statistical effect size

**α:** Alpha

**β:** Beta

**γ:** Gamma

# Acknowledgements

I would like to thank my supervisor Carol Trim for all the advice and support, helping me develop my research and scientific skills. You have not just been a supervisor to me but also a close and supportive friend. Thank you for having patience and believing in me, even when I didn't always believe in myself. It has been a pleasure to be part of your lab and I am honoured to know you as a scientist and friend. Huge thanks go to my industrial supervisor Steve Trim who has been extremely influential and provided endless advice and consultation about various scientific techniques including plate-based screening, HPLC and qPCR. You always made time for me and went above and beyond to support me with infinite patience and wisdom. Thank you for welcoming me into your team and teaching me a huge proportion of what I know about venom and drug discovery. I would also like to thank Cornelia Wilson for being a hugely supportive second supervisor, lab manager and friend and for offering endless advice and support in the lab. You were a constant presence at the lab and a truly inspirational scientist. Thank you to Mike Weed for being my chair for the first three years of my PhD and ensuring I had regular supervisory meetings and for providing enthusiasm and encouragement about my field of study.

My gratitude goes to Canterbury Christ Church University for funding and supporting my studies. Thank you to Kent Cancer Trust whose kind donation provided funding for parts of this project. Without your generous gift, this research would not have been possible. Thank you to Venomtech Ltd. for access to equipment and materials essential for undertaking this research including all the venom used in this study. Many thanks to the University of Manchester and the University of Kent for generous donations of cell lines, without which I would not have had the resources required to proceed with this study. Thanks go to my fellow PhD students and Life Science Instructors for providing a supportive community and for relief after the particularly difficult weeks! Particular thanks go to Danielle McCullough and Alice Tirnoveanu for being an incredibly supportive friends and colleagues, helping me navigate through the PhD process.

Thank you to Christian Schmidt and Lucien Douglas-Green for being incredibly supportive partners at their respective points throughout this process. Your love and support have been essential to make it this far and I simply couldn't have done it without you. Thank you for pushing me to be my best and for catching me when I fell. I would like to thank my incredibly supportive family, particularly my mum, Sally Knight, who wasn't able to see me complete my PhD journey, but I know would be very proud of me.

# Abstract

Pancreatic and colorectal cancer are aggressive, difficult to target cancers with new therapeutic options desperately required to improve prognosis. Animal venoms are useful in drug discovery due to huge variety of bioactive components evolved over millions of years. Venom has already been utilised in drug discovery with several venom derived drugs currently on the market.

A plate-based resazurin assay was used to determine cytotoxicity of a panel of cobra venoms against BxPC-3 pancreatic and SW620 colorectal cancer cell lines *in-vitro*. African spitting cobra (*Afronaja*) venoms displayed selective toxicity against SW620 cells compared to non-spitting and Asian cobra venoms. Five *Afronaja* venoms were fractionated using HPLC then the venom components were rescreened in a miniaturised resazurin assay. Dose response curves for both lines were performed using selected fractions, giving $IC_{50}$ values between 17-225 µg/ml.

Four venom fractions were analysed using Mass Spectrometry (MS) and identified as cytotoxins. MS outputs plus bioinformatic techniques predicted likely sequences for each fraction. A structure activity relationship was performed and AA residue 7 and AAs 26-29 were identified as conferring the most significant selective toxicity.

Finally, a preliminarily qPCR study assessed the effect of each fraction on the tested cell lines. This study investigated 20 genes commonly mis-regulated in pancreatic and colorectal cancer. The cell lines were exposed to three venom fractions and regulation of each gene assessed compared to untreated controls. SMAD4, Tp53, WNT1 and EGFR genes were significantly upregulated following addition of venom and MMP9 and Bcl-2 were significantly downregulated in this preliminary study. A larger scale qPCR study should be performed to confirm these findings and assess other potential genetic alterations caused by the venom fraction to assess the potential for development into an anti-cancer treatment option.

# Chapter 1 - Introduction

## 1.1 The Research Problem

Pancreatic cancer (PC) is a rare and aggressive form of cancer which has a particularly poor prognosis. Lack of specific symptoms of PC leads to late diagnosis, often after the disease has already progressed or metastasised. This is coupled with resistance to most forms of treatment including chemotherapy, radiotherapy and hormonal therapy. For most patients the best treatment option is surgery which increases average survival from ~5 months to 23 months when combined with adjuvant therapy (Gillen *et al.*, 2010). However, only 10-20% of diagnosed patients are eligible for surgical intervention due to their cancer already being locally advanced or metastasised. This leaves the remaining patients with few to no treatment options, illustrating the urgent need for development of novel treatments for PC and highlighting the importance of early diagnosis.

Traditional drug discovery focuses on the detection and optimisation of small chemical drug molecules. Most of the currently available cancer therapies are not specific enough and lead to numerous side effects. Many promising novel therapies and combinations of existing therapies fail to deliver in a clinical setting. Venoms and other natural products have been used for various conditions in the past including for cancer therapies. However, these have not yet been fully explored in the context of PC.

## 1.2 Pancreatic Cancer

### 1.2.1 Incidence, Morbidity and Mortality Rates

Pancreatic cancer (PC) is an aggressive and often fatal disease. Although it is only the tenth most common type of cancer, accounting for just 3% of new cancer cases, it is the sixth most common cause of cancer related death leading to 6% of cancer related deaths (Cancer Research, 2018b). It is predicted that by 2030, PC will become the second most common cause of cancer related death, overtaking other cancer types including breast, prostate and colorectal (Rahib *et al.*, 2014). In the UK between 2016-2018, approximately 10,500 patients were diagnosed with PC and in the same time period, 9,400 patients died from PC (Cancer Research, 2018b). Of those diagnosed, one quarter (25.4%) of patients in England survive for more than one year after diagnosis. This figure decreases to just 7.3% of patients surviving for more than five years and 5% of patients surviving for more than ten years after diagnosis (Cancer Research, 2018b). This illustrates the bleak outlook that currently faces patients. Despite increasing amounts of research being invested in PC, these survival figures have barely changed for more than four decades.

Colorectal cancer (CRC) is currently the fourth most common type of cancer and the second leading cause of cancer related death in the UK, accounting for 11% of new cancer cases and 10% of cancer-related deaths (Cancer Research, 2018a). There were approximately 42,900 cases of CRC diagnosed between 2016-2018 with 16,600 deaths in the same time frame. Mortality rates in the UK are predicted to fall by 23% by 2035 (Cancer Research, 2018a). Approximately 78.3% of patients diagnosed with CRC survive at least one year with 58.4% patients surviving for at least 5 years and 52.9% of patients surviving for more than 10 years after diagnosis. CRC treatment has improved considerably in the last four decades and survival has more than doubled in this time period (Cancer Research, 2018a).

Currently, no routine screening programme is in place for the diagnosis of PC. Diagnostic tools exist for referred patients, normally after exhibiting symptoms. However, the lack of specific symptoms displayed in PC make this form of diagnosis difficult and time consuming. This coupled with the rapid tumour progression normally displayed in PC combine to lead to late stage diagnosis with over 50% of patients having cancer that is already metastatic by time of diagnosis (Gillen *et al.*, 2010). This leaves very few treatment options available and most of these patients undergo palliative care only. For those who are diagnosed before the point of metastasis, only around 10% of PCs are detected early enough to be eligible for a resection (Hidalgo *et al.*, 2015). Of those who do undergo resection, 90% still succumb to the disease due to local recurrence or metastases away from the original site without adjuvant therapy (Hidalgo *et al.*, 2015). Despite this, resection is currently by far the most effective treatment option for PC and coupling surgery with chemotherapy and/or immunotherapy provides the best outlook for patients.

Screening of patients who have a family history of PC may prove to be beneficial, however only 10% of PCs are thought to be familial (Alsubai *et al.*, 2016). Other factors are at play asides from inherited genetic mutations and as such, whole population screening has been suggested to be unnecessary and not cost effective at this time (Geurts *et al.*, 2015). CRC on the other hand is thought to have a stronger familial link with 25% of cases caused by inherited genetic mutations (Stoffel and Kastrinos, 2014).

Since 2006, all patients in the UK between the age of 60-74 are invited to participate in a CRC screening program involving a home stool test every 2 years. If tested positive, patients are invited to attend a follow-up colonoscopy. Screening in this manner has helped to detect CRC at an earlier stage, improving morbidity and mortality levels in these patients (Logan *et al.*, 2012). Treatment for CRC depends upon the specific type and the stage of cancer.

### 1.2.2 Causes and Risk Factors

Age has a strong correlation with development of both PC and CRC. For PC, patients become most at risk in the 60-70 year age bracket, although PC has shown the phenomenon of 'genetic anticipation' with some patients with a family history of PC showing a two decade reduced age of onset compared with their affected relatives (Rulyak *et al.*, 2003).

CRC risk is significantly increased in patients over 50 years old (Levin *et al.*, 2008). Patients previously suffering with inflammatory bowel disease also have an increased risk of developing CRC in later life. Patients diagnosed with ulcerative colitis have a cumulative risk of CRC of 2% 10 years after diagnosis, 8% 20 years after diagnosis and 18% 30 years after diagnosis (Eaden *et al.*, 2001). Patients diagnosed with Crohn's Disease for 10 years have a cumulative risk of CRC of 2.9% (Canavan *et al.*, 2006). These statistics contrast with the general population risk of 0.7% per 10 years (Cancer Research, 2018a).

Smoking is a strong independent risk factor for the development of both PC and CRC. Smokers were shown to develop PC a decade earlier that non-smokers and male smokers aged less than 50 years were the most susceptible to this apparent increased risk (Rulyak *et al.*, 2003). It appears that it is the nicotine contained within cigarettes and other smoking products that increases the risk of developing PC, with nicotine also appearing to accelerate the rate of transformation of healthy cells to PC cells in mouse models (Hermann *et al.*, 2014). In addition, nicotine appears to promote the number and aggressiveness of circulating cancer cells (Hermann *et al.*, 2014). Tobacco smoking has been associated with a 10.8% increase in risk of developing CRC, with nicotine (and its metabolites) again suspected of being the cause of this increased risk (Botteri *et al.*, 2008, Cross *et al.*, 2014). Some of these risk factors appear to be additive, for example those who already have a family history of PC appear to have a further increased risk if they are a smoker (Rulyak *et al.*, 2003).

Another lifestyle factor thought to have an impact in PC is choice of diet and obesity. A strong association between raised BMI and increased risk of developing PC has been shown (Arslan *et al.*, 2010). A study conducted in Taiwan in 2011 concluded that diabetes of less than two years was found to be a risk factor for PC and may even be an early form of PC (Liao *et al.*, 2012). Long standing diabetes, however, did not correlate with risk of PC, and other factors such as age and presence of chronic pancreatitis or other co-morbidities had a larger effect on PC risk (Liao *et al.*, 2012). Conversely, some studies have found that diabetes does not appear to be a risk factor for the development of PC (Rulyak *et al.*, 2003) and so further study in this area is warranted.

### 1.2.3 Types and Subtypes of Pancreatic Cancer

There are two broad categories of PC – tumours of the endocrine pancreas (causing islet cell cancer) and tumours of the exocrine pancreas (causing pancreatic ductal adenocarcinoma (PDAC) and acinar cancer). PDAC is the most common and aggressive of these cancer types and accounts for around 90% of PC cases. Long before diagnosis of PC, pancreatic duct epithelial cells form precursor lesions known as pancreatic intrathelial neoplasias (PanIN). The three types of PanIN (PanIN-1, PanIN-2 and PanIN-3) reflect the corresponding severity (Murphy *et al.*, 2013). These non-invasive PanINs progress over time, becoming more severe, increasing in grade until they eventually culminate in the formation of an invasive neoplasia which goes on to become PDAC (Koorstra *et al.*, 2008).

Jones *et al*. (2008) conducted an in-depth genetic analysis of 24 human PCs. The average number of somatic mutations found in PDACs appears to be far less than in breast cancer or CRC, possibly due to the infrequent division of pancreatic epithelial cells compared to breast or colorectal cells. They concluded that there is an average of 63 genetic alterations in PDAC cells compared to noncancerous pancreatic cells, most of which are point mutations. These genetic alterations span 12 core signalling pathways that appear mutated in at least 67% of pancreatic tumours.

Collisson *et al*. (2011) conducted a study looking at the transcriptional profiles of a combination of primary pancreatic tumours and human and mouse cell lines of PDAC. Their findings suggested the existence of three sub-types of PC – classical, quasi-mesenchymal and exocrine-like. The classical type of PDAC shows high levels of expression of adhesion and epithelial genes with specific overexpression of GATA Binding Protein 6 (GATA6) and elevations in KRAS mRNA levels. The quasi-mesenchymal subtype on the other hand showed lower levels of GATA6 and KRAS but increased levels of mesenchyme associated genes. Patients with classical subtype showed improved survival following surgery. There was also a difference in response to different chemotherapeutic agents; patients with quasi-mesenchymal subtype were on average more sensitive to treatment with gemcitabine than those with classical subtype PDAC. Conversely, those with a classical subtype showed a better response to the epidermal growth factor receptor (EGFR) inhibitor erlotinib than those with quasi-mesenchymal subtype. This suggests that categorising patients by PDAC subtype could be beneficial for selecting the most appropriate therapy for them and thus improving overall survival whilst also minimising toxicity.

A virtual microdissection of PDACs was performed by Moffitt *et al.* (2015) which also suggested the presence of sub-types of PC. The group identified tumour- and stroma-specific subtypes with basal-like and classical tumour types. This produced a total of four subtypes (basal-like tumour normal stroma, basal-like tumour activated stroma, classical tumour activated stroma and

classical tumour normal stroma). Tumours with a basal-like characteristics had significantly poorer prognosis than those of the classical type with activated stroma appearing to be responsible for disease progression due to an inflammatory stromal response.

Further genetic analyses were conducted on pancreatic tumours by Bailey *et al.* (2016), which provided an even more detailed analysis of 456 PDACs. The team identified 32 recurrent mutations categorised into 10 pathways – KRAS, TGF-β, WNT, NOTCH, ROBO/SLIT signalling, G 1/S transition, SWI-SNF, chromatin modification, DNA repair and RNA processing. They identified four different sub-types of PC – Squamous (most similar to Collisson *et al*. (2011) quasi-mesenchymal subtype), pancreatic progenitor (Collisson *et al*. (2011) classical), aberrantly differentiated endocrine exocrine (ADEX) (Collisson *et al*. (2011) exocrine-like) plus an additional fourth novel subtype, immunogenic. Details of several genes involved in both PDAC and CRC are described in section 6.1.

The combination of these virtual, genetic and transcriptional analyses helps to build a deeper understanding of PC in its many guises and may aid future treatment choices for patients.

### 1.2.4 Diagnostic Tools and Biomarkers

One of the greatest challenges in reducing mortality rates in PC is in early diagnosis. At the point of diagnosis, 50-60% of patients have PC which is already metastatic and are not eligible for surgical intervention (Gillen *et al.*, 2010). Those who are diagnosed early and are able to undergo resection and/or chemotherapy have a much better prognosis with life expectancy rising from 5-9 months with no resection or chemotherapy to potentially two years (23.6 months) when receiving both (Gillen *et al.*, 2010). It has been suggested that there is more than a decade when PDAC is in a potentially curable stage which provides an ideal window for diagnosis (Yachida *et al.*, 2010). There also appears to be 6.8 years on average between the initial tumorous cell and the start of metastasis which would provide an additional time period for early diagnosis (Yachida *et al.*, 2010). Currently, however, patients are usually not being diagnosed until the final two years of tumorigenesis, after metastasis, when treatment options are severely limited. This illustrates the key role of early diagnosis to identify PC whilst it is still in a treatable state. PC also has rapid tumour progression, showing a need for swift treatment using more effective treatment options once diagnosis has occurred.

Blood tests may contain useful biomarkers which suggest the early signs of PC. Some biomarkers which may be present in PC patients include immunoreactive elastase (IRE), carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA 19-9), which remains to be the only PC biomarker approved for diagnosis (Hasan *et al.*, 2019). Although useful in diagnosis, these biomarkers are not particularly specific to PC and can be present in a range of other cancers and

some patients with PC will not have these biomarkers at all. Additionally, use of these biomarkers lack sensitivity, especially at the early stages of PC when it is crucial for an accurate diagnosis. Therefore, blood tests are used in conjunction with other tests to confirm whether it is likely to be PC.

In a trial conducted in 1994, IRE was present in 70% of PCs and CA 19-9 was present in 73%. IRE elevation was observed more in early, resectable head cancers whereas CA 19-9 was present more in body-tailed cancers and cancers that were unresectable (Nakae *et al.*, 1994). Another study concluded that IRE was not helpful in the diagnosis of PC as it did not distinguish between PC and chronic pancreatitis (Gullo *et al.*, 1989). Furthermore, it could not distinguish between early and advanced PC (Gullo *et al.*, 1989). IRE as a diagnostic marker for PC is not seen much in the literature after 1990 (Hayakawa *et al.*, 1990).

A single nucleotide polymorphism has been discovered to be present in the gene encoding the cholecystokinin (CCK) receptor, a G-protein coupled receptor involved in digestion. Normal function of the CCK protein stimulates gut motility and pancreatic enzyme secretion whilst inhibiting gastric acid secretion (Rehfeld, 2017). Expression of this mutated receptor appears to be specific to PC as control samples from healthy patients as well as patients with other GI cancers were compared with PC and only PC was positive for expression of the mutant receptor (Alsubai *et al.*, 2016, Nakae *et al.*, 1994).

Mothers against decapentaplegic homolog 4 (SMAD4) provides an interesting opportunity for diagnosis as inactivation of the SMAD4 gene appears to be specific for PC. As there is a strong relationship between SMAD4 gene status and Smad4 protein expression, monitoring levels of SMAD4 could provide an easy monitoring method which could also aid in determining whether a tumour in another organ is likely to be a pancreatic metastasis or a secondary tumour (Maitra *et al.*, 2006).

Short sections of non-coding RNA, known as microRNA (miRNA), have been found to have a significant role in various functions of the cell and they are thought to contribute to tumour formation when dysregulated (Bartel, 2004, Rachagani *et al.*, 2010). miRNAs have a role in some cancers by changing the epigenetic regulation of key cancer genes which may promote cell proliferation and tumour growth or reduce cell survival and therefore reduce tumour development (Rachagani *et al.*, 2010, Yanaihara *et al.*, 2006).

Some miRNAs found to be specific in PC may act as biomarkers for potential novel diagnostic tools, or as novel therapeutic options (Rachagani *et al.*, 2010). For example, miR-375 and miR-376 were both found to be expressed at a higher level in the pancreas and pancreatic islet cells

than they were in liver, brain or heart tissue in mice (Rachagani *et al.*, 2010). They may have a role in distinguishing between high- and low-risk neoplasms of the pancreas and thus aid in decisions about treatment. MiRNA profiles have already been identified and are being used for diagnosis and to predict prognosis in patients with lung cancer (Yanaihara *et al.*, 2006). Using miRNAs specific to PC for diagnosis purposes would greatly aid in the early treatment and better prognosis for PC patients. miRNA that was originally found only in PC tissue cells has since been detected in other bodily fluids including blood, serum, saliva and stool samples (Hernandez and Lucas, 2016). Used in combination with CA 19-9 antigen testing, miR-205, miR-210, miR-492 and miR-1427 testing show 91% sensitivity and 100% specificity for PDAC (Hernandez and Lucas, 2016).

Small extracellular vesicles known as exosomes containing proteins, nucleic acids and even miRNA are involved in cellular communication and have a role in cell proliferation, metastasis and angiogenesis in PC (Ariston Gabriel *et al.*, 2020). PC exosomes also express glypican-1 and the level of glypican-1 in patient tumours correlate with tumour burden and overall survival (Melo *et al.*, 2015). PC releases these exosomes into both the tumour microenvironment and the wider circulation, making them excellent candidates for diagnostic purposes and could help with the early detection of PC, enabling appropriate treatment to be expediated (Ariston Gabriel *et al.*, 2020).

### 1.2.5 Current Treatment

#### *1.2.5.1 Surgery*

Surgical resection is currently the only treatment with curative potential. Surgery for PDAC may involve removal of all or part of the pancreas, depending on the location and extent of the tumour. This, however, relies upon early presentation and diagnosis of disease and is therefore not an option for most patients (Brand, 2001). For patients whom surgical resection is appropriate, there is a very low actual cure rate of 3-5% at 5 years (Valsecchi *et al.*, 2014).

#### *1.2.5.2 Chemotherapy*

Current chemotherapeutic treatment for PC revolve around a central core of gemcitabine therapy which may also incorporate an additional agent or agents. Another option for patients, particularly those who are younger and/or fitter, is FOLFIRINOX therapy. This combination therapy provides a higher overall survival for PC patients but at the expense of higher toxicity with a considerably worse side effect profile than gemcitabine (Conroy *et al.*, 2011). The currently used chemotherapy drugs are listed in Table 1.1. These include gemcitabine, the current gold standard treatment option, as well as several combination therapies that have been approved for use in the treatment of advanced and metastatic PDAC.

#### *1.2.5.3 Gemcitabine*

Originally approved for treatment of advanced PC in 1997, gemcitabine (2',2'-difluoro-2'-deoxycytidine, or dFdC) works as an anti-metabolite and is a fluorinated analogue of deoxycytidine, a natural pyrimidine nucleoside of normal DNA (Mini *et al.*, 2006). As such, gemcitabine becomes incorporated into the growing DNA chain during DNA replication in place of deoxycytidine. Before this incorporation can occur, gemcitabine must be converted from its prodrug form to its active form. Firstly, gemcitabine is taken into the cell and is phosphorylated by deoxycytidine kinase to gemcitabine monophosphate. Gemcitabine monophosphate is then further phosphorylated to gemcitabine di-, then triphosphate which is its active drug form (Heinemann *et al.*, 1988). Once activated, gemcitabine can be incorporated into the cellular DNA by DNA polymerase, eventually leading to cell apoptosis. Crucially, DNA synthesis continues for at least one nucleotide after incorporation of gemcitabine making DNA polymerase unable to remove gemcitabine from the DNA chain (Huang *et al.*, 1991). This is in contrast to earlier anti-metabolites such as ara-CTP which terminates the DNA strand immediately after its incorporation but may consequently be removed from the DNA in 37% of cases (Huang *et al.*, 1991).

*Table 1.1. Currently used chemotherapeutic agents for the treatment of advanced and metastatic PC.*

*Table includes the year of approval, class of drug, trials and reference to the relevant trial. *FOLFIRINOX is a combination therapy containing 5-Fluorouracil, folinic acid, irinotecan and oxaliplatin.*

| Drug | Approval Year | Class | Reference |
|---|---|---|---|
| Gemcitabine | 1996 | Nucleoside analogue | (Burris *et al.*, 1997) |
| 5-Fluorouracil as adjuvant | 2004 | Cytotoxic | (Neoptolemos *et al.*, 2004) |
| Gemcitabine + cisplatin | 2006 | Platinum containing anticancer drug | (Heinemann *et al.*, 2006) |
| Gemcitabine + erlotinib | 2007 | Nucleoside analogue + tyrosine kinase inhibitor | (Moore *et al.*, 2007) |
| Gemcitabine as adjuvant | 2007 | Nucleoside analogue | (Oettle *et al.*, 2007) |
| Gemcitabine + capecitabine (GEM-CAP) | 2009 | Pro-drug of 5-fluorouracil | (Cunningham *et al.*, 2009) |
| FOLFIRINOX* | 2010 | Combination therapy | (Conroy *et al.*, 2011) |
| Gemcitabine + nab-paclitaxel | 2013 | Taxane - interferes with breakdown of microtubules during cell division | (Von Hoff *et al.*, 2013) |
| Nanoliposomal irinotecan + fluorouracil + folinic acid | 2015 | Topoisomerase 1 inhibitor | (Wang-Gillam *et al.*, 2015) |

#### 1.2.5.4 FOLFIRINOX

FOLFIRINOX is a combination therapy which consists of the administration of the four agents 5-fluorouracil, leucovorin, irinotecan and oxaliplatin which act synergistically as an effective treatment for PC. FOLFIRINOX has been shown to have a superior effect in the treatment of advanced PC, increasing median survival from 6.8 months in the gemcitabine treatment group to 11.1 months in the FOLFIRINOX group (Conroy *et al.*, 2011). Although a more effective therapeutic option for some patients, FOLFIRINOX was shown to have an increased toxicity profile over gemcitabine therapy including increased neutropenia (45.7% vs 21% gemcitabine), sensory neuropathy (9% vs 0%) and diarrhoea (12.7% vs 1.8%) (Conroy *et al.*, 2011). This restricts the administration of FOLFIRINOX to patients who are under the age of 76 and deemed healthy enough in terms of Eastern Cooperative Oncology Group (ECOG) score to receive this more toxic therapy. Unfortunately, due to the nature and often late diagnosis of PC, many patients are not eligible for FOLFIRINOX therapy.

#### 1.2.5.5 Gemcitabine plus additional agents

Numerous trials have taken place to examine the effect of adding additional therapeutic agents to gemcitabine therapy. Some of these additional agents include cisplatin (Heinemann *et al.*, 2006), oxaliplatin (Louvet *et al.*, 2005)**,** 5-fluorouracil (Berlin *et al.*, 2002), irinotecan (Rocha Lima *et al.*, 2004) and pemetrexed (Oettle, 2005), but none of these combinational therapies made a significant improvement on PC survival rates.

Gemcitabine plus capecitabine (GEM-CAP) went through phase III clinical trials in 2009. It was concluded that GEM-CAP provided a significant benefit to life expectancy and overall survival with few, controllable side effects (Cunningham *et al.*, 2009). It has been suggested that GEM-CAP should be considered a first line therapy for locally advanced and metastatic PC (Cunningham *et al.*, 2009).

#### 1.2.5.6 Erlotinib

Gemcitabine was trialled with and without the addition of erlotinib, an EGFR inhibitor (Moore *et al.*, 2007). Erlotinib is an EGFR tyrosine kinase inhibitor drug that has been approved for use in non-small cell lung cancer. It also has a licence approval for use in combination with gemcitabine in advanced-stage PC which has grown, spread or cannot be resected in patients who have not yet received chemotherapy (http://www.tarceva.com/patient/). Gemcitabine can help to overcome resistance to erlotinib by downregulation of Akt in EGFR over-expressing cells (Bartholomeusz *et al.*, 2011). When gemcitabine and erlotinib are combined, patients receive a 22% overall improvement in survival compared to patients who received gemcitabine alone (Moore *et al.*, 2007). Combination therapy provided a small improvement in life expectancy (6.24 months versus 5.91 months with gemcitabine alone) and an increase in one year survival

rate from 17% to 23% compared to gemcitabine alone (Moore *et al.*, 2007). The disadvantages to this combination therapy include a marginal increase in incidence of diarrhoea and appearance of skin rash when erlotinib is added to gemcitabine therapy. There is, however, a distinct cost barrier to the widespread use of this combination therapy with a dramatic reduction in cost or improvement in efficacy required before this becomes feasible as a commonly prescribed treatment option (Miksad *et al.*, 2007).

#### *1.2.5.7 Nab-paclitaxel*

Nab-paclitaxel consists of albumin-bound paclitaxel particles. In 2013, a regime containing gemcitabine plus nab-paclitaxel was trialled against gemcitabine alone in patients with metastatic PDAC (Goldstein *et al.*, 2015). Nab-paclitaxel plus gemcitabine shows significant improvement in survival when compared with gemcitabine alone, although side effects such as neuropathy and myelosuppression are worsened when these therapies are combined (Von Hoff *et al.*, 2013, Al-Hajeili *et al.*, 2014). This combination therapy also resulted in a 90% reduction in levels of CA 19-9 (Von Hoff *et al.*, 2013), which has been suggested to correlate with improvement in survival in PC (Reni *et al.*, 2009).

#### *1.2.5.8 Telomerase inhibitors*

Telomerase inhibiting drugs could be a potential therapeutic option for PC patients. Telomere shortening can be observed in cells undergoing replicative crisis whereby the telomeres shorten with each progressive replication, eventually leading to the inability of DNA polymerase to replicate the cells. Some malignant cells (and almost all immortal cells) can bypass this crisis state by producing telomerase, an enzyme able to catalyse the addition of telomeric sequences to chromosome ends. As such, this presents a potential route to develop anti-telomerase drugs for the treatment of cancer (Danovi *et al.*, 2008).

### 1.2.6 Failed Therapies

#### *1.2.6.1 Why are current treatments not enough?*

Surgery, gemcitabine and FOLFIRINOX make up the core trio of treatment options for the treatment of resectable, advanced and metastatic PC. Since its original approval in 1997, multiple chemotherapeutic agents have been added to gemcitabine therapy in attempts to improve the overall survival of patients. Most of these agents failed to significantly improve survival rates in patients over gemcitabine alone or were deemed to have higher than acceptable toxicity. The reason for failure for each combination is described below and summarised in Table 1.2.

Matrix metalloproteinases (MMPs) are proteolytic enzymes present in the cell which are thought to be involved in the degradation of extracellular matrix and tumour invasion when an

imbalance occurs between activated MMPs and their inhibitors. Expression of certain MMPs have been shown to be significantly higher in PC than in the healthy pancreas. As such, MMP inhibitors (MMPIs) have been developed to reduce tumour growth and metastasis. Bramhall *et al*. (2002) performed a phase II clinical trial on the use of gemcitabine plus marimastat, a potent MMP inhibitor, as a treatment for PC. Although the drug combination was well tolerated, it performed poorly in clinical trials and did not confer a survival benefit compared to gemcitabine alone.

Tipifarnib inhibits the action of farnesyl protein transferase which is thought to have a role in membrane anchorage of Ras proteins which is required for their activity. A phase III clinical trial was conducted comparing gemcitabine plus tipifarnib with gemcitabine alone (Van Cutsem *et al.*, 2004). Despite good activity against PC cell lines and acceptable toxicity, the trial indicated that gemcitabine plus tipifarnib does not significantly improve overall survival in advanced PC.

Two clinical trials investigated the addition of the COX-2 inhibitor celecoxib to gemcitabine therapy. COX-2 expression in PC is believed to be significant in tumour progression and resistance to other therapeutics. Therefore, the addition of a COX-2 inhibitor such as celecoxib was thought to improve the effects of existing therapies such as gemcitabine. Trials included the addition of celecoxib to gemcitabine (Dragovich *et al.*, 2008) and in addition to the existing combination of gemcitabine and cisplatin combination therapy (El-Rayes *et al.*, 2005).

Cilengitide targets integrins αvβ3, αvβ5 and α5β1 to produce an anti-angiogenic effect. It was combined with gemcitabine and reached phase II clinical trials against advanced, unresectable cancer. Despite being well tolerated with a good safety profile, the combination did not significantly affect markers of angiogenesis with no differences in efficacy of treatment or quality of life measures compared to gemcitabine monotherapy (Friess *et al.*, 2006).

Bevacizumab is a vascular endothelial growth factor (VEGF) inhibitor which reduces binding of VEGF to cell surface receptors, reducing angiogenesis and limiting blood supply to tumours. It was added as a combination therapy with gemcitabine for advanced PC and trialled in phase III clinical trials (Kindler *et al.*, 2007). The combination, however, failed to improve survival in these patients compared to the gemcitabine plus placebo control group. This combination was added to and a phase III trial of gemcitabine, bevacizumab and the tyrosine kinase inhibitor erlotinib was tested (Van Cutsem *et al.*, 2009). This combination also failed to lead to a significant survival improvement compared to gemcitabine monotherapy.

*Table 1.2. Failed PC treatments.*

*The table shows the drug/drug combination, what year the therapy failed, the class of drug, why it failed, what stage of development it was at during failure and a reference to study which decided on failure of the therapy in PC.*

| Drug | Year failed | Class | Reason for failure | Phase at failure | Reference |
|---|---|---|---|---|---|
| **Gemcitabine plus marimastat** | 2002 | MMP inhibitor | Performed poorly in clinical trials and was terminated | II | (Bramhall *et al.*, 2002) |
| **Gemcitabine + tipifarnib** | 2004 | Farnesyl protein transferase inhibitor | Does not significantly improve overall survival in advanced PC | III | (Van Cutsem *et al.*, 2004) |
| **Gemcitabine + cisplatin + celecoxib** | 2005 | Platinum co-ordination compound and COX-2 inhibitor | The addition of celecoxib did not improve the effect of gemcitabine + cisplatin | II | (El-Rayes *et al.*, 2005) |
| **Gemcitabine + cilengitide** | 2006 | Angiogenesis inhibitor | No clinically important differences in quality of life, efficacy or safety between active and placebo groups | II | (Friess *et al.*, 2006) |
| **Gemcitabine + Bevacizumab** | 2007 | Angiogenesis inhibitor | The addition of Bevacizumab to gemcitabine does not improve survival in advanced PC | III | (Kindler *et al.*, 2007) |
| **Gemcitabine + celecoxib** | 2008 | COX-2 inhibitor | Does not show significant improvement on survival of patients with metastatic PC | II | (Dragovich *et al.*, 2008) |
| **Gemcitabine + Bevacizumab + erlotinib** | 2009 | Angiogenesis inhibitor plus tyrosine kinase inhibitor | Does not lead to a significant improvement in overall survival | III | (Van Cutsem *et al.*, 2009) |
| **Gemcitabine + cetuximab** | 2010 | Anti-EGFR anti-body | Did not improve outcome compared to gemcitabine alone | III | (Philip *et al.*, 2010) |
| **Gemcitabine + axitinib** | 2011 | VEGF receptor inhibitor | Does not improve overall survival in advanced PC | III | (Kindler *et al.*, 2011) |
| **Cetuximab, Gemcitabine and radiation** | 2012 | Anti-EGFR anti-body | Significant toxicity observed in test population | I | (Chakravarthy *et al.*, 2012) |
| **Gemcitabine + sorafenib** | 2012 | G + Multi-kinase inhibitor (RAS/REF/MEK/ERK pathways) | Did not improve progression-free survival | III | (Gonçalves *et al.*, 2012) |
| **Gemcitabine + cisplatin + sorafenib** | 2014 | Multi-kinase inhibitor (RAS/REF/MEK/ERK pathways) | Shown to be ineffective in advanced PC | II | (Cascinu *et al.*, 2014) |
| **Gemcitabine + cixutumumab + erlotinib** | 2014 | EGFR inhibitor + IFG-1R inhibitor | Shown to be ineffective in advanced PC | Ib and II | (Philip *et al.*, 2014) |
| **Gemcitabine + Ganitumab** | 2015 | Fully human monoclonal antibody that targets IGF-1 and -2 | Does not significantly improve overall survival in advanced PC | III | (Fuchs *et al.*, 2015) |
| **Gemcitabine + Evofosfamide** | 2016 | DNA alkylating agent | Does not improve overall survival in advanced PC | III | (Van Cutsem *et al.*, 2016) |

A combination of gemcitabine plus the VEGF receptor (VEGFR) inhibitor axitinib was trialled in patients with advanced PC. This combination did not improve overall survival in patients and adds weight to the argument that targeting VEGFR is not an effective strategy in PC (Kindler *et al.*, 2011).

There is some debate about targeting members of the ErbB family of tyrosine kinases. Lapatinib reversibly inhibits HER2 and, less potently, EGFR. It is used in breast cancer but has a potential application in EGFR based PCs (Walsh *et al.*, 2013). Conversely Safran *et al*. (2011) states that lapatinib has no value in PC as it is not effective and future trials of HER2 inhibitors are not warranted for PC. Gemcitabine plus the anti-EGFR antibody cetuximab was trialled in phase III trials against gemcitabine alone (Philip *et al.*, 2010). This combination did not improve clinical outcome compared to gemcitabine alone. This combination later had a phase I clinical trial with the addition of radiation based on work by Buchsbaum *et al.* (2002), however significant toxicity was observed in the test population and trials of this combination did not progress further (Chakravarthy *et al.*, 2012). The seemingly low efficacy of these ErbB receptor tyrosine kinase inhibitors may be due to the different expression levels of these tyrosine kinases in different subtypes of PC detailed above in section 1.2.3. This could lead to a low overall efficacy in PC, but with a much higher efficacy in subtypes of PC specifically overexpressing members of the ErbB family.

Sorafenib is a multi-kinase inhibitor and targets the RAS/REF/MEK/ERK pathway, specifically inhibiting RAF-1 kinase and the VEGF-2 receptor (Cascinu *et al.*, 2014, Siu *et al.*, 2006). Despite showing a good safety profile in phase I trials (Siu *et al.*, 2006), phase II and III trails had disappointing results as the combination of gemcitabine and sorafenib (with and without the addition of cisplatin) were shown to be ineffective in advanced PC (Cascinu *et al.*, 2014, Kindler *et al.*, 2012, Gonçalves *et al.*, 2012).

Cixutumumab was combined with gemcitabine and erlotinib to test whether adding an insulin like growth-factor (IGF)-1R inhibitor to an EGFR inhibitor (erlotinib) and the gold standard gemcitabine improved overall survival in patients with metastatic PC (Philip *et al.*, 2014). Again, there were disappointing results with the combination failing to show an improvement in advanced PC.

Ganitumab, a fully human monoclonal antibody that targets IGF-1 and -2, was combined with gemcitabine and trialled in patients with metastatic PC. In phase III clinical trials it was concluded that the therapy did not provide significant improvement in overall survival compared to

gemcitabine alone and it was further suggested that use of IGF1R targeted therapies were not recommended for treatment of metastatic PC (Fuchs *et al.*, 2015).

Evofosfamide is a prodrug which becomes activated to the DNA alkylating agent bromo isophosphoramide mustard under conditions of hypoxia such as within pancreatic tumours (Duan *et al.*, 2008). Despite being granted FDA Fast Track designation in May 2015, Evofosfamide failed to improve overall survival in patients with advanced PC in combination with gemcitabine at Phase III clinical trials (Van Cutsem *et al.*, 2016).

A study suggested that use of ACE inhibitors (ACEIs) and angiotensin II receptor blockers (ARBs) in combination with gemcitabine could have a benefit in PC (Nakai *et al.*, 2010), however further studies found that these agents did not have a significant impact on overall survival (Winston *et al.*, 2015). There is still some uncertainty surrounding the issue as another study found that calcium channel blockers (CCBs) combined with aspirin, but not ACEIs or ARBs following surgery can have a potent anti-stromal effect which translates to a survival benefit in PC patients, giving 1414 days median survival compared to 528 days for those not taking either drug (Tingle *et al.*, 2015). The effect of CCBs may be due to reduction of effects of CCK hormones on pancreatic cells, resulting in decreased carcinogenesis and metastasis.

Gemcitabine plus masitinib, a kinase inhibitor which selectively targets macrophages and mast cells was shown to have a generally acceptable safety profile and confer a significant survival benefit over gemcitabine monotherapy in patient populations with an overexpression of acyl-CoA oxidase-1 and in patients suffering from pain associated with their cancer. However the combination was not recommended for patients without pain or who did not have the deleterious genomic biomarker that conferred a survival advantage (Deplanque *et al.*, 2013, Deplanque *et al.*, 2015).

A meta-analysis of 9 randomised phase III clinical studies compared the use of gemcitabine plus a targeted therapy to gemcitabine mono therapy. The comparison included the addition of erlotinib, cetuximab, rigosertib, elpamotide, bevacizumab, aflibercept, axitinib, masitinib and ganitumab to gemcitabine therapy. The meta-analysis concluded that there was no significant improvement in survival of patients given the additional targeted therapeutic compared to gemcitabine alone in any of the presented studies (Ottaiano *et al.*, 2017).

The reasons for these failures are varied and some are unknown. Some possible reasons include difficulty in delivery of the agents to the actual pancreatic tumours. This may be due to the poorly vascularised nature and/or hypoxic environment of pancreatic tumours. Although there has been some progress in trialling novel therapies for PC, a lot of these therapeutic options are

based on combination therapies using other therapeutic agents designed for other indications or unrelated cancers and not specifically for the intricate environment presented by pancreatic tumours. There is an urgent unmet need for novel therapeutic options that consider the varying PC subtypes shown to be present amongst the population. This combined with more effective diagnosis and swift commencement of treatment is essential to provide a better clinical outcome for patients diagnosed with PC.

### 1.2.7 Drug Resistance

Drug resistance is a lack of response to therapy, either immediately or after an initial period of positive response, which eventually leads to failure of treatment (Michaelis *et al.*, 2019). This enforces an upper limit to the dosage able to be given to patients without the emergence of intolerable side-effects or acute toxicity to the patient themselves (Fenton *et al.*, 2018).

As illustrated in Table 1.1 and Table 1.2, most current and attempted drug therapies for PC rely heavily on the inclusion of gemcitabine. There is, however, evidence that some Cathepsin-D (CatD) overexpressing PDACs show resistance to gemcitabine therapy and patients showing high CatD expression have a much poorer prognosis that those with normal or low CatD levels (Mahajan *et al.*, 2020). Despite being nano-sized and bound to albumin, there is also evidence to suggest acquired drug resistance emerging to nab-paclitaxel in various cancer types (Vallo *et al.*, 2017, Zhao *et al.*, 2015).

Drug resistances such as these illustrate the requirement for novel therapies with an alternative mechanism of action as complementary therapies for the treatment of PC. The following section explores the utility of animal venoms and their components as drug treatments and potential candidates for various types of cancer.

## 1.3 Venom

### 1.3.1 What is Venom?

Animal venoms are a multifunctional mixture of different components that act in a synergistic manner to subdue prey or defend against potential predators and rivals. Venom and venom administering apparatus have evolved independently multiple times over millions of years in many different phyla in response to predation and to aid in prey capture and digestion (Oldrati *et al.*, 2016). Multiple convergent evolution episodes give rise to a vast and diverse range of different components found across venomous species. This gives rise to a vast, largely unexplored library of pharmacologically active molecules which have evolved to be bioactive and have the potential to be harnessed as therapies for a range of human disease (Escoubas and King, 2009).

Diet of individuals has been found to have an impact on venom diversity with members of closely related subspecies having quite different venom components due to a divergence of prey (Sanz *et al.*, 2006). In addition to interspecies diversity, venom may differ even within one species. This may be due to various ecological and environmental factors, such as geographical isolation (Senji Laxme *et al.*, 2021) or may be due to factors such as age or sexual dimorphism (Antunes *et al.*, 2010). There are even differences documented between litter mates and between the venom of hatchlings compared to after their first shed (Wray *et al.*, 2015). It is therefore important in antivenom production (and in venom research) to select individuals of various ages, sexes and from various locations to ensure that pooled venom represents all the components that may be present within venom from that species.

Snake venoms, along with arachnid and cone snails, are amongst some of the most ancient and best characterised animal venoms with an abundance of literature available on both the evolution of different venom components and studies on their pharmacological effects (Oldrati *et al.*, 2016). This is due to their large size (and thus plentiful venom yields) combined with the significant human health burden that can result from the serious effects of their bite (Oldrati *et al.*, 2016).

### 1.3.2 Venom Components

Venom components can be broken down into three main categories: proteins, low molecular mass organic compounds and inorganic compounds. Up to 95% of the dry weight of venom is thought to be made up of proteins and peptides, with other components being minor constituents (Liu *et al.*, 2014).

The protein category consists of enzymatic and non-enzymatic proteins, enzymatic inhibitors and peptides. Enzymatic proteins can be further split into oxidases and hydrolases. The only

oxidases found in venoms thus far are L-amino acid oxidases (LAAOs) whilst hydrolases are much more abundant in snake venom and include phospholipases, proteases, peptidases, acetylcholinesterases and hyaluronidases (Utkin, 2015).

Non-enzymatic proteins include three finger toxins, disintegrins, proteinase inhibitors, C-type lectins, nerve growth factors, vascular endothelial growth factors, cysteine-rich secretory proteins (CRISPs), amongst others. These non-enzymatic proteins are discussed in detail by McCleary and Kini (2013).

Peptides and polypeptides tend to be found in greater abundance in invertebrate venoms, such as spider, scorpion and cone snail and may include antimicrobial peptides, neurotoxins and immune-modulating peptides. Many of these peptides modulate ion channels and have been used widely in the study of ion-channel function (Kalia *et al.*, 2015).

Non-protein components include salts, alkaloids, polysaccharides, terpenes, organic acids and biogenic amines, free nucleosides. magnesium, cobalt, iron, and potassium, amongst other components (Ramos and Selistre-De-Araujo, 2006). A full breakdown of snake venom components is detailed in Figure 1.1.

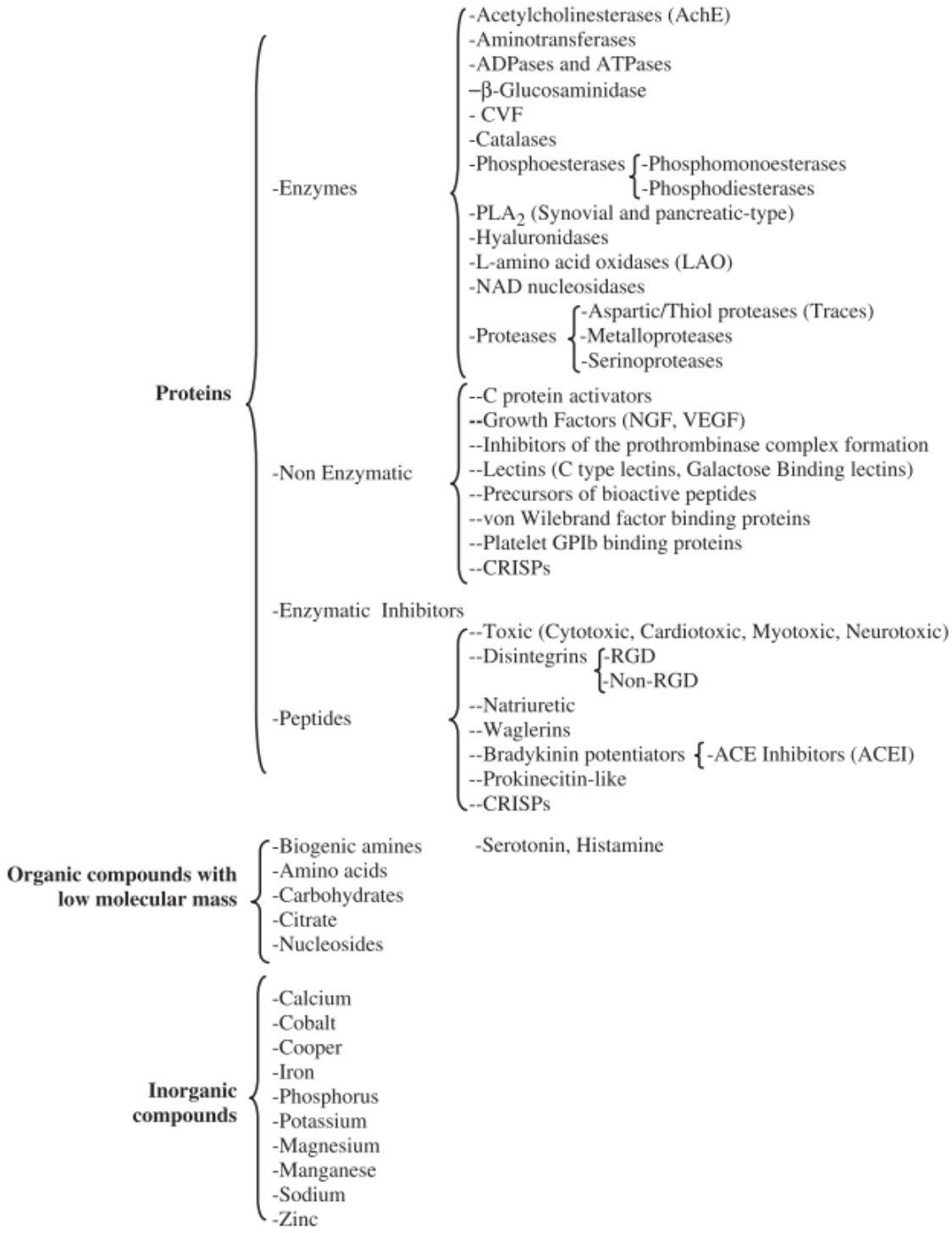



*Figure 1.1. An overview of the components found in snake venoms.*

*Figure reprinted from Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, Vol 142 Issues 3-4, O.H.P.Ramos and H.S.Selistre-de-Araujo, Snake venom metalloproteases — structure and function of catalytic and disintegrin domains , P328-346, Copyright (2006), with permission from Elsevier. (Ramos and Selistre-De-Araujo, 2006)*

### 1.3.3 Venoms as Drugs

The huge diversity of bioactive molecules present in venoms provide an enormous library of potential drug-like molecules. Venoms and venom components which have gained market authorisation as drugs fall into three main categories - cardiovascular, pain and type II diabetes treatments. These categories are briefly discussed below:

#### *1.3.3.1 Cardiovascular*

The first major success story for venom-based drug discovery was Captopril, the world's first angiotensin converting enzyme (ACE) inhibitor. ACE inhibitors are now a commonly prescribed drug class which effectively reduce blood pressure by regulating the renin-angiotensin system (RAS), preventing the conversion of angiotensin I to angiotensin II by ACE. Angiotensin I, an inactive precursor molecule which has no intrinsic biological activity but is converted to angiotensin II. Angiotensin II increases aldosterone production, increasing sodium retention in the kidneys, increases vasopressin production and causes vasoconstriction potentially leading to hypertension when the RAS system is over activated (Rang *et al.*, 2008).

Captopril is a synthetic inhibitor drug molecule based on short peptides found in *Bothrops jararaca* (Brazilian viper) venom which potentiate bradykinin activity, acting as a potent inhibitor of ACE (Bakhle, 2020). The viper peptide was shown to reduce the conversion of angiotensin I to angiotensin II via inhibition of ACE, and as a result reduce hypertension in test animals and eventually humans. This was then developed into an orally bioavailable form of the viper peptide which can be given to patients in the form of a tablet rather than injection (Smith and Vane, 2003). Since the development and approval of captopril in 1981, many improved ACE inhibitors including ramipril, enalapril and lisinopril have been developed and are still prescribed extensively for hypertension and cardiac failure today.

Integrins play an important role in the blood clotting process by acting as surface receptors, recruiting fibrinogen when a platelet encounters a damaged blood vessel. Fibrinogen is then converted to fibrin and encourages aggregation of platelets, forming clots. Eptifibatide (Integrilin®) was discovered as part of a drug discovery campaign by Scarborough *et al*. (1991) aimed to discover a disintegrin that specifically targeted the platelet aggregating glycoprotein IIb-IIIa complex but not other integrins. Following a screen of 62 snake venoms, barbourin was discovered in the venom of *Sistrurus miliarius barbouri*. Eptifibatide was developed and optimised from this lead molecule. Eptifibatide bind specifically to the glycoprotein IIb-IIIa integrin receptor and blocks the binding of fibrinogen, reducing the localised coagulation of blood. Tirofiban (Aggrastat®) also acts as a glycoprotein IIb-IIIa receptor antagonist to reduce

platelet-mediated blood clotting and was developed from echistatin from *Echis carInatus* venom (Koh and Kini, 2012).

Snake venoms are well known for containing a wealth of both pro- and anti-coagulant components used as key mechanisms in prey subjugation. Both properties have been utilised throughout drug discovery for various applications, for example with Batroxobin (Defibrase® or Reptilase®) and Ancrod (Viprinex™). Ancrod was isolated from *Agkistrodon* (now *Calloselasma*) *rhodostoma* venom whilst Batroxobin was isolated from the venom of *Bothrops atrox*. Both molecules are hemotoxic serine protease molecules with similar function to thrombin as a defibrinogenating agent. Both molecules cleave fibrinogen, reducing its circulating concentration, reducing the formation of insoluble clots whilst also catalysing soluble clot formation which may be broken down by plasmin (Koh and Kini, 2012). Since 2017, Ancrod was no longer marketed for any clinical use. Defibrase® only has market authorisation in China and Japan whilst Reptilase® has market authorisation in Japan, Korea and India for various emergency bleeding and both internal and external haemorrhage. Reptilase® is also used internationally to determine 'Reptilase time' - a measure of coagulation time which can help to determine the presence or absence of heparin in a clinical sample (Karapetian, 2013).

#### *1.3.3.2 Pain*

Ziconotide is the first commercially available drug product based on ω-conotoxin peptide from *Conus magus* (cone snail) venom. Many cone snail venom peptides have analgesic properties as they have been shown to interact with voltage gated ion channels such as sodium, potassium and calcium, as well as nicotinic acetylcholine receptors. This gives a distinct mechanism of action compared to opioids which are traditionally given to relieve severe pain such as cancer or neuropathic pain, as well as an improved safety profile with less risk of side effects and addiction generally observed with opioid use. Similar cone snail peptides are undergoing clinical trials for treatment of neuropathic pain with Conantokin-G additionally being trialled for treatment of epilepsy and κ-Conotoxin PVIIA in preclinical trials for cardio protection against myocardial infarction (Gorson and Holford, 2016).

#### *1.3.3.3 Type II Diabetes*

Exenatide is a synthetic peptide based on the hormone Exendin-4 from the saliva of *Heloderma suspectum* (Gila Monster). The drug acts in a similar manner to the human hormone glucagon-like peptide-1 (GLP-1). GLP-1 is secreted in response to food intake and stimulates the production of insulin. In type II diabetes, GLP-1 secretion is often impaired but the response to GLP-1 is maintained, making GLP-1 receptors an attractive target for drug intervention (Werner *et al.*, 2010). Exenatide is able to mimic the activity of the intrinsic hormone by acting on

pancreatic GLP-1 receptors to increase insulin production and reduce glucagon activity, but has a significantly improved half-life compared to native human GLP-1 (Cvetkovic and Plosker, 2007). Exanatide is available as two formulations - twice-daily administered Byetta® and the once-weekly Bydureon®. Both formulations are administered as subcutaneous injections.

Lixisenatide is another synthetic peptide based on the hormone Exendin-4 isolated from *Heloderma suspectum* saliva and also acts as a GLP-1 agonist to preserve insulin secretion and the function of pancreatic β-cells (Werner *et al.*, 2010). Lixisenatide also reduces food transit rate, decreasing absorption of glucose and appetite, leading to weight loss. A summary of the currently approved pharmaceutical products derived from venom components is shown in Table 1.3.

## 1.3.4 Venoms Components as Cancer Treatments

Venoms make attractive potential cancer therapies for several reasons. Many venoms are generally cytotoxic and some have increased toxicity to specific tumour cells compared to normal, non-cancerous cell types. They may lead to a reduction of tumour cell invasiveness, metastasis, cell migration and proliferation. They may also lead to the induction of apoptosis in tumour cells, reducing their numbers and preventing them from reproducing (Vyas *et al.*, 2013).

The below paragraphs detail some of the major components of snake venoms and how they have been studied for their anti-tumorigenic properties in the literature.

### *1.3.4.1 L-amino acid oxidases*

L-amino acid oxidases (LAAO) are enzymatic proteins found across many different phyla including both land and marine animals, plants, fungi and bacteria. Snake venoms provide a rich source of LAAOs which have been studied extensively. LAAOs catalyse oxidative deamination of L-amino acids, leading to the formation of hydrogen peroxide ($H_2O_2$), ammonia and α-ketoacids (Rodrigues *et al.*, 2009). LAAOs inhibit platelet aggregation and have anti-protozoal activity and apoptotic properties (Franca *et al.*, 2007, Sakurai *et al.*, 2001, Samel *et al.*, 2006).

The effect of LAAO from *Calloselasma rhodostoma* snake venom (CR-LAAO) on the myeloproliferative neoplasm model cell lines HEL 92.1.7 and SET-2 was studied (Tavares *et al.*, 2016). CR-LAAO appeared to have cytotoxic activity on these JAK2-mutated cell lines with $IC_{50}$ values of 0.15 µg/mL from HEL 92.1.7 cells and 1.5 µg/mL for SET-2 cells. The production of $H_2O_2$ was thought to be the cause of cytotoxicity and this was confirmed by mitigation of toxicity upon treatment with catalase. CR-LAAO also stimulated apoptosis in both cell lines in a concentration dependent manner, confirmed by caspase-3 and -8 activation and cleaved Poly(ADP-ribose) polymerase (PARP).

*Table 1.3. Current FDA approved venom derived pharmaceuticals.*

*Approval years are FDA approval unless otherwise stated. FDA = U.S. Food and Drug Administration, EMA = European Medicines Agency.*

| Generic Name | Brand Name | Species | Drug Class | Disease(s) / Use(s) | Approval Year |
|---|---|---|---|---|---|
| Captopril | Capoten® (Bristol-Myers Squibb) | *Bothrops jararaca* (Lancehead viper) | ACE inhibitor | Hypertension, Congestive heart failure | 1981 |
| Eptifibatide | Integrilin® (Millennium Pharmaceuticals, Inc.) | *Sistrurus miliarius* (Southeastern Pygmy Rattlesnake) | Antiplatelet (glycoprotein IIb/IIIa inhibitor) | Acute coronary syndrome | 1998 |
| Tirofiban | Aggrastat® (Medicure International, Inc.) | *Echis carinatus (*Saw-scaled viper) | Antiplatelet (glycoprotein IIb/IIIa inhibitor) | Non-ST elevation acute coronary syndrome | 1998 |
| Ziconotide | Prialt® (Elan Pharmaceuticals, Inc.) | *Conus magus* (Cone Snail) | Analgesic (ω-conotoxin peptide) | Pain | 2004 |
| Exanatide | Byetta® (AstraZeneca) | *Heloderma suspectum (*Gila Monster) | GLP-1 analogue | Type II diabetes | 2005 |
| | Bydureon® (extended release) (AstraZeneca) | | | | 2011 (EMA) 2012 (FDA) |
| Lixisenatide | Lyxumia® (Europe) Adlyxin® (USA) (Sanofi S.A.) | *Heloderma suspectum (*Gila Monster) | GLP-1 receptor agonist | Type II diabetes | 2013 (EMA) 2016 (FDA) |
| Ancrod | Viprinex™ (Neurobiological Technologies, Inc) | *Calloselasma rhodostoma* (Malayan Pit Viper) | Anticoagulant | Acute ischaemic stroke | No longer marketed |
| Batroxobin | Defibrase ® (Tobishi Pharmaceutical, China and Pentapharm, Switzerland) | *Bothrops moojeni* (Brazillian Lancehead) | Defibrinogenating agent | Thrombosis | |
| | Reptilase ® (Tobishi Pharmaceutical, China) | *Bothrops atrax* (Common Lancehead) | Haemostatic/Coagulant | Haemorrhage, fibrinogen deficiency test | |

*Bothrops leucurus* venom (*Bl*-LAAO) produce enough $H_2O_2$ in cell culture medium to induce apoptosis in MKN-45 (stomach cancer), HUTU (adenocarcinoma), RKO (colorectal cancer) and LL-24 (human fibroblast) cell types (Naumann *et al.*, 2011). BI-LAAO also activates caspase 3, 8 and 9 and causes cytotoxicity mediated by $H_2O_2$ production in chronic myeloid leukaemia (CML) cell lines (Burin *et al.*, 2016). LAAO from *Bothrops moojeni* snake venom (BmooLAAO-I) was found to have selective antitumorigenic activity with concentration dependent reactive oxygen species (ROS) produced in leukemic cells, but not non-tumorous HEK cells (Burin *et al.*, 2020). A decrease in the anti-apoptotic Bcl-2 protein was also observed, suggesting increased sensitivity of these cells to apoptosis (Burin *et al.*, 2020).

Purified ACTX-8, an LAAO from *Agkistrodon acutus* snake venom, shows effects on the mitochondrial pathway, including the release of cytochrome C into the cytosol and mitochondrial membrane potential (MMP) dissipation (Zhang and Wei, 2007). Rusvinoxidase, an LAAO isolated from *Daboia russelii russelii* venom, causes changes in cell size and morphology, membrane integrity and DNA fragmentation in MCF-7 breast cancer cells (Mukherjee *et al.*, 2015). Rusvinoxidase also appears to affect the extrinsic signalling pathway by activating caspase-8 which goes onto activate caspase-7 (but not caspase-3) and upregulates production of pro-apoptotic protein Bax whilst downregulating antiapoptotic proteins such as Bcl-XL (Mukherjee *et al.*, 2015).

#### *1.3.4.2 Phospholipase $A_2$*

Phospholipase $A_2$ ($PLA_2$) enzymes cause hydrolysis of phospholipids and have essential and varied biological functions. Mammalian $PLA_2$ enzymes have important roles in cell proliferation, smooth muscle contraction, maintenance of phospholipids and membrane repair, but are also attributed to some diseases such as chronic inflammation diseases and hypersensitisation (Kini, 2003). Snake venom $PLA_2$s, unlike regular mammalian $PLA_2$s, have evolved a highly toxic function.

Many venom components have been shown to act synergistically to potentiate the effect of other venom components. Isolated $PLA_2$ from *Cerastes cerastes* Tunisian snake venom was found to have anti-angiogenic properties by inhibiting the effects of integrins and demonstrating dose-dependant inhibition of cell adhesion and migration (Kessentini-Zouari *et al.*, 2010). $PLA_2$ isolated from *Macrovipera lebetina transmediterranea* inhibited cell adhesion and migration of a range of human tumour cells without causing cytotoxicity up to 2 µM concentration (Bazaa *et al.*, 2009). $PLA_2$ isolated from *Daboia russelii siamensis* venom inhibited the migration of SK-MEL-28 skin melanoma cells and provided a 65% inhibition of colonisation of B16F10 murine

skin melanoma cells in the lungs of BALB/c mice (Khunsap, 2011). A catalytically inactive *Bothrops jararacussu* $PLA_2$ homologue, BthTX-I has also been isolated (Gebrim *et al.*, 2009).

When it comes to cobra venom, there appears to be a distinct difference in the abundance of $PLA_2$ depending on the geographical location and behaviour of the cobra species. $PLA_2$ has been shown to potentiate the effect of cytotoxins (CTs) to increase sensory neuron activation and lead to increased pain (Kazandjian *et al.*, 2021). The emergence of spitting behaviour in African *Naja* (*Afronaja*) species and Asiatic *Naja* (*Naja*) lead a high abundance of enzymatically active $PLA_2$, suggesting the function of a mixture of $PLA_2$ and CTs in defence (Kazandjian *et al.*, 2021). The African non-spitting cobra species, however, differ in their $PLA_2$ abundance with the forest inhabiting *Naja* (*Boulengerina*) continuing to provide a rich supply of $PLA_2$ in their venom, whereas the plains species of African *Naja* (*Uraeus*) appear to have lost the majority of the $PLA_2$ in their venoms, correlating with the lower reported cytotoxicity and increased neurotoxicity of venom from cobras from these species (Tan *et al.*, 2019).

#### *1.3.4.3 Non-enzymatic proteins*

##### 1.3.4.3.1 Cytotoxins

Cobra cytotoxins/cardiotoxins (CTs) are 60 amino acid highly conserved proteins that form part of the 'three-finger toxin' superfamily (Konshina *et al.*, 2017). Three-finger toxins are characterised by their characteristic loops ("fingers") formed by anti-parallel β-strands extending from the core (Kini and Doley, 2010, Konshina *et al.*, 2017). Cardiotoxins specifically are only found in the venom of elapid snakes from the *Naja* and *Hemachatus* species (Breckeridge and Dufton, 1987). The structure of cytotoxins arises from eight conserved cysteine residues forming four disulphide bonds (between C1-C3, C2-C4, C5-C6 and C7-C8) which stabilise the hydrophobic centre (Kessler *et al.*, 2017, Reeks *et al.*, 2015).

CTs were first documented as early as 1905 as being components unique amongst elapid venoms for having cardiotoxic properties with attempts to isolate these uncharacterised components beginning as early as 1942 and on through the 1960's (Dufton and Hider, 1988). Due to their potentially devastating effects on human health, the evolution of structure and function of CTs have been studied throughout every decade since then, continuing to the present day.

Cytotoxicity of cobra venom appears to have evolved primarily as a defensive mechanism as opposed to prey subjugation, for which cobras have already evolved venom with abundant neurotoxins extremely effective for this purpose (Panagides *et al.*, 2017). These cytotoxins have co-evolved to supplement their camouflaging colouration and defensive hooding behaviour, evolved to deter potential predators from attacking (Panagides *et al.*, 2017). Following transcriptomic analysis, three-finger toxins (3FTx) were found to be the most abundant toxin

family in the venom of cobra species with $PLA_2$s being the second most abundant with spitting cobra lineages differing significantly from the non-spitting lineages (Kazandjian *et al.*, 2021).

Although the mechanism of action of CTs has not yet been elucidated, they display a cytotoxic effect on many cell types including cardiomyocytes, lymphocytes, erythrocytes and spleen cells, as well as various tumour types (Feofanov *et al.*, 2005). Lysosomes also appeared to be a primary target for the CTs and their accumulation in lysosomes correlated with their cytotoxic effect (Feofanov *et al.*, 2005).

There is also a body of evidence to suggest that CTs exert some effects on mitochondrial membranes. Konshina *et al.* (2011) used CT4 from *Naja kaouthia* venom to investigate the relationship between CTs and phosphatidylserine (PS). They found that the membrane damaging effect of CTs was increased when zwitterionic liposomes had PS incorporated into them. Zhang *et al*. (2019a) used cardiotoxin VII4 from the venom of *Naja mossambica* to investigate the effect of CTs on mitochondrial membranes. They found that CTs selectively migrated to mitochondria of their test cells, resulting in fragmentation and decrease of energy production. They also found that the studied cardiotoxin binds to cardiolipin to form non-bilayer structures in mitochondrial membranes but did not bind to phosphatidylcholine. Cytotoxins I and II were also found to disrupt mitochondrial membrane integrity, forming abnormal non-bilayer structures (Gasanov *et al.*, 2015).

#### 1.3.4.3.2 Disintegrins

Saxatilin, a disintegrin isolated from *Gloydius saxatile* venom inhibited cell invasion in a dose dependent manner by reducing tumour necrosis factor alpha (TNF-α) and reducing MMP-9 activity in MDAH 2774 ovarian cancer cells (Kim *et al.*, 2007). Saxatilin was also found to reduce angiogenesis due to suppression of VEGF via the Akt pathway in NCI-H460 lung cancer cells (Jang *et al.*, 2007).

Tzabcanin, a disintegrin purified from *Crotalus simus tzabcan* (middle American rattlesnake) had cytotoxic effects against A375 melanoma cells but not A549 lung cancer cells and reduced the migration and invasiveness of both tumour cell lines by inhibiting the integrin $\alpha_v\beta_3$. It could also be used as a diagnostic tool for detecting cancers that overexpress $\alpha_v\beta_3$ (Saviola *et al.*, 2016) .

A recombinant version of Concortrostatin showed anti-adhesive and anti-angiogenic properties equivalent to native Concortrostatin, a disintegrin isolated from *Agkistrodon contortrix contortrix* (Southern Copperhead) venom (Minea *et al.*, 2005). R-viridistatin 2, a recombinant disintegrin from *Crotalus viridis viridis* (Prairie rattlesnake) venom metalloproteinase II was able to inhibit cell adhesion, migration and invasive ability of a variety of human carcinoma cell lines

(Lucena *et al.*, 2012). The ability to produce large quantities of bio-equivalent product is an essential step to potentially translate a naturally produced compound such as Concortrostatin from the lab to the clinic.

### 1.3.5 Cobra Venom

There are a wide variety of cobra species diverging from a common ancestor which provides a wide variety of subtlety different venom components. These components may be compared from one species to another to determine the subtle changes in protein sequence between closely related species which give rise to large changes in activity. There is a relative abundance of studies already conducted using cobra venoms testing their propensity to cause genetic and morphological changes in cancer cells. There is also a rich source of structural information already available about some cobra venom proteins which would make any structural or sequence comparisons between different species easier with 40% of all published three-finger snake toxin sequencing originating from species from the *Naja* (cobra) genus (Gorson and Holford, 2016).

Figure 1.2 shows a phylogenetic tree detailing the evolution of members of the *Naja* genus with an approximate timeline of evolution. Cobra species may be divided into four subgenera depending on their habitat and classification as either 'spitting' or 'non-spitting' species. These subgenera include Asian cobras (*Naja*), non-spitting cobras populating the open formations and savannas in Africa (*Uraeus*), African forest inhabiting species (*Boulengerina*), and the newly appointed subgenera *Afronaja* from African spitting cobras (Wallach *et al.*, 2009).

The projection of venom ("spitting") has evolved independently three times within closely related elapid snakes including the *Afronaja* (African spitting) and *Naja* (Asian spitting) subgenera of cobras, and in the *Hemachatus* (rinkhals) (Kazandjian *et al.*, 2021). Below are some literature sources from other laboratories also studying the effect of cobra venom in cancer cells. All species of cobra have venom dominated by 3FTX, with many species also having an abundance of $PLA_2$ in their venom (Kazandjian *et al.*, 2021).

The effect of *Naja oxiana* venom toxins on HepG2, HCF7 and DU145 cancer cell lines and the non-cancerous MDCK normal cell line was investigated (Ebrahim *et al.*, 2016). Condensation of chromatin, cytoplasm blebbing and irregularly shaped cells were observed following treatment with the venom. Caspase 3 was activated at higher doses, leading to the initiation of apoptosis. Cytotoxins I and II from *Naja oxiana* venom were investigated and found to induce apoptosis via lysosomal damage and through significant activation of Caspase-3. When the dose is increased, however, the mechanism of action rapidly changes to a necrotic pathway and apoptosis is no longer observed (Ebrahim *et al.*, 2015, Feofanov *et al.*, 2004).

A CT isolated from *Naja kaouthia* venom was shown to induce apoptosis in leukaemic cells mediated by caspase 3 and 9 with an increased BAX-BCl-2 ratio, indicating a shift towards initiation of apoptosis. The CT showed selectivity in the tumour cells with low cytotoxicity in normal leukocyte cells (Debnath *et al.*, 2010).

The effect of NN-32 toxin from *Naja naja* venom on MCF-7 and MDA-MB-231 breast cancer cells was investigated. MTT was used to monitor cell cytotoxicity and cytotoxic effects were observed in both breast cancer cell lines ($IC_{50}$ values of 2.5 and 6.7 µg/ml for MCF-7 and MDA-MB-231 cell lines respectively) but without causing significant cytotoxicity in MCF-10A normal breast epithelial cells. This selective toxic effect was confirmed using anti-proliferative, LDH release and NR uptake assays (Attarde and Pandit, 2017).

CTXIII isolated from *Naja atra* venom was found to induce apoptosis in A549 adenocarcinoma cells through the activation of Caspase 3 and 9, loss of mitochondrial membrane potential and phosphatidylserine externalization (Su *et al.*, 2010). Upregulation of the apoptotic proteins BAX and Bad with a concomitant downregulation of anti-apoptotic proteins including BCl-2 was observed. CTX III was found to affect many genes within the EGFR cascade such as the PI3K/AKT and JAK2/STAT3 pathways, as well as reducing phosphorylation of EGFR itself (Su *et al.*, 2010). CTX III from *Naja atra* was also found to reduce cell invasion and migration in MDA-MB-231 breast cancer cells without cytotoxicity mediated by activation of matrix metalloprotease 9 (MMP-9) with concurrent activation of the downstream EGFR genes ERK 1/ 2 and PI3K/Akt. Interestingly, treatment with an EGFR inhibitor reversed effects of MMP-9, PI3K/Akt and ERK1/ 2 activation, suggesting the anti-metastatic properties of CTX III are exerted through the activation of EGFR and its downstream effectors (Tsai *et al.*, 2012). Phospholipase $A_2$ also isolated from *Naja atra* venom was found to induce apoptosis through JNK pathway activation leading to downregulation of Bcl-2 and an increase in translocation of Bax (Chen *et al.*, 2010).

With such a vast array of venom components having the potential to be used in cancer therapies, it is important to efficiently screen many venom components at once. High throughput screening (HTS) may be utilised to perform simultaneous screening of many hundreds to thousands of venom components against a caner target or cell line of interest and will be discussed further in the following section.

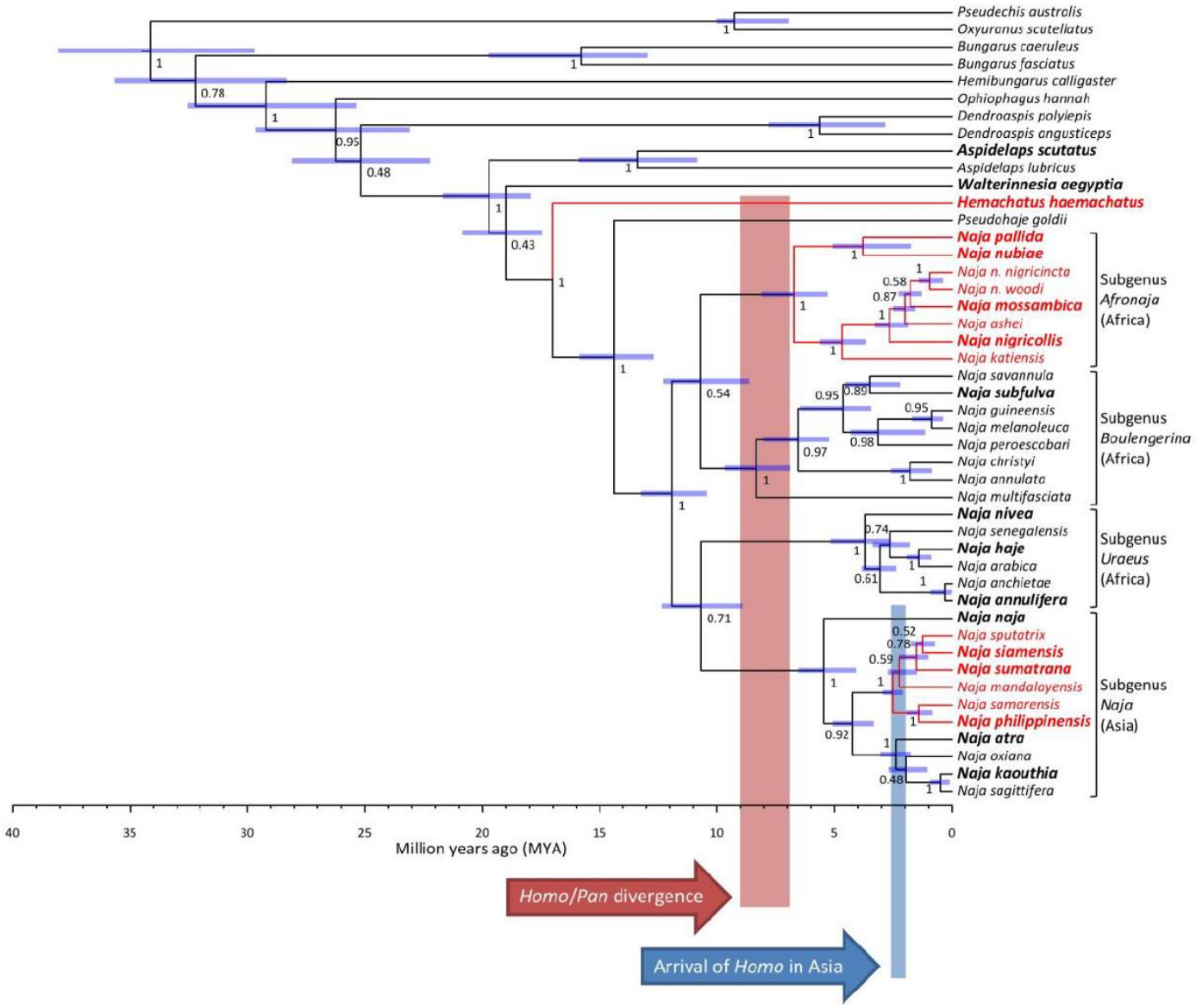



*Figure 1.2. Phylogenetic tree illustrating estimated dates of divergence of* Naja *clade from other related elapids.*

*Time (MYA) indicated at bottom with the divergence Homo/Pan indicated in red and the arrival of Homo in Asia indicated in blue. Node labels indicate posterior probabilities and red tipped labels denote spitting species. Figure reprinted from supplementary materials from: Kazandjian TD, Petras D, Robinson SD, et al. Convergent evolution of pain-inducing defensive venom components in spitting cobras. Science. 2021;371(6527):386-390 with permission from AAAS.*

## 1.4 Drug Discovery Process

### 1.4.1 High Throughput Screening

High throughput screening (HTS) is a method used to efficiently assess the effect of a large number of compounds against a model for a disease in order to identify compounds which are causing inhibition, changes in gene regulation, apoptosis, or another such desired biological effect. HTS incorporates automation in order to screen a large magnitude (thousands to millions) of compounds using biological screening techniques (Dandapani *et al.*, 2012). The aim is to discover chemical or biological compounds which could be optimised to become a potential new lead candidate for drug discovery (Babaoglu *et al.*, 2008). It is important to be able to robustly distinguish between compounds which are eliciting an effect from those which are not and so the assay must be designed carefully to ensure active compounds are found if they exist in the screen (Macarron and Hertzberg, 2011). Cell-based HTS methods may involve the measurement of second messengers, reporter gene assays or cell proliferation assays that monitor the effect of external factors on cell growth response (Sundberg, 2000).

In an ideal HTS campaign, active compounds would be easily identifiable and highly distinct from the untreated control wells. In addition, compounds would ideally be sorted into 'active' and 'inactive' categories with perfect accuracy. In reality, however, this is not always possible and variability and faults with assay design can lead to less than perfect compound identification. The two major problems during HTS campaigns are the presence of false positives and false negatives.

False negatives are compounds which elicit an effect against the intended target but are missed during a HTS campaign. Possible reasons for false negatives include quenching of fluorescent signal in a fluorescence reporter assay (Su *et al.*, 2015) or primary screening at too low a concentration for a response to be detected (Macarron *et al.*, 2011).

False positives are compounds that are erroneously identified as active when they are in fact not, often due to the compound interfering with the assay in some manner to appear as a hit compound. False positives could be caused by fluorescent interference in the form of autofluorescence (fluorescent signal emitted by the compound itself at a similar wavelength and intensity as the reporter molecule). It has been suggested that up to 50% of false positives could be due to fluorescence interference, potentially even masking the real fluorescence signals generated by active compounds (Su *et al.*, 2015).

Other potential causes of false positives include enzyme inhibition or denaturation (if running an enzyme based assay), or the compound interfering with the assay by its intrinsic reactivity,

for example by the generation of hydrogen peroxide which could interfere with the assay by causing redox reactions (Sink *et al.*, 2010). Some compounds are real hits in a screen, but end up being identified as promiscuous entities, or "frequent hitters" that come up often in many HTS campaigns and bind non-specifically to multiple targets making them inappropriate as lead compounds (Yang *et al.*, 2020b). These compounds may form aggregates which inhibit the target molecules non-specifically (Mcgovern *et al.*, 2002).

Although false positives are less detrimental to a HTS campaign as these will be re-screened and profiled and can be discarded later, they should be identified as early as possible in the screening campaign or else they can be expensive if taken forwards to structural optimisation only to fail later (Sink *et al.*, 2010). False negatives on the other hand can be harmful to the campaign as once screened, these compounds may never be revisited for the same target and so a potentially effective lead drug compound could be missed entirely (Posner *et al.*, 2009). It is imperative that the rate of both false positives and false negatives is as low as possible and HTS screens should be set up and analysed in such a way as to minimise the risk of this possibility.

HTS has thorough quality assurance (QA) methods applied to ensure that hits are high quality and data are free from errors. This includes the use of standards relevant to the type of assay undertaken (Coma *et al.*, 2009), high accuracy in liquid handling systems and robotics including technologies such as acoustic dispensing, high quality detection methods and instruments including state of the art plate readers with multiple reading capacities (Macarron *et al.*, 2011).

Although replicates are normally omitted in high volume screens due to the additional cost they would incur, compound libraries usually have multiple compounds from each chemotype in the library meaning that activity is rarely missed (Macarron *et al.*, 2011). The typical HTS workflow is shown in Figure 1.3. Compounds identified as 'active' at the primary screening stage have dose response curves performed to determine their potency. Following this, structure-activity relationship (SAR) analysis may be undertaken to determine which parts of the active compound(s) is/are contributing to the desired effect. Changes may be made to improve the active compound(s) to improve potency, leading to a lead compound which may be taken forwards for further investigation and eventually clinical trials as a potential drug entity. SAR is described further in Chapter 5.

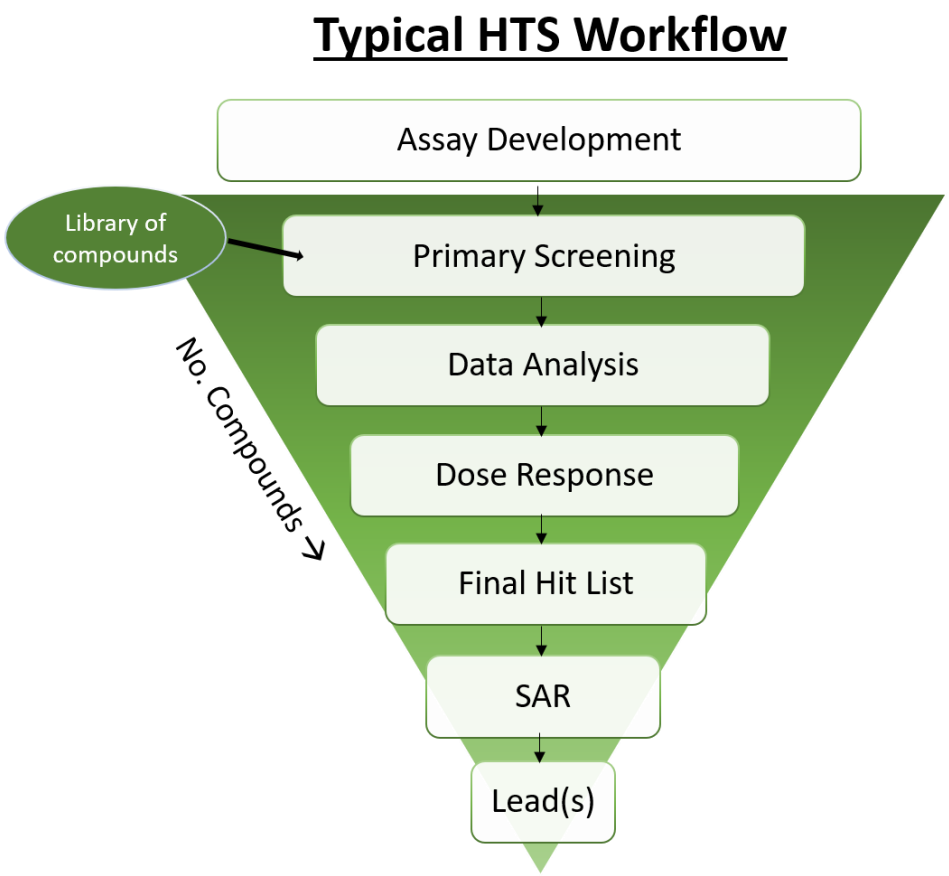



*Figure 1.3. Flowchart of the typical HTS workflow*

*The process starts with developing and validating an appropriate assay and concludes with the identification of lead compound(s). The compound library is introduced during the primary screen and the number of compounds remaining decreases at every stage of the process until just one or two lead compounds remain. SAR = structure activity relationship.*

### 1.4.2 Types of Screening

There are two main types of screening approaches in drug discovery: target-based and phenotypic screening. Target-based drug discovery (or 'hypothesis-driven' approach) relies upon knowing the identity and/or role of a specific drug target and designing an assay capable of identifying molecules that specifically interact with the target of interest. In contrast, phenotypic screening requires no such prior knowledge and instead can be used to screen potential drug molecules in disease models which are incompletely understood and for which a specific target has not been identified (Moffat *et al.*, 2017). Phenotypic screening is enjoying a recent resurgence after three decades of target-based discovery being the dominant screening approach.

Phenotypic screening has an advantage for finding first-in class drug molecules and drugs with novel mechanisms. Target based screening, on the other hand, tend to find 'best-in-class' drugs and is particularly effective for developing highly effective follower drugs (Swinney, 2013). For the treatment of PC in particular, therapeutics with novel mechanisms of action are desperately

needed and molecular and genetic targets contributing to the progression of PC are poorly understood compared to many other cancers. Therefore, a phenotypic screening campaign presents a potential means of identifying a lead molecule(s) with drug-like potential. In particular, phenotypic screening of natural products is proving to have strong advantages with 28 out of 50 'first in class' drugs gaining FDA approval between 1999-2008 originating from phenotypic screens (Harvey, 2014).

### 1.4.3 HTS Data Analysis

Following optimisation of the assay format, the robustness of the assay should be assessed before primary screening begins. A well accepted measure of assay robustness is the Z' ("Z-prime") assay, also known as "Z-factor" (Zhang *et al.*, 1999). This assay evaluates the difference between the mean signals generated by the positive and negative control but also considers the variability of measurements. It will be referred to hereafter simply as Z'.

The equation described by Zhang *et al*. (1999) is shown in Equation 1.1. The Z' is dimensionless and its value can never exceed 1 (but may fall below 0). A value of 1 in the Z' assay represents an 'ideal assay' which either has no variation (SD=0) or the dynamic range is infinite. A Z' of between 0.5-1 shows that there is a large separation band and that the assay is suitable for use in HTS. A Z' of 0-0.5 represents an assay with small separation between controls and should only be used for smaller scale screening or difficult to optimise assays, but is not suitable for HTS. A value below 0 shows the assay is not suitable for screening assays as it would be practically impossible to distinguish active compounds from the background signal.

$$Z - factor = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$

*Equation 1.1. How to calculate the Z-factor.*

*$\mu_p$= mean of the positive control, $\mu_n$ = mean of the negative control, $\sigma_p$ = standard deviation of the positive control, $\sigma_n$ = standard deviation of the negative control. The Z-factor cannot exceed 1 due to the denominator being an absolute value.*

There is currently a lack of universally applied processes for data analysis in HTS and statistics only serve a limited function at present (Malo *et al.*, 2006). Often, HTS data has very subjective means of analysis, such as when a compound looks considerably more active than the bulk of other compounds, it is deemed as active (Malo *et al.*, 2006). Whilst this can be useful for quickly detecting highly active compounds, it lacks repeatability and reliability and may lead to weaker compounds being missed. This type of analysis can also vary significantly from analyst to analyst. Activity is sometimes capped, for example as the "top 1% of activity", but this can be quite

arbitrary and could lead to compounds being missed in screens with many actives, and weakly active compounds being taken forwards in screens with few actives (Malo *et al.*, 2006).

When running a HTS campaign, despite best efforts to keep all variables constant, there will inevitably be at least a small amount of plate to plate variability. Data may be 'normalised' to remove systematic variation between plates and make individual plates more comparable to each other. Normalisation may be controls-based or non-controls-based. Ideally, a combination of both methods should be used in order to maximise reliability of results and reduce the risk and impact of false positive and false negatives (Brideau *et al.*, 2003). The following describe some of the most frequently used normalisation methods:

Percent of control

This is a method for converting the raw plate reading to a measure of activity/inhibition.

$$\text{Percent of control} = \frac{Xi}{\mu} \times 100$$

Where $xi$ is the raw measurement of the compound being investigated (the $i$th compound) and μ is the mean of the untreated control raw measurements.

Normalised percent inhibition

$$\text{Normalised percent inhibition} = \frac{\mu H - Xi}{\mu H - \mu L}$$

Where $Xi$ is the raw measurement of the compound being investigated (the $i$th compound) and μ is the mean of the high (μH) and low (μL) control values (Brideau *et al.*, 2003).

Z score

The Z score is a non-controls-based normalisation method and uses the mean of the measurements within the plate as the baseline for which to compare the response for each compound. Since HTS 'hits' are essentially outliers, the Z score is a statistics-based test to help to detect these 'abnormal' values (Brideau *et al.*, 2003).

$$Z = \frac{Xi - Xx}{Sx}$$

Where $Xi$ is the raw measurement of the compound being investigated (the $i$th compound), $Xx$ is the mean of all measurements within the plate and $Sx$ is the standard deviation of all measurements in the plate.

A typical workflow for HTS data analysis is shown in Figure 1.4. There are many methods for normalising data before analysing to find 'hit' compounds for HTS campaigns. Shun *et al.* (2011)

compared the appropriateness of different methods including percentage inhibition and Z score for HTS and quality control assays. The team concluded that there is no one ideal method for HTS data analysis and each data set should be assessed on a case by case basis.

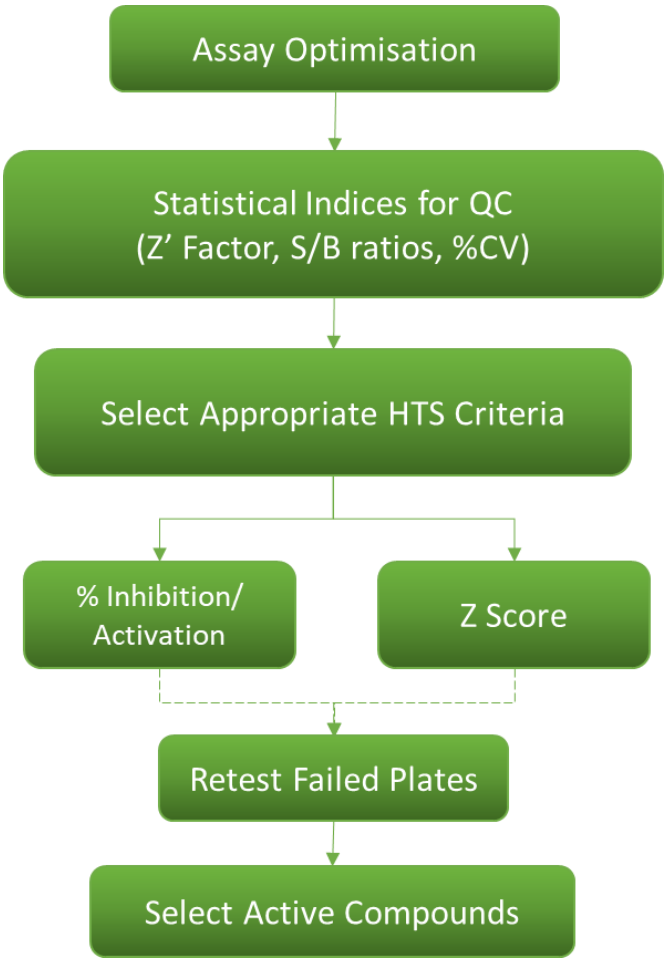



*Figure 1.4. Flowchart showing a simplified data analysis overview for HTS data.*

Once active compounds have been selected by an appropriate method, chemical filtering may be applied to reduce the number of compounds based on the desired physicochemical properties of the project team or company. For drugs designed to be given by the oral route, this may include size and solubility of the compound, oral absorption, mucosal permeability and other such factors (Alqahtani *et al.*, 2021). In 2001, Lipinski *et al.* developed a guide for desirable physicochemical properties of a potential drug candidate referred to as the 'Rule of Five'. In this paper, it is stated that potential drug-like molecules will have poor solubility and biological permeability if they have a molecular weight of more than 500, more than 5 H-bond donors or 10 H-bond acceptors, or the calculated LogP (cLogP) is greater than 5 (Lipinski *et al.*, 2001). Natural products are considered to be one of the best exceptions to this 'Rule of Five', whereby

chemical modifications to natural products can lead to potential drug lead molecules (Lipinski, 2016).

Emerging HTS platforms include microfluidic droplet mode HTS allowing extremely small (down to picolitre) volumes to be effectively screened by utilising water-in-oil droplets to separate reagents from neighbouring droplets. These systems may even be utilised for cell-based assays, significantly lowering the reagent usage and allowing 500 fold throughput compared to current capacities (Du *et al.*, 2016). However, it is difficult to change the cell media in this format, leading to a lack of essential nutrients and the build-up of toxic materials (Du *et al.*, 2016). Microfluidics may be useful in producing improved absorption, distribution, metabolism, and excretion (ADME) data as well as toxicity information. It also reduces the rate of false positives which can be a barrier in plate based HTS and gives a larger amount of information about the activity and safety profile of the compound (Mullin, 2004).

*In silico* screening can be helpful in reducing the starting number of compounds in which to screen *in vitro*. This can involve collecting potential hit compounds that may be screened and verified *in vitro* (Shoichet, 2004), or using a negative screening effort which starts by eliminating compounds which have no change of binding to the particular target. From there, a large compound library can be whittled down to a smaller number (e.g. 2-5000 compounds) of diverse compounds based on what is the most likely to 'hit' the target which may then be screened *in vitro* (Mullin, 2004).

### 1.4.4 The Use of HTS In The Search For Therapeutics For Treating Pancreatic Cancer

HTS has been widely utilised in the past three decades including in the search for novel cancer therapeutics. These searches have been fruitful with the development of several kinase drugs including Gefitinib, Erlotinib, Sorafenib, Dasatinib and Lapatinib being the result of HTS campaigns (Macarron *et al.*, 2011). Each of these oncology drugs was discovered by a different pharmaceutical company, indicating the widespread use of HTS for advanced drug discovery.

For PC in particular, screens looking at combining existing chemotherapeutics for treatment of PC have been studied, reinvigorating old compounds for a novel application (Langdon *et al.*, 2017). Guzmán *et al* (2017) screened over 1,000 marine natural products against PC cells in culture. Ten of the screened products were found to inhibit IL-8 production, some of which also induced cytotoxicity in the cell lines tested.

In addition, virtual screening of PC cells has been performed (Song *et al.*, 2020). The identified hits were validated *in vitro* and confirmed to have inhibitory effects on the migration and invasive abilities of PC cells in culture.

### 1.4.5 The Use of HTS For Screening Venoms

There have been a few venom studies which have utilised HTS in the form of plate-based toxicity studies. *Crotalus* venom was screened against a range of cancer cells including PC cells to assess cytotoxicity (Teixeira *et al.*, 2016). L-amino acid oxidase from *Crotalus durissus terrificus* venom was found to be cytotoxic against multiple cancer cell lines with particular sensitivity exhibited by glioma and pancreatic carcinoma cells. A HTS campaign was performed using fractionated scorpion venoms from the *Androctonus* genus to assess their effects on colorectal cancer cells. Primary colon cancer cells were inhibited by treatment with Gonearrestide, a scorpion venom-derived peptide (Li *et al.*, 2018a). HTS identification of snake venom proteins and peptides were performed to assess the effect of 20 snake venoms on blood coagulation. A miniaturised plasma coagulation assay was used to rapidly identify bioactive components of interest in 384-well format (Slagboom *et al.*, 2020).

## 1.5 Thesis Aims and Structure

This thesis has been constructed with the aim of answering the research question mentioned at the start of this chapter and has been broken down into the following sections to address the question in a logical manner.

### 1.5.1 The Research Question

Can animal venoms be utilised for the treatment of PC?

**Chapter 1:** This chapter aims to provide background information and an introduction to the subject and literature. This work has been guided by scientific findings, built upon in each chapter. Some subjects of the introduction have been guided by later results chapters to help place the findings into a wider literature context. Therefore, some of the details discussed in the introduction are present due to later identification of entities unknown at the commencement of the research.

**Chapter 2:** Both general and chapter specific materials and methods are found in this chapter.

**Chapter 3:** The aim of this chapter was to design an appropriate assay to allow effective screening of whole and fractionated venoms. The first section of this chapter demonstrates that the resazurin assay is an effective an appropriate method to determine cell cytotoxicity. The

second section of this chapter indicates 5 out of 19 tested cobra venoms are selectively cytotoxic against SW620 cancer cells.

**Chapter 4:** The aim of this chapter was to fractionate the whole venoms selected in Chapter 3 and rescreen them for activity to select potential lead compounds. The first section of this chapter demonstrates that a combination of ion exchange and reverse phase chromatography provide an effective method of separating whole venoms. The second section of this chapter demonstrates how a number of venom fractions have selective activity against BxPC-3 pancreatic cancer cells.

**Chapter 5:** The aim of this chapter was to identify the venom fractions chosen in Chapter 4. The first section of this chapter uses mass spectrometry to identify peptide fragments, then matches these to venom proteins available in public databases and suggests potential sequences of the venom fractions. The second section of this chapter explores structure activity relationships and indicates which amino acid residue positions are most important for activity against cancer cells.

**Chapter 6:** The aim of this chapter was to explore what effects the cardiotoxins identified in Chapter 5 are having on relevant cancer genes. The first section demonstrates the optimal parameters for assessing gene expression by qPCR with reference to the MIQE guidelines. The second section of this chapter investigates which of the genes are being influenced by the selected cytotoxic venom components.

**General Discussion:** This chapter aims to summarise the findings of the previous chapters and describe their weaknesses, implications and future research which could be performed to build upon this study.

# Chapter 2 - Materials and Methods

## 2.1 Cell Culture

### 2.1.1 Cell Lines Utilised in this Study

BXPC-3, adherent epithelial pancreatic adenocarcinoma cells originating from the pancreas of a 61-year-old female, obtained as a gift from the University of Manchester (Manchester, UK).

SW620, adherent epithelial colorectal adenocarcinoma cells derived from metastasis in lymph node from a 51-year-old Caucasian male, obtained as a gift from the University of Kent (Canterbury, UK).

### 2.1.2 STR Profiling

Both these cell lines were authenticated using STR profiling in March 2018. The BxPC-3 cell line was confirmed to be generated from the same source material as the original cell lines profiled on the Cellosaurus database (see Appendix I for STR profiling report).

Until this point, the SW620 cell line was considered to be the Mia PaCa-2 pancreatic adenocarcinoma cell line and all work performed was under this belief. The STR profiling indicated only a 27% identity with Mia PaCa-2 cell line when compared with the Cellosaurus database. The profiling confirmed that this cell line was, in fact, the SW620 colorectal cancer cell line with 100% match to the Cellosaurus database for this cell line (see Appendix I for STR profiling report).

Most of the laboratory-based work had been completed by the time this discovery was made, and resources did not permit sourcing new Mia PaCa-2 cells, or a suitable alternative PC cell line from the ATCC collection for comparison. Therefore, this work now compares the effect of cobra venoms on PC and CRC cell lines, as opposed to comparing two different PC cell lines.

### 2.1.3 Routine maintenance of cells in culture

SW620 cells were cultured using Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% Foetal Calf Serum (FCS) and 2 mM L-Glutamine, all from Thermo Fisher Scientific (Massachusetts, USA). BXPC-3 cells were cultured using Roswell Park Memorial Institute (RPMI-1640) media supplemented with 10% FCS, 2 mM L-Glutamine and 1% penicillin-streptomycin, all from Thermo Fisher Scientific (Massachusetts, USA).

Both cell lines were cultured as adherent monolayers and incubated at 37°C in a 95% air, 5% $CO_2$ atmosphere. Tissue culture flasks of 25 $cm^3$ to 75 $cm^3$ from Thermo Fisher Scientific (Massachusetts, USA) were used for cell growth and maintenance before plating.

Cells were passaged as required using Gibco trypsin- ethylene diamine tetra acetic acid (EDTA) (0.05% trypsin, 0.53 mM sodium EDTA) (1x) (Thermo Fisher Scientific, Massachusetts, USA). For general cell maintenance, cells were washed twice with 5 ml PBS, then 0.5 ml 0.005% trypsin-EDTA was applied to cells in 25 $cm^3$ flasks, or 1 ml to 75 $cm^3$ flasks and swirled gently to cover the bottom of the flask. Flasks were incubated at 37°C, 5% $CO_2$ for 2-5 minutes depending on cell line and state of confluency. Once cells had detached sufficiently from the bottom of the flask, appropriate cell media was added to total 10 ml in 25 $cm^3$ flasks, or 15 ml in 75 $cm^3$ flasks. Aliquots were then taken to new flasks for maintenance or counted and plated into 96-well plates for assays. Routine passaging of cells were performed as a 1 in 5 split for both cell lines.

To maintain adequate lower passage cell stocks, cells were routinely frozen down by washing cells twice with 5 ml PBS, then adding 0.2 ml 0.005% trypsin-EDTA and incubating at 37°C, 5% $CO_2$ until cells were detached. 1 ml of 90% FCS and 10% Dimethyl sulfoxide (DMSO) was added to cells and mixed thoroughly. The cell mixture was pipetted into 2ml cryotubes and frozen at -20°C for 30 minutes, then transferred to the -80°C freezer for storage for up to one year or to liquid nitrogen stores for longer term cell banking.

### 2.1.4 Preparation of cells for cell-based assays

Before counting, cells were washed twice with PBS, 500 µl of 0.005% trypsin-EDTA added and incubated at 37°C, 5% $CO_2$ for 2-5 minutes depending on cell line and state of confluency. Once cells had detached sufficiently from the bottom of the flask, appropriate cell media was added to total 10 ml. The cell mixture was transferred to a 15 ml centrifuge tube (Thermo Fisher Scientific, Massachusetts, USA) and mixed gently to fully suspend the cells. A sample of 20 µl harvested cells were combined with 12 µl PBS and 8 µl of 0.4% trypan blue in a 1.5 ml microcentrifuge tube. Cells were counted using a Marienfeld Superior Neubauer-improved haemocytometer (Thermo Fisher Scientific, Massachusetts, USA) and the cell concentration calculated. This was done by multiplying the observed cell number by 2 (to account for the dilution factor of PBS and trypan blue), then multiplying by $10^4$ to calculate for the scale of the haemocytometer.

The appropriate number of cells were plated into 96-well cell culture treated flat bottom Falcon® plates (Corning, New York, USA), or 384-well flat bottom cell culture treated plates (Greiner Bio-One, Kremsmünster, Austria) according to cell line optimisation results in section 3.2.1.4.

## 2.2 Resazurin Assay Optimisation

### 2.2.1 Preparation of Resazurin for assays

Resazurin sodium salt was obtained as powder from Sigma-Aldrich (Missouri, USA). Initially, a resazurin concentration of 2.7 mM as described by Sarker *et al*.(2007) was used for all cell lines.

Following optimisation, the final concentration chosen was 160 µM made with fresh cell media immediately before use from a 1.6 mM resazurin stock solution made with PBS stored in aliquots at -20°C.

### 2.2.2 Source of Venom and Venom Preparation Details

All venom was kindly provided by Venomtech Limited (Sandwich, UK). Venoms were obtained either wet (frozen) or lyophilised. Protein concentrations of both thawed and resuspended venoms were obtained using a DS-11 spectrophotometer (Denovix, Wilmington, USA) using the 280 nm 'protein' function.

A dose response assay was performed to decide which concentration of *Naja nigricollis* venom to use as a positive control. A 6-point, 1 in 10 serial dilution of venom was performed in water and PBS starting from 10 mg/ml *Naja nigricollis* venom, then SW620 cells were treated with each dilution series for 2 hours. Following this, 160 µM resazurin was added to the cells and the fluorescence value read using a FLUOstar Omega (BMG LABTECH, Ortenberg, Germany) plate reader as detailed in section 2.2.3 below. The percent growth for each dilution was calculated by dividing the fluorescence value of the venom treated cells by the fluorescence value of the untreated control cells and multiplying by 100. This value was subtracted from 100 to give the % inhibition. Graphs were plotted using Graphpad Prism 9.0.0 (GraphPad Software, San Diego, USA). A further 8-point, 1 in 2 dose response assay was performed over a narrower dose range starting from 100 µg. The % inhibition was calculated and plotted in the same manner as described above.

To determine an appropriate amount of time to expose cells to venom samples, SW620 and BxPC-3 cells were exposed to 100 µg/ml Naja nigricollis venom for up to 6 hours (0.5h, 1h, 2h, 3h, 4h, 5h and 6h). Venom was applied to the 6 hour cell sample, then an hour later venom was applied to the 5h cell sample, and so on. Then venom samples were removed and discarded, then resazurin was applied to the cells for 2h. The signal to background (S/B) ratio was calculated by dividing the fluorescence values of the venom treated wells by untreated control cells.

*Naja nigricollis* venom was used as a positive control for plate based assays and was filter sterilised using 0.2 µm Minisart syringe filter (Sartorius AG, Göttingen, Germany) and applied at a concentration of 100 µg/ml, diluted in PBS or appropriate cell media. Other venoms were not filter sterilised due to low volumes available but were each tested in a resazurin assay to test for potential contamination.

### 2.2.3 Plate Reader Settings for Resazurin Based Assays

FLUOstar Galaxy and FLUOstar Omega (BMG LABTECH, Ortenberg, Germany) plate readers were used to read absorbance and fluorescence values from 96 and 384-well assay plates. Following

fluorescence optimisation, an excitation wavelength of 544 nm and emission wavelength of 590 nm were selected for fluorescence assays. Plate reader settings of 100 flashes per well with a constant 800-unit gain were used.

### 2.2.4 Optimisation of Cell Seeding Densities and Wavelengths for Cell Assays

#### *2.2.4.1 96-well Cell Seeding Density Optimisation*

To determine the optimal cell density to seed for the assay, each cell line was independently plated at seeding densities of 2.5 x $10^4$, 5 x $10^4$, 1 x $10^5$ and 2 x $10^5$ cells per well in 96-well format and incubated at 37°C, 5% $CO_2$ for 24 hours. Cells were treated in triplicate with either 50 µl of 100 µg/ml *Naja nigricollis* venom (positive control) or 50 µl of appropriate cell media (negative control) and incubated at 37°C, 5% $CO_2$. Following 2 hours of exposure, the venom/media was removed and discarded and 50 µl of 160 µM resazurin solution (Sigma Aldrich, Missouri, USA) was added to each well. Absorbance plate readings were taken at 595 nm and fluorescent plate reads were taken at excitation wavelengths of 355, 485 and 544 nm, with an emission wavelength of 590 nm. Readings were taken at 0, 15, 30, 45 and 60 minutes, then every thirty minutes until 300 minutes were reached. The optimum number of cells (2 x $10^5$ cells/well for SW620 and 1 x $10^5$ cells/well for BxPC-3 cells) were plated for future 96-well assays.

#### *2.2.4.2 384-well Cell Seeding Density Optimisation*

A similar optimisation process was undertaken for 384 well plates. BxPC-3 cells were plated at seeding densities of 6.5 x $10^3$, 1.25 x $10^4$, 2.5 x $10^4$ and 3.3 x $10^4$ cells per well and SW620 cells were plated at seeding densities of 2 x $10^4$, 2.5 x $10^4$, 3.3 x $10^4$ and 5 x $10^4$ cells per well and incubated at 37°C, 5% $CO_2$ for 24 hours. Cells were treated in triplicate with 25 µl of either 100 µg/ml *Naja nigricollis* venom (positive control) or appropriate cell media (negative control) for 2 hours, then venom/media was removed and discarded. 25 µl of 160 µM resazurin was then added to each well. Plates were read at fluorescence excitation wavelength 544 nm and emission wavelength 590 nm after 0, 15, 30, 45, 60, 90, 120 and 150 minutes. The optimum number of cells (5 x $10^4$ cells per well for SW620 cells and 2.5 x $10^4$ cells per well for BxPC-3 cells) were plated for all future 384 well assays.

#### *2.2.4.3 Determining Between the use of Absorbance or Fluorescence Plate Reading for Resazurin Based Assays*

To determine the optimum plate reader wavelength for cell viability testing, various fluorescent and absorbance wavelengths were tested. The excitation wavelength spectra for resorufin (the fluorescent product of converted resazurin) spans between approximately 480-600 nm so excitation wavelengths of 485 and 544 nm were tested with 355 nm as a comparison. The emission wavelength spectra for resorufin peaks at 584 nm, so 590 nm was chosen as it was the closest wavelength available on the chosen plate reader.

The absorbance spectra for resorufin peaks at 573 nm, so 595 nm was chosen as the closest wavelength to measure absorbance from plate-based assays.

### 2.2.5 Optimisation of Resazurin Concentration

To determine the optimum concentration of resazurin to use for plate based assays, a 1 in 4 serial dilution of resazurin was performed to produce concentrations of 10.8 mM, 2.7 mM, 675 µM, 168.75 µM, 42.19 µM and 10.55 µM to cover and extend beyond the range of resazurin concentrations often used in the literature (Zrimšek *et al.*, 2004, Sarker *et al.*, 2007, O'brien *et al.*, 2000). A 21.6 mM resazurin stock solution was made using powdered resazurin sodium salt (Sigma Aldrich, Missouri, USA) plus colourless DMEM media (Thermo Fisher Scientific, Massachusetts, USA) plus 10% FCS (Thermo Fisher Scientific, Massachusetts, USA). Appropriate dilutions were made from the stock concentration.

These doses of resazurin were tested on SW620 cells plated at a density of $2x10^5$ cells/well in order to determine the optimal resazurin concentration.

Heat maps were generated in order to visually observe differences between the positive and negative control wells of different number of plated cells and different concentrations of resazurin added to each plate. The heat maps were generated using the green-yellow-red colour scale found in the conditional formatting tab in Microsoft® Excel® for Office 365 (Microsoft Corporation, Washington, USA).

### 2.2.6 Z' Assay

#### *2.2.6.1 96-well format*

Assay robustness was determined by Z' assay for SW620 and BxPC-3 cell lines. Z' assays were performed by plating the previously determined optimum cell number into 96-well plates ($2 \times 10^5$ cells/well for SW620, $1 \times 10^5$ cells/well for BxPC-3 cells). Half of the wells (top left and bottom right diagonal quarters) had 50 µl of 100 µg/ml *Naja nigricollis* venom added as a positive control. To the other half, appropriate cell media was added as a negative control. Following 2 hours of exposure, venom/media was removed and discarded and replaced with 50 µl of 160 µM resazurin solution. Absorbance and fluorescence readings were taken at time points 1, 2, 2.5, 3 and 4 hours for SW620 cells and at 1, 2 and 4 hours for BxPC-3 cells.

The equation

$$Z - factor = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$ (where σ = sample standard deviation and µ =sample mean)

was applied to the results to calculate the Z' value for each cell line, time point and absorbance/fluorescence wavelength.

#### *2.2.6.2 384-well format*

The Z' assay was repeated for 384-well format. The method detailed above was followed but with cell numbers of $5x10^4$ cells/well for SW620 cells and $2.5x10^4$ cells/well for BxPC-3 cells. For the 384-well assays, 25 µl of 160 µM resazurin was added to each well. Data were collected from some additional time points for this assay compared with the 96-well version.

## 2.3 Cobra Screen

### 2.3.1 Cobra Screen Assay Details

Nineteen cobra venoms were prepared for the cobra screen. Details of these cobras including number of individuals contributing to venom pools, their gender(s) and locale information are shown in Table 2.1. As most individuals were captive bred, sometimes for many generations, it is not possible to provide more precise locale information for many of the species listed. As this is not an anti-venom project reliant on accurate locale data to ensure suitability of potential anti-venom treatment in specific geographic regions, plus sequence data was later acquired by mass spectrometry analysis, this was deemed to be acceptable for this project.

These venoms were either frozen whole venom, diluted to 100 mg/ml stock or lyophilised venom resuspended to 100 mg/ml stock. Venom stocks were diluted to 100 µg/ml working stock and plated into rows A and E of a compound plate. 10 µg/ml concentrations were made by transferring from 100 µg/ml plated venom to create a 1/10 dilution. A positive control of 100 µg/ml *Naja nigricollis* venom and a negative control of appropriate cell media were used. All wells were performed in quadruplicate.

Prepared venom samples were applied to cells for 2 hours, then removed and discarded. 160 µM resazurin solution was then added to each well of the plate and incubated for two and a half hours. Plate reader measurements and images were taken after 1, 2 and 2.5 hours. Statistics were performed according to section 2.4.

*Table 2.1. Cobra species included in this study including their common name, number of specimens and their gender(s) (M = male, F = female, U = unknown), plus the locale information for each species.*

| Species | Common name | Number of individuals (gender) | Locale |
|---|---|---|---|
| *Naja mossambica* | Mozambique spitting cobra | 2 (M/F) | Southeast Africa |
| *Naja nigricincta nigricincta* | Zebra Spitting Cobra | 2 (M/F) | Southern Africa |
| *Naja nigricollis* | Black Necked Spitting Cobra | 2 (M/F) | Sub-Saharan Africa |
| *Naja nubiae* | Nubian Spitting Cobra | 2 (M/F) | Northern Africa |
| *Naja pallida* | Red Spitting Cobra | 2 (M/F) | East Africa |
| *Naja samarensis* | Siamese Samar Cobra | 2 (U/U) | Southern Philippines |
| *Naja siamensis* | Indo Chinese Spitting Cobra | 2 (M/M) | Southeast Asia |
| *Naja sputatrix* | Javan Spitting Cobra | 2 (M/M) | Indonesia |
| *Naja annulifera* | Snouted Cobra | 2 (M/F) | Southern Africa |
| *Naja anchietae* | Anchieta's Cobra | 2 (M/U) | Southern Africa |
| *Naja haje* | Egyptian Cobra | 2 (M/F) | Northern Africa |
| *Naja haje legionis* | Moroccan Cobra | 1 (F) | Morocco |
| *Naja melanoleuca* | Forest Cobra | 2 (M/F) | Africa |
| *Naja nivea* | Cape Cobra | 2 (M/F) | Southern Africa |
| *Naja atra* | Chinese Cobra | 2 (M/F) | China & Taiwan |
| *Naja kaouthia* | Monocled Cobra | 2 (M/F) | South & Southeast Asia |
| *Naja naja* | Spectacled Cobra | 2 (M/F) | Sri Lanka |
| *Naja naja karchachensis* | Pakistan Black Cobra | 2 (M/F) | Pakistan |
| *Ophiophagus hannah* | King Cobra | 1 (M) | Malaysia |

## 2.4 Statistical Analysis

A variety of statistical analysis methods were employed throughout this work depending on the distribution of the data and the hypothesis to be tested. The data were tested for normality to determine the most appropriate statistical test to apply. Quantile-quantile plots (QQ-plots) were plotted for visual understanding of the data and Anderson-Darling, D'Agostino & Pearson, Shapiro-Wilk and Kolmogorov-Smirnov normality tests were run on each data set to test for normality. The statistical test was chosen based on the workflow shown in Figure 2.1.

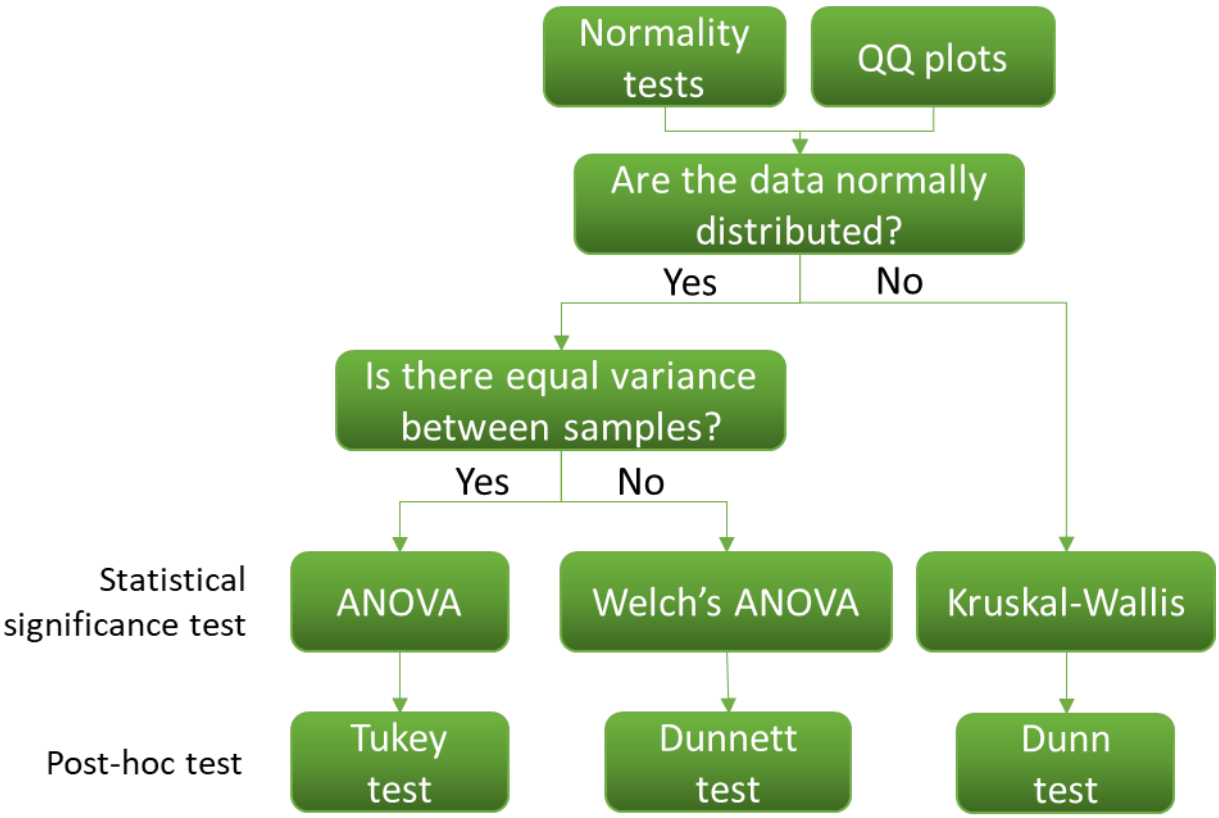



*Figure 2.1. Workflow used to select the most appropriate statistical test for each set of data.*

*Appropriate parametric or non-parametric tests were selected based on the distribution of the data and whether the samples had equal variance.*

Data with a normal distribution were tested for equal variance between samples and either a one-way ANOVA with Tukey post-hoc test or a Welch's ANOVA with Dunnett post-hoc test were run. Data found not to have a normal distribution were tested with the non-parametric Kruskal-Wallis and non-parametric Dunn post-hoc test. All statistical tests were performed using Graphpad Prism 9.0.0 (GraphPad Software, San Diego, USA).

## 2.5 High Performance Liquid Chromatography

### 2.5.1 First Dimension Reverse Phase HPLC

The buffers used throughout both first- and second-dimension high performance liquid chromatography (HPLC) are displayed in Table 2.2.

*Table 2.2. Compositions of buffers.*

*The composition of the buffers used throughout the thesis are displayed below. TFA($C_2HF_3O_2$) = Trifluoric Acid, ACN($C_2H_3N$) = acetonitrile, $NaH_2PO_4$ = Sodium Dihydrogen Orthophosphate Dihydrate, $Na_2HPO_4$ = Sodium phosphate dibasic dihydrate, NaCl = Sodium Chloride. All buffers made up to volume with HPLC grade water. RP = Reverse Phase, IE = Ion Exchange.*

| Buffer | TFA ($C_2HF_3O_2$) (% v/v) | ACN ($C_2H_3N$) (% v/v) | $NaH_2PO_4$ (mM) | $Na_2HPO_4$ (mM) | NaCl (M) |
|---|---|---|---|---|---|
| **RP buffer A** | 0.05 | - | - | - | - |
| **RP buffer B** | 0.045 | 80 | - | - | - |
| **RP buffer C** | 0.2 | 30 | - | - | - |
| **RP buffer D** | 0.1 | 30 | - | - | - |
| **IE buffer A** | - | - | 17 | 3 | - |
| **IE buffer B** | - | - | 15 | 5 | 1 |

#### *2.5.1.1 Sample Preparation for First Dimension Reverse Phase HPLC*

Dried cobra venoms were weighed out to give between 12-40 mg venom and the appropriate volume of reverse phase (RP) buffer A (see Table 2.2) was added to give a 100 mg/ml concentration. The samples were vortexed briefly then incubated on the bench for 5 minutes to aid passive suspension and reduce foaming. The samples were then gently pipetted up and down to mix thoroughly. Wet venoms were diluted with RP buffer A (see Table 2.2) to 100 mg/ml. Once re-suspended or diluted, the samples were centrifuged for 5-10 minutes (1000 rpm) using a benchtop centrifuge to pellet insoluble debris and reduce air bubbles. The sample was carefully transferred to a HPLC vial avoiding the pellet and any air bubbles. If the volume of sample prepared was less than 200 µl, a microvial insert was used.

#### *2.5.1.2 First Dimension Reverse Phase HPLC Protocol*

The reverse phase HPLC column (Vydac 218TP54 5 µm 250x4.6mm, Jaytee part number GP-STD-125) was used on an Agilent 1100 HPLC system. The column was conditioned by flowing 95% RP buffer A with 5% Buffer B at a flow rate of 1 ml/min for at least 10 minutes prior to sample loading (see Table 2.2 for buffer compositions). For analytical HPLC, 50 µg protein was loaded and for preparative runs, up to 100 µl (≤10 mg protein) of sample was injected and in both cases, 'Venom RP gradient' protocol (Table 2.3) was run.

*Table 2.3. 'Venom RP gradient' protocol.*

*This protocol was used for first dimension separation using reverse phase chromatography. Flow rate = 1ml/min. RP Buffer A = 0.05% TFA, 99.95% $H_2O$, RP Buffer B = 0.045% TFA, 80% ACN, 19.955% $H_2O$. Pump mode = gradient, fraction collector mode = peak based, max peak duration = 4 mins, working mode = threshold/slope, up-slope threshold 0.1U/s, down slope 0.5U/s threshold, collection threshold 2U.*

| Time | % RP Buffer A | % RP Buffer B |
|---|---|---|
| **0** | 95 | 5 |
| **10** | 95 | 5 |
| **20** | 80 | 20 |
| **100** | 40 | 60 |
| **120** | 0 | 100 |
| **130** | 0 | 100 |
| **131** | 95 | 5 |
| **140** | 95 | 5 |

#### *2.5.1.3 Lyophilisation of Reverse Phase Fractions*

Following separation, samples were lyophilised according the following manual protocol: The freeze drier (AdVantage Pro, VirTis SP Scientific, Philadelphia, USA) shelf temperature was set to -50°C and the condenser to -70°C. The vacuum pressure was set to 50 mBar to create a frost seal and the samples were left to freeze at -50°C (approx. 1.5 hours). Once frozen, the vacuum was switched on and allowed to equilibrate at ~500 µBar. After 5 hours, the temperature was changed to -40°C. The following day (approx. 24 hours later), the temperature was set to -30°C. Following the weekend (approx. 66 hours later), the temperature was changed to -20°C, then -10°C (3.5 hours later), then 0°C (1.5 hours later), then +4°C (one hour later). Once fully dried, the samples were stored at -20°C.

#### *2.5.1.4 Sample Preparation for First Dimension Fraction Screening*

Lyophilised venom fractions were resuspended in 100 µl phosphate buffered saline (PBS) (Thermo Fisher Scientific, Massachusetts, USA) and protein concentration measured by DS-11 spectrophotometer using the 'protein' setting (Denovix, Wilmington, USA). Samples were diluted to 10 µg/ml concentration with PBS and plated into a 96-well compound plate.

*Naja nigricollis* fractions were prepared the previous day and stored at 4 °C overnight. *Naja nigricincta* fractions were prepared fresh on the day of assay and were used immediately. Initial assays with *Naja mossambica, Naja pallida* and *Naja nubiae* venoms had compound plates prepared in advance and frozen at -20 °C. Compound plates for repeated assays were prepared fresh on the day.

#### *2.5.1.5 First Dimension Reverse Phase Fraction Screening Assay*

Following separation of venom components by first dimension reverse phase HPLC, fractions were tested on the SW620 cell line to assess toxicity of each fraction. $2x10^5$ SW620 cells per well

were plated into each well of a 96-well clear flat-bottomed Falcon® plate (Corning, New York, USA) and allowed to adhere overnight. The cell media was then removed from each well and discarded. 50 µl of each fractionated venom sample was pipetted into corresponding wells in triplicate and incubated (37°C, 5% $CO_2$) for 2 hours. After this time, the venom fractions were removed and discarded and 50 µl of 160 µM resazurin solution was added to each well and incubated (37°C, 5% $CO_2$) for a further 2 hours. The plates were then read using the FLUOstar Omega plate reader (BMG Labtech, Ortenberg, Germany) set for fluorescence at excitation 544 nm and emission 590 nm.

### 2.5.2 Second Dimension Size Exclusion Chromatography

#### *2.5.2.1 Sample Preparation for Second Dimension Size Exclusion Chromatography*

'Hit' compounds which were identified from the first dimension RP HPLC run were fractionated further using size exclusion chromatography (SEC). The first dimension 'hit' fractions can be found in Table 2.4 below.

*Table 2.4. Active venoms identified during the first dimension reverse phase HPLC screen.*

*The active venoms are displayed below by the name of the species of cobra and which reverse phase fraction number contained the 'hit' compound(s).*

| **Species** | Reverse phase fraction # |
|---|---|
| ***Naja mossambica* (N.mos)** | 13 |
| ***Naja nigricincta* (N.nct)** | 8, 9, 10 |
| ***Naja nigricollis* (N.nig)** | 11, 12 |
| ***Naja nubiae*(N.nub)** | 10 |
| ***Naja pallida* (N.pal)** | 10 |

Samples were made up to a maximum of 100 µl using IE buffer A (see Table 2.2 for buffer compositions).

#### *2.5.2.2 Second Dimension Size Exclusion HPLC Protocol*

The 'Venom SEC ISO3D' protocol used for second dimension size exclusion chromatography can be found in Table 2.5. An injection volume of 100 µl and flow rate of 2 ml/min was used on isocratic pump mode with a 10-minute run time. The column was conditioned by running IE buffer A through the column for at least 10 minutes prior to sample loading. The fraction collector was set to 'peak-based' with a max peak duration of 2 minutes (4 ml). The threshold/slope working mode was used with an up-slope threshold of 0.1 U/s, down slope threshold of 0.5 U/s and sensitive collection threshold of 1.5 U with peak-based detection at A280 nm.

*Table 2.5. Venom SEC ISO3D protocol*

*This protocol was used for second dimension SEC. Buffer compositions can be found in Table 2.2.*

| Step | Time (min) | Flow (ml/min) | % IE Buffer A | % IE Buffer B | Max pressure (bar) |
|---|---|---|---|---|---|
| **1** | 0 | 2 | 100 | 0 | 230 |
| **2** | 10 | 2 | 100 | 0 | 230 |

#### *2.5.2.3 Second Dimension Size Exclusion Fraction Screening Assay Protocol*

Second dimension reverse phase/size exclusion fractions were screened for cytotoxicity against SW620 cells. The protein concentration of the second dimension fractions were measured using the DS-11 spectrophotometer (Denovix, Wilmington, USA). Samples with enough protein content were diluted to both 10 µg/ml and 5 µg/ml and added to a 96-well low protein binding compound plate. Samples with less protein available were diluted to 5 µg/ml concentration, then plated into the compound plate. Where the protein concentration was too low to be accurately read by the spectrophotometer, the samples were applied directly to the compound plate without further dilution. SW620 cells were plated and screened according to 'First Dimension RP Fraction Screening Assay'.

### 2.5.3 Investigating the Effect of Concentrated Ion Exchange Buffers on SW620 Cells

The effect of concentrated IE Buffer A was investigated to observe its effect on SW620 cell viability. IE Buffer A (see Table 2.2) was lyophilised from a known volume and resuspended in water to create various concentrated samples between 20X concentrated down to 0.1X concentrated to mimic the salt concentrations likely to be present in the lyophilised 2D fractions. Cells were plated and screened according to 'First Dimension RP Fraction Screening Assay'.

### 2.5.4 First Dimension Size Exclusion Chromatography Optimisation and Cell Assay

Several parameters were tested to determine the optimum conditions for using size exclusion chromatography (SEC) as the first dimension for HPLC.

A series of SEC buffers and protocols were tested to optimise the first dimension HPLC. The buffer compositions can be found in Table 2.2 and the run parameters can be found in Table 2.6 below. Run 1 used the original IE buffer A condition that was successfully used for the second-dimension size exclusion chromatography separation. Run 2 had the run time increased to 40 minutes as the 20-minute run cycle resulted in half a peak being cut off suggesting that sample was still on the column when the run ended. The downslope threshold was adjusted to from 0.5 U/s 0.2 U/s try to collect peaks which were less steep to improve sample collection. The buffer was then changed to reverse phase-based buffers and the collection wavelength changed from 280 nm to 215 nm.

*Table 2.6. First dimension size exclusion optimisation protocols.*

*Protocol details for the optimisation runs for first dimension size exclusion chromatography. Buffer compositions can be found in Table 2.2.*

| Run # | Flow rate (ml/min) | Run time (mins) | RP buffer A | RP buffer B | RP buffer C | RP buffer D | IE buffer A | IE buffer B | Up-slope (U/s) | Down-slope (U/s) | Collection wave-length (nm) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 1 | 20 | 0 % | 0 % | 0 % | 0 % | 100 % | 0 % | 0.1 | 0.5 | 280 |
| **2** | 1 | 40 | 0 % | 0 % | 0 % | 0 % | 100 % | 0 % | 0.1 | 0.2 | 280 |
| **3** | 1 | 40 | 0 % | 0 % | 0 % | 0 % | 80 % | 20 % | 0.1 | 0.2 | 280 |
| **4** | 1 | 40 | 62.5 % | 37.5 % | 0 % | 0 % | 0 % | 0 % | 0.1 | 0.2 | 280 |
| **5** | 1 | 40 | 0 % | 0 % | 100% | 0% | 0 % | 0 % | 0.1 | 0.2 | 280 |
| **6** | 1 | 40 | 0% | 0% | 0% | 100% | 0% | 0% | 0.1 | 0.2 | 215 |
| **7** | 0.5 | 40 | 0% | 0% | 0% | 100% | 0% | 0% | 0.1 | 0.2 | 215 |
| **8** | 1 | 40 | 0 % | 0 % | 100% | 0% | 0 % | 0 % | 0.1 | 0.2 | 215 |

Samples of *Naja nubiae* and *Naja nigricincta* were prepared from whole, lyophilised stocks of venom. 5 mg of whole venom was loaded per full run to provide maximum yield without overloading the column. The optimised protocol used RP buffer C (see Table 2.2 for buffer composition) with isocratic pump mode. The run parameters can be found in Table 2.7. The column was conditioned by flowing RP buffer C through the column for at least 10 minutes prior to sample loading.

*Table 2.7. Optimised first dimension size exclusion protocol.*

*The optimised first dimension size exclusion chromatography protocol for separation of venom proteins for screening. Up-slope=0.1 U/s, Down-slope=0.2 U/s.*

| Step | Time (min) | Flow rate (ml/min) | RP buffer A | RP buffer B | RP buffer C | Max pressure (bar) | Collection wavelength (nm) |
|---|---|---|---|---|---|---|---|
| **1** | 0 | 1 | 0% | 0% | 100% | 230 | 215 |
| **2** | 40 | 1 | 0% | 0% | 100% | 230 | 215 |

Following fractionation, samples were lyophilised according to the lyophilisation protocol described in section 2.5.1 and stored at -20°C.

The *Naja nubiae* and *Naja nigricincta* first dimension fractions were screened against SW620 cells. SW620 cells were plated and assayed in the same manner described in 'First Dimension Reverse Phase Fraction Screening Assay Protocol' in section 2.5.1.

### 2.5.5 Second Dimension Reverse Phase HPLC and Cell Assay

Active fractions identified from the first-dimension size exclusion chromatography screen were further fractionated by reverse phase HPLC. The reverse phase HPLC column (Vydac 218TP54 5 µm 250x4.6mm, Jaytee part number GP-STD-125) was conditioned for at least 10 minutes prior

to sample loading by running 95% RP buffer A (see Table 2.2 for buffer composition) through the column at a flow rate of 2 ml/min.

Two shortened HPLC protocols were used for the second-dimension separation. If sample to be injected totalled 100 µl or less, the full reverse phase protocol (Table 2.8) was run. If sample to be injected totalled more than 100 µl, one or more 'loading protocol' (Table 2.9) was run prior to the full reverse phase protocol (Table 2.8) in order to load the required quantity of protein onto the column.

*Table 2.8. Second dimension reverse phase HPLC protocol.*

*% RP Buffer B indicated the percentage of reverse phase buffer B that was used for each section of the protocol.*

| **Time (minutes)** | % RP Buffer B | Flow rate (ml/min) |
|---|---|---|
| **0** | 5 | 2 |
| **2.5** | 20 | 2 |
| **22.5** | 60 | 2 |
| **27.5** | 100 | 2 |
| **28** | 5 | 2 |
| **30** | 5 | 2 |

*Table 2.9. Protocol for 'loading' of samples that totalled more than 100µl total volume.*

| **Time (minutes)** | % RP Buffer B | Flow rate (ml/min) |
|---|---|---|
| **5** | 5 | 1 |

Following separation, the second-dimension reverse phase fractions were lyophilised in the same manner as described in the lyophilisation protocol and stored at -20°C.

SW620 cells were plated and the second-dimension reverse phase fractions screened in the same manner as described in 'First Dimension Reverse Phase Fraction Screening Assay Protocol'.

### 2.5.6 First Dimension Ion Exchange Chromatography

Whole, lyophilised venoms were resuspended in HPLC grade water and loaded into micro-vial inserts for HPLC loading. Volumes giving 10 mg of each venom were injected onto the HPLC. All five of the 'hit' venoms identified in the Cobra screen were separated (*Naja pallida*, *Naja nigricincta*, *Naja nigricollis*, *Naja nubiae* and *Naja mossambica*).

An ion exchange HPLC column (TSKgel SP-5PW 10um 7.5 x7.5mm Jaytee part number GP-STD-133) was conditioned for at least 10 minutes prior to sample loading by flowing 100% IE buffer A through the column at a flow rate of 1 ml/min (see Table 2.2 for buffer composition). A maximum of 100 µl was loaded onto the column and the protocol shown in Table 2.10 was run.

*Table 2.10. First dimension ion exchange protocol.*

*% IE B indicates the percentage of ion exchange buffer B that was used for each section of the protocol. The remainder of the buffer was ion exchange buffer A. See Table 2.2 for buffer compositions.*

| **Time (min)** | % IE B | Flow rate (ml/min) |
|---|---|---|
| **0** | 0 | 1 |
| **10** | 0 | 1 |
| **70** | 75 | 1 |
| **80** | 90 | 1 |
| **90** | 90 | 1 |
| **95** | 0 | 1 |
| **99** | 0 | 1 |
| **100** | 0 | 1 |

First dimension ion exchange fractions were lyophilised in the same manner as described in the lyophilisation protocol in section 2.5.1 and stored at -20°C.

### 2.5.7 Second Dimension Reverse Phase HPLC

First dimension vials were combined to give 10-11 first dimension samples per cobra species. As much of each sample was loaded into second dimension reverse phase HPLC as possible. The protocol detailed in Table 2.8 was followed for second dimension reverse phase separation of the first-dimension ion exchange fraction samples. Fractions were lyophilised according to the lyophilisation protocol in section 2.5.1 and stored at -20°C.

## 2.6 Two-Dimensional Cobra Venom Screening

Cells were plated into 384-well plates according to section 2.1.4. Per well, $5x10^4$ SW620 cells or $2.5x10^4$ BxPC-3cells were plated, according to cell number optimisation in section 3.2.1.4.2. Cells were allowed to adhere at 37°C, 5% $CO_2$ for 24 hours before the assays were started.

Dried second dimension fractions were resuspended in 50 µl of HPLC grade water. Each fraction was measured for protein concentration by DS-11 spectrophotometer (Denovix, Wilmington, USA). A compound library block was then assembled by diluting each sample to 0.1 mg/ml (*N.nct* and *N.mos*) or 0.4 mg/ml (*N.nig*, *N.nub* and *N.pal*) and either used the same day (*N.nig*, *N.nub* and *N.pal*) or frozen at -20°C and thawed immediately before use (*N.nct* and *N.mos*).

Compound plates were prepared on the morning of the assay by diluting samples in PBS to 10 µg/ml in low protein binding 96-well plates (Corning, New York, USA). Cell media was removed from the cell plates and discarded. 50 µl of each 2D venom fraction was pipetted into the corresponding well of the cell plate test plate. The samples were left on the cells for 2 hours and then removed and discarded. 50 µl of 160 µM resazurin was added to each well of the plates

and incubated at 37°C, 5% $CO_2$ for 2 hours. Plate readings were then taken immediately and every 30 minutes until 2 hours were reached.

### 2.6.1 Two Dimension Fraction Screen Data Analysis

The raw fluorescence values for each first-dimension cobra venom fraction screen were normalised using the 'Z score' method. For this particular data set, a 'Z score' of ±3 was set as an appropriate threshold according to Brideau *et al*. (2003). The alternative would have been to rank the results and select an appropriate proportion to take forwards for further work (Brideau *et al.*, 2003). Statistical analysis was performed according to section 2.4.

### 2.6.2 Second Dimension Fraction Dose Response Assays

Dose response assays were performed for the fractions identified as 'active' in the 2D fraction screens. BxPC-3 and SW620 cells were plated into 384-well plates according to section 2.1.4 and allowed to adhere at 37°C, 5% $CO_2$ for 24 hours before the assays were started. Second dimension venom fraction samples were prepared by performing a 10-point, two-fold serial dilution starting at 100 µg/ml down to 0.2 µg/ml in a 96-well low protein binding compound plate (Corning, New York, USA). Media was removed from the cells and discarded. 50 µl of each sample was transferred from the compound plates to the corresponding wells of the test plates. The samples were incubated for 2 hours, then removed and discarded. 50 µl of 160 µM resazurin was added to each well of the plates and incubated at 37°C, 5% $CO_2$ for 2 hours. Plate readings were taken immediately, then every 15 minutes until 2 hours were reached.

## 2.7 Nano Differential Scanning Fluorimetry (NanoDSF)

*Naja nigricollis* venom was prepared in a variety of buffers to assess thermal stability when exposed to different diluents. The diluents used were as follows: ultra-pure water, 10% DMSO, 90% isopropyl alcohol (IPA), reverse phase buffer C (30% ACN, 0.2% TFA) and reverse phase buffer D (30% ACN, 0.1% TFA). Reverse phase buffer B (80% ACN, 0.05% TFA) was also tested, however due to the volatile nature of the sample buffer, it was not possible to record a melt temperature for this sample.

All samples had a final concentration of 1 mg/ml and were run in triplicate. Approximately 10 µl was drawn up into each glass nanoDSF capillary tube and inserted into the NanoTemper Prometheus (NanoTemper Technologies GmbH, Munich, Germany). Samples were heated from 15-95°C with a 0.2°C/minute ramp (40 minute total run).

## 2.8 Mass Spectrometry

Mass spectrometry analysis was performed by Peak Proteins (The Biohub, Alderley Park, Cheshire, SK10 4TG) on the 4th October 2019. Peptide mapping and intact mass analysis were performed separately. The methods used are described as follows.

### 2.8.1 Peptide Mapping

Each sample was reconstituted in 50 µl of 100 mM ammonium bicarbonate. To this, 5 µl of 100 mM DTT in 100 mM ammonium bicarbonate was added and incubated at 65°C for 30 minutes. 5 µl of 500 mM iodoacetamide in 100 mM ammonium bicarbonate was added and incubated at RT for 30 minutes in the dark. 10 µl of 25 ng/µl trypsin in 50 mM ammonium bicarbonate was added and incubated overnight at 37°C. 10 µl of the digest was taken and 10 µl of 0.1% TFA was added.

10 µl of sample was loaded onto the Sciex Exion LC and a 10-minute RP gradient was run. Buffer A was 0.1% formic acid and buffer B was 0.1% formic acid in 100% ACN. The gradient used started at 5% B, increasing to 45% over 15 minutes and then a 95% B wash followed by re-equilibration at 5% B. The flow was set to 300 µl. The column used was a Phenomenes Lunar 1.6 µm, PS C18, 100A 150x2.1 mm. The flow from the column was passed into the Sciex X500B mass spectrometer set to positive ion mode for data collection. The source was set to 400°C, 5500 V with gas at 30 psi to enable ionisation of the elute. A TOF mass window of 300-1800 Da was collected scanning at 1.2 seconds. Up to 10 MSMS were collected per scan using an information dependant acquisition method. A positive calibration mix was used to calibrate the X500B and the experiment had an error of approximately 1 ppm.

The collected .wiff files were converted to Mascot Generic Files (.mgf) and scanned against the SwissProt database using Mascot Daemon version 2.8 (Matrix Sciences, Boston, USA). Parameters were set for Trypsin digest (cleaving after K or R, provided they were not followed by a P residue). Modifications used were Carbamidomethyl as a fixed modification on Cys residues and oxidation on Met residues. Up to 2 missed cleavages were allowed with an allowed error of 10-20ppm. Example Mascot Daemon search criteria can be seen in Figure 2.2.

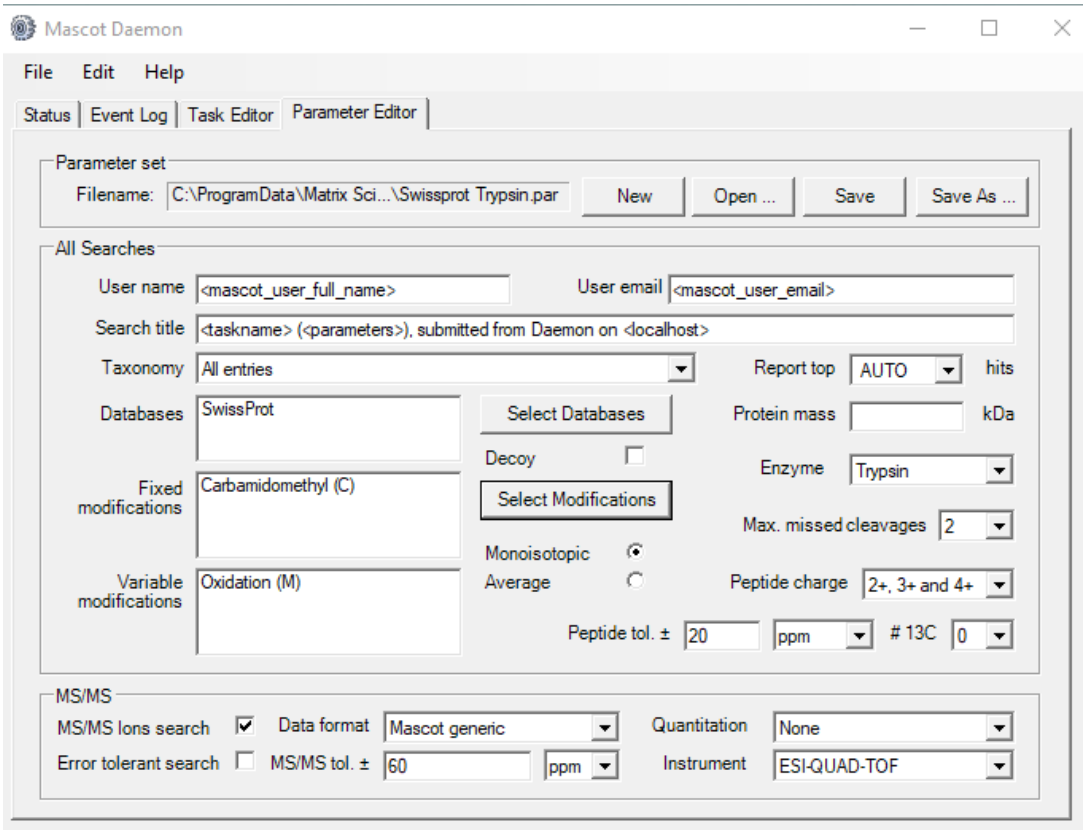



*Figure 2.2. Mascot Daemon search parameters.*

*Search parameters used to query the MS data generated by Peak Proteins. All database querying was performed by Peak Proteins.*

Mascot searched the theoretical ions generated from the Swissprot database against the experimental ions and matched them together in a search result. These results show the peptide identified along with the MSMS ions mapped against the sequence. The results were then validated manually. Mascot generated a probability score based around the number of ions detected and the quality of the data for the peptide matched. Only peptides with a score above 20 were validated.

### 2.8.2 Intact Mass Analysis

Samples were reconstituted in 40 µl of 0.1% formic acid/5% ACN. 10 µl sample was loaded onto the Sciex Exion LC and a 5-minute RP gradient was run. Buffer A was 0.1% formic acid and buffer B was 0.1% formic acid in 100% ACN. A Phenomenex Jupiter 5 µm, C4, 300A 50x2.1 mm column was used with a gradient of 5% B increasing to 45% B over 3 minutes and then a 95% B wash and re-equilibration at 5% B with a flow of 500 µl. The flow from the column passed into the Sciex X500B mass spectrometer in positive ion mode for collection. The source was set to 400°C, 5500V with gas at 50 psi to enable ionisation. A TOF mass window of 500-3000 Da scanning at 0.5 seconds was collected. A positive calibration mix was used to calibrate the X500B and the

error for the experiment was estimated at 0.5 Da. The TIC was deconvoluted using BioToolKit software.

## 2.9 Bioinformatic Analysis

### 2.9.1 BLASTp Searches

Protein searches were conducted using the NCBI protein basic local alignment search tool (BLASTp) to search for similar protein sequences to those input into the search box. A screenshot of the standard protein BLAST (BLASTp) web client used can be found in Figure 2.3.

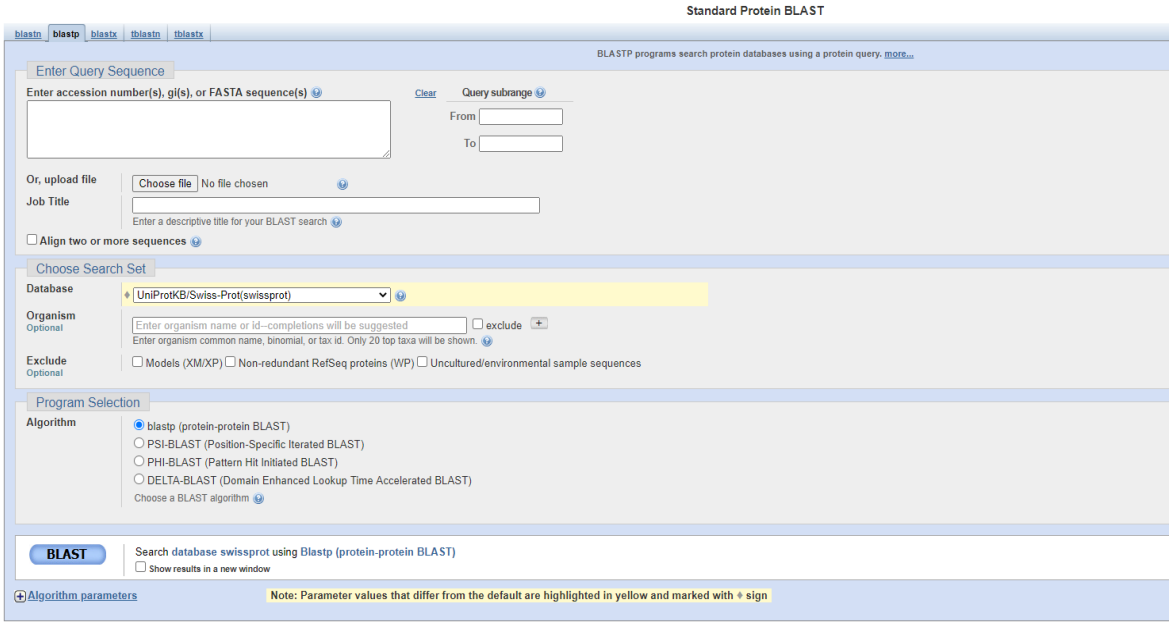



*Figure 2.3 The BLASTp suite web client.*

*The web client may be found at [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). The UniProtKB/SwissProt database was selected as the chosen database.*

### 2.9.2 Multiple Sequence Alignments

The FASTA format of each of the selected cardiotoxins were collected into a document and multiple sequence alignments were performed using The European Bioinformatics Institute (EMBL-EBI) Clustal Omega web tool (Sievers *et al.*, 2011) found at https://www.ebi.ac.uk/Tools/msa/clustalo/. The protein tool was used with the pasted FASTA sequences. The output selected was Pearson/FASTA as shown in Figure 2.4 and the resulting alignment file was saved as a FASTA format file.

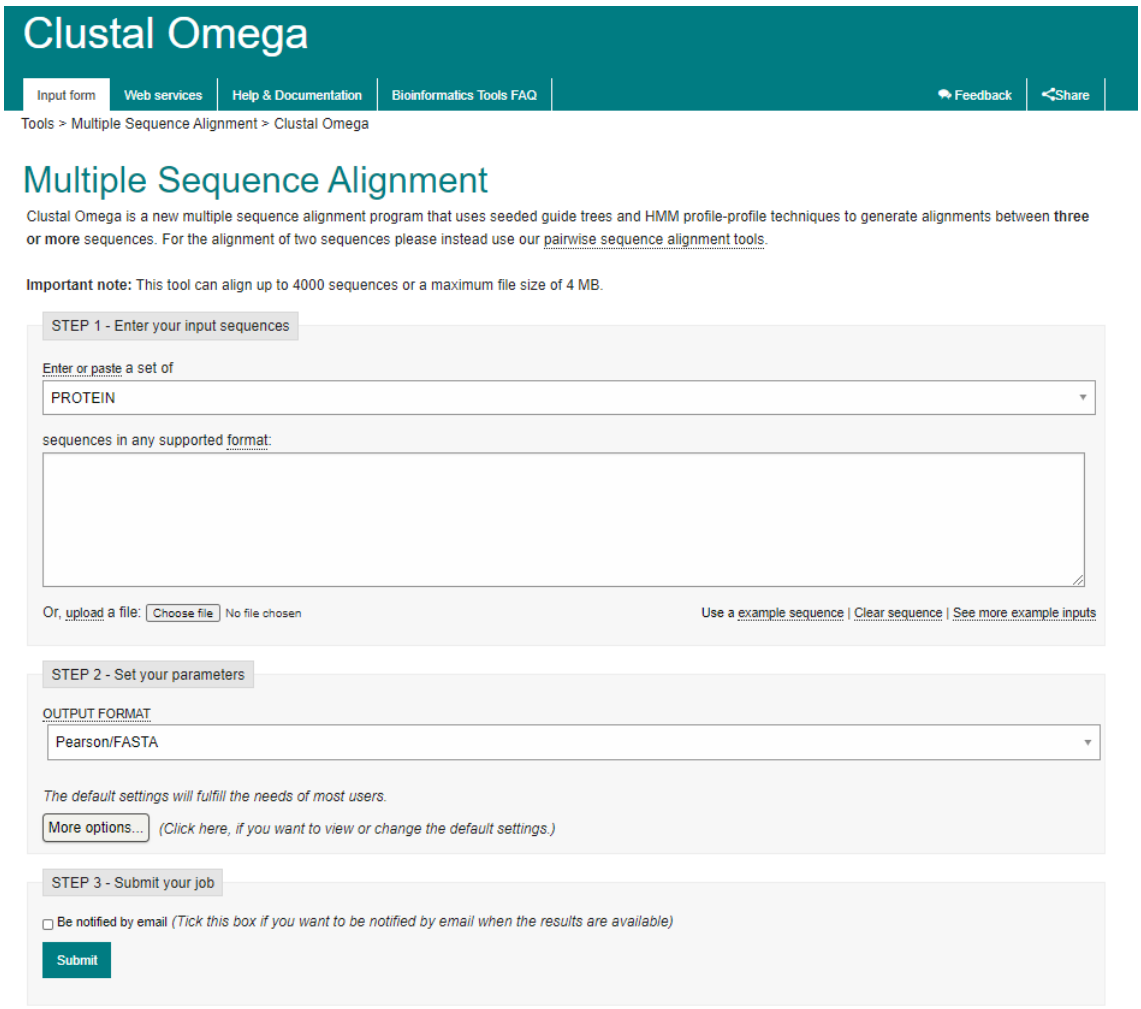



*Figure 2.4. The Clustal Omega Multiple Sequence Alignment web client.*

*The web client may be found at https://www.ebi.ac.uk/Tools/msa/clustalo/ . The settings used for the multiple sequence alignments are shown in the screenshot above.*

The aligned FASTA file was imported into Molecular Evolutionary Genetics Analysis version 7 (MEGA7) (downloaded at www.megasoftware.net) to display the alignments (Kumar *et al.*, 2016). Screenshots of sequences aligned in MEGA7 were used for figures throughout Chapter 5.

### 2.9.3 Structure Activity Relationships (SAR)

The protein sequences of interest were collected in FASTA format and aligned as described in 2.9.2. The WebProAnalyst web application (http://wwwmgs.bionet.nsc.ru/mgs/programs/panalyst/) was used to perform SAR analysis (Ivanisenko *et al.*, 2005). The FASTA format aligned sequences were pasted into the 'sequence alignment' text box. Into the 'protein activity values' box, the $LD_{50}$ or $IC_{50}$ values were pasted in space delimited format. Analysis was performed for SADC/SACC from fragment position 1 to 100 with a slide window of 1. A screenshot of the WebProAnalyst web client may be found in Figure 2.5. The output was pasted into a Microsoft® Excel® for Office 365 document (Microsoft Corporation, Washington, USA) and the resulting SADC values were coloured according to their value between 0-1.

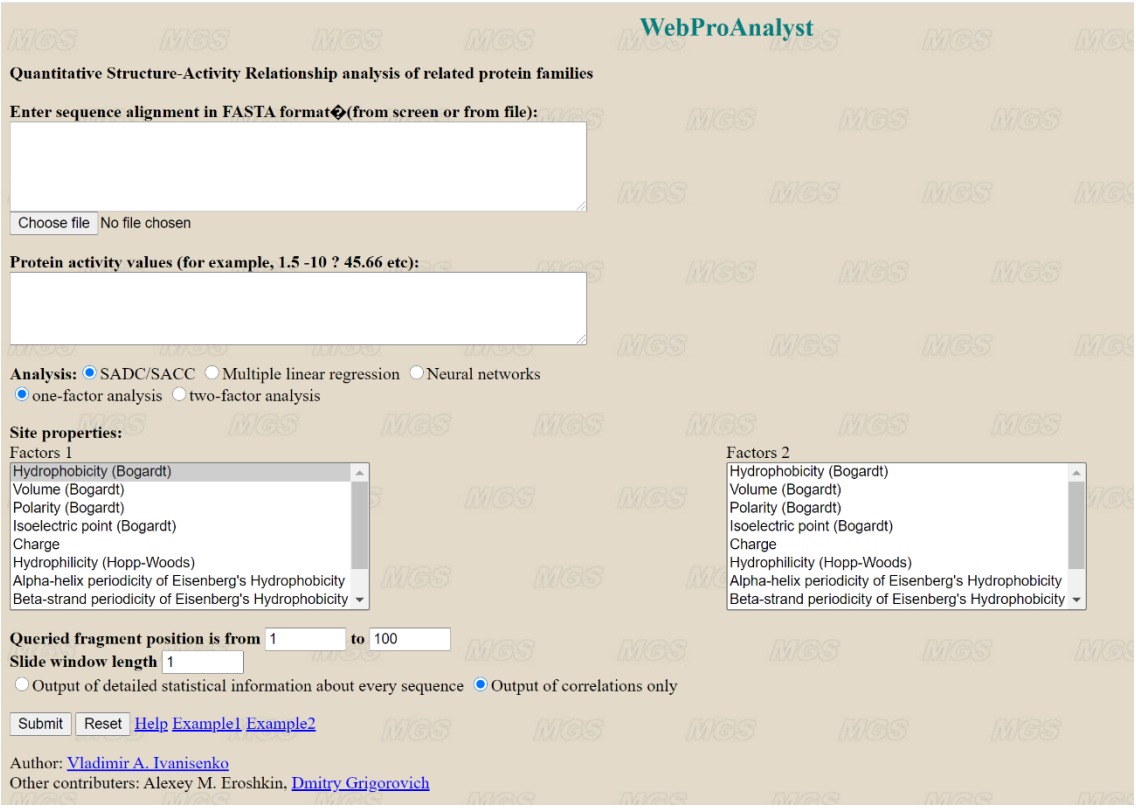



*Figure 2.5. The WebProAnalyst webclient used for SAR analysis.*

*The webclient may be found at: http://wwwmgs.bionet.nsc.ru/mgs/programs/panalyst/.*

### 2.9.4 3D Structure Modelling

Three-dimensional structure modelling was performed using The PyMOL Molecular Graphics System, Version 2.0 (Schrödinger, LLC). Biological assemblies were imported into PyMOL by accessing File->Get PDB…, typing the desired PDB code and selecting 'Download' as shown in Figure 2.6.

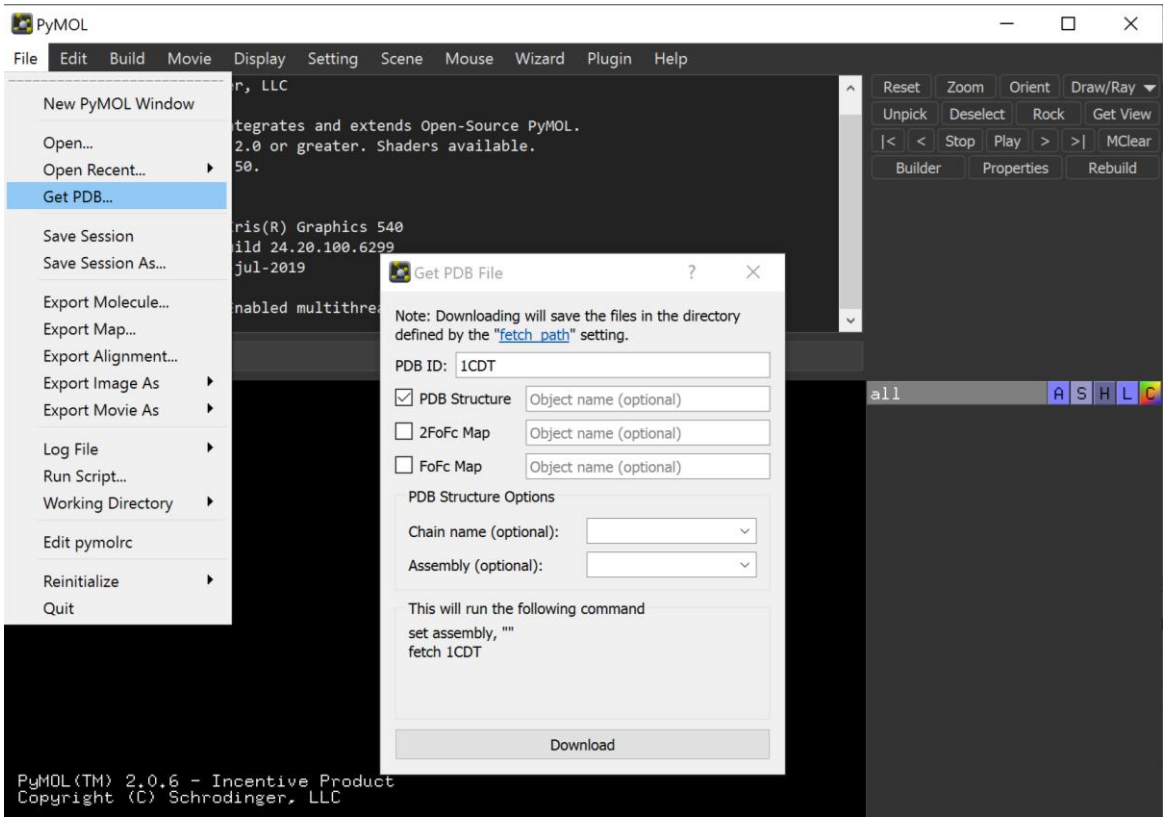



*Figure 2.6. Importing PDB files into PyMOL using the 'Get PDB…' function.*

*Any valid PDB code can be entered into the PDB ID box (in this case '1CDT' was chosen) and imported into the software by selecting 'Download'.*

To make the 3D structures easier to visualise, water molecules and ligands (in this case, phosphorus molecules) were removed from the structure. For structure 1CDT, one of the dimer units was removed to make visualisation clearer and labelling of significant residues easier to distinguish.

Images were exported using the 'Export Image As' function as may be seen on the 'file' menu in Figure 2.6. Images were exported as ray-traced png image files with opaque backgrounds.

## 2.10 qPCR

### 2.10.1 qPCR Optimisation

The following sections detail the experimental procedures used prior to qPCR experimentation investigating the effect of venom fractions on the cancer cell lines to design and optimise the qPCR assays and ensure good quality, reliable results.

#### *2.10.1.1 Primer Design*

Primers were designed using the following parameters:

- Amplicon of 70-150 bp (up to 200 bp in some cases)[1]
- Primer pair annealing temperature of approximately 60 °C[1]
- Primer length of approximately 20 bp
- No more than two consecutive CC or GG runs anywhere in the primer
- No more than two C/Gs in last 5 bp of each primer
- Primer pairs should be separated by at least one intron in corresponding genomic DNA

[1] =Suggested in the Bio-Rad SSoAdvanced™ Universal SYBR® Green Supermix leaflet.

Where primers were present and met the majority of primer design parameters, primers were taken from the RTprimerDB (http://www.rtprimerdb.org/). Otherwise, primers were designed through the NCBI Primer-BLAST tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Details of the selected primers can be found in Table 2.11.

25 nm of each dry, desalted primer was ordered from the Invitrogen custom DNA oligos website (https://www.fishersci.co.uk/shop/oligoPunchout.do). Primer stocks were made by adding appropriate volume of Invitrogen ambion TE buffer (Fisher, Loughborough, UK) to lyophilised primers to give 50 µM concentration. Working stocks were made by diluting the 50 µM stocks to 5 µM with nuclease free water and storing in aliquots at -20°C.

#### *2.10.1.2 RNA Purification and Quality Assessment*

QIAGEN RNeasy Mini Kit (Qiagen, Hilden, Germany) was used to purify RNA from up to $1x10^7$ BxPC-3 and SW620 cells according to manufacturer's instructions. On-column DNA digestion was performed using QIAGEN RNase-free DNase set (Qiagen, Hilden, Germany) according to manufacturer's instructions. RNA concentration and quality of samples were measured using the DS-11 spectrophotometer (DeNovix, Wilmington, USA) using the "RNA" application. The 260/280 and 260/230 ratios were measured to monitor potential contamination or lack of purity in the samples. 10 mm absorbance values were plotted against wavelength to observe the peak wavelengths for each sample.

#### *2.10.1.3 RNA to cDNA Reverse Transcription and Quality Assessment*

RNA was converted to cDNA using iScript™ Select cDNA synthesis kit (Bio-Rad, California, USA). Each reaction contained 4 µl 5X iScript select reaction mix, 2 µl random primers, 1 µl iScript reverse transcriptase and a combination of RNA sample and water to a total of 13 µl. Reactions contained 1 µg RNA unless the concentration was too low to allow this, in which case 13 µl RNA (0 µl water) was added to the reaction. The reverse transcription reactions were performed in a Techne $^{3}$Prime PCR machine (Techne, Staffordshire, UK). Samples were incubated at 25°C for 5 minutes, 42°C for 30 minutes, then heat inactivated at 85°C for 5 min.

The concentration and quality of the reverse transcribed cDNA were measured using the DS-11 spectrophotometer (DeNovix, Wilmington, USA) using the "ssDNA" application. The 260/280 and 260/230 ratios were measured to monitor potential contamination or lack of purity. 10 mm absorbance values were plotted against wavelength to observe the peak wavelengths for each sample.

*Table 2.11. Details of primers designed for genes included in gene study.*

| Target gene | RTPrimer DB ID | Primer | Sequence (5'-3') | Annealing temp (°C) | Amplicon size (bp) | Accession no. |
|---|---|---|---|---|---|---|
| Beta-Actin | 1 | F | CTGGAACGGTGAAGGTGACA | 60 | 140 | NM_001101.3 |
| | | R | AAGGGACTTCCTGTAACAATGCA | 60 | | |
| Tubulin | | F | GAAATCGTGCACATCCAGGC | 59.9 | 144 | NM_001293213 |
| | | R | GTACACAGAGATGCGGTCCA | 59.47 | | |
| GAPDH | 3 | F | TGCACCACCAACTGCTTAGC | 60 | 87 | NM_002046.5 |
| | | R | GGCATGGACTGTGGTCATGAG | 60 | | |
| 18S | 3879 | F | CGGACAGGATTGACAGATTG | 60 | 83 | NR_145820.1 |
| | | R | CAAATCGCTCCACCAACTAA | 60 | | |
| EGFR | 1906 | F | GGAGAACTGCCAGAAACTGACC | 60 | 106 | NM_005228.4 |
| | | R | GCCTGCAGCACACTGGTTG | 60 | | |
| HER2/ErbB2 | 1926 | F | CCAGCTGGCTCTCACACTG | 60 | 74 | NM_004448.3 |
| | | R | AGCCCTTACACATCGGAGAAC | 60 | | |
| HER3/ErbB3 | | F | AGGGACCGAGATGCTGAGAT | 60.11 | 134 | NM_001982.3 |
| | | R | GGAGCACAGATGGTCTTGGT | 59.67 | | |
| HER4/ErbB4 | 1931 | F | GGCAGATGCTACGGACCTT | 60 | 64 | NM_001042599.1 |
| | | R | CTGAGCAGCCTCCAGCAC | 60 | | |
| AKT1 | 4395 | F | CCCGGCCCAACACCTT | 60 | 84 | NM_001014431.1 |
| | | R | GCTCCTCAGGAGTCTCCACATG | 60 | | |
| KRAS | | F | ACACAAAACAGGCTCAGGACT | 59.79 | 118 | NM_033360 |
| | | R | TGTCGGATCTCCCTCACCAA | 60.25 | | |
| SMAD4 | 4805 | F | GCTGCTGGAATTGGTGTTGATG | 60 | 108 | NM_005359.5 |
| | | R | AGGTGTTTCTTTGATGCTCTGTCT | 60 | | |
| TP53 | | F | AGTCTAGAGCCACCGTCCAG | 60.68 | 145 | NM_001126114.2 |
| | | R | CTCCTCCATGGCAGTGACC | 59.78 | | |
| CDKN2A | 1708 | F | CATAGATGCCGCGGAAGGTC | 60.95 | 139 | NM_001195132.1 |
| | | R | GCCAGCTTGCGATAACCAAA | 59.47 | | |
| MAPK14 | | F | CCGCTTATCTCATTAACAGGATGC | 59.85 | 135 | NM_001315.2 |
| | | R | TCTTCTCCAGCAAGTCGACAG | 59.73 | | |
| Caspase 3 | | F | AGAACTGGACTGTGGCATTGAG | 60 | 191 | XM_011532301.1 |
| | | R | GCTTGTCGGCATACTGTTTCAG | 60 | | |
| Caspase 7 | | F | GAGCACGGAAAAGACCTGGA | 59.97 | 137 | NM_001227.3 |
| | | R | GTGAGCATGGAGACCACACA | 59.96 | | |
| MMP-9 | 3495 | F | CCTGGAGACCTGAGAACCAATC | 58 | 79 | NM_004994.2 |
| | | R | CCACCCGAGTGTAACCATAGC | 58 | | |
| WNT1 | | F | ACTGCACGAGTGTCTGTGAG | 59.97 | 85 | NM_005430 |
| | | R | TGCTAGCGAGTCTGTTTGGG | 60.04 | | |
| NOTCH1 | | F | CAGCCCGGTGAGACCTG | 59.35 | 193 | NM_017617 |
| | | R | CACAGCTGCAGGCATAGTCT | 60.11 | | |
| Bax | 1539 | F | CCTTTTCTACTTTGCCAGCAAAC | 60 | 148 | NM_138761.3 |
| | | R | GAGGCCGTCCCAACCAC | 60 | | |
| Bad | | F | TCCGGAGGATGAGTGACGAG | 60.75 | 128 | NM_004322 |
| | | R | ATCCCACCAGGACTGGAAGA | 59.88 | | |
| Bcl-2 | 582 | F | TCCGCATCAGGAAGGCTAGA | 60 | 113 | NM_000633.2 |
| | | R | AGGACCAGGCCTCCAAGCT | 60 | | |
| STAT3 | 4814 | F | GATCCAGTCCGTGGAACCAT | 60 | 73 | NM_003150.3 |
| | | R | ATAGCCCATGATGATTTCAGCAA | 60 | | |
| TGF-β | | F | TGATGTCACCGGAGTTGTGC | 60.6 | 125 | NM_000660.6 |
| | | R | TGAACCCGTTGATGTCCACTT | 59.86 | | |

### 2.10.2 qPCR Experimental Parameters

The general recipes for each 10 µl qPCR reaction are shown in Table 2.12 below. Both options give the same overall proportions, but 'reaction 2' was preferred when the INTEGRA VOYAGER multichannel pipette (INTEGRA Biosciences AG, Switzerland) was available. This pipette has the capacity to draw up and dispense multiple reactions at once which reduces the risk of contamination, but has a minimum dispense volume of 2 µl. Master mixes were prepared for each primer pair with a 10% excess volume. For each reaction, 5 µl of BioRad SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad Laboratories, California, USA) was used as the source of dNTPs, DNA polymerase, $MgCl_2$, SYBR® Green I and ROX normalisation dyes. To this, 5 µM stocks of forward and reverse primers were added to each master mix to give a final concentration of 300 nM of each primer. Nuclease-free water was added to each reaction to make a final volume of 10 ul (volumes found in Table 2.12). Master mixes were mixed gently, then centrifuged briefly immediately before use. Any mixture to be used for future plates remained in the tube and was covered with aluminium foil and stored at 4°C and used within 24 hours.

To reduce the risk of potential contamination, all qPCR reactions were prepared in a Thermo-Scientific MSC-Advantage Class II Biological Safety Cabinet (Thermo Fisher Scientific, Massachusetts, USA) and separate pipettes were reserved solely for qPCR reagents and cDNA. The laminar flow hood was cleaned thoroughly with 5% detergent solution followed by Nucleoclean™ decontamination spray (Millipore Corporation, Massachusetts, USA) and 60 minutes of exposure to UV light. This was done each time between RNA purification, reverse transcription steps and qPCR set up. The laminar flow hood was additionally sprayed with Nucleoclean™ decontamination spray and exposed to UV light for 60 min between preparation of each qPCR plate. Whilst plating, all other qPCR plating steps were completed before the cDNA tubes were opened to prevent contamination of stock solutions.

A no template control (nuclease-free $H_2O$) triplicate was included for each primer pair tested for each plate to monitor for contamination. Two reference genes (GAPDH and β-Actin) were used as inter-plate calibrator genes in the 2D Venom qPCR Assays to ensure plates could be cross-compared, if desired.

*Table 2.12. General recipe for one 10 µl qPCR reaction.*

*Each reaction contained 10 ng template cDNA. Reaction 1 was the recipe used when the template cDNA conc. was 10 ng/µl, Template 2 was the recipe used when the template cDNA conc. was 5 ng/µl.*

| | Reaction 1 Volume (µl) | Reaction 2 Volume (µl) |
|---|---|---|
| SsoAdvanced™ Universal SYBR® Green Supermix | 5 | 5 |
| Forward primer (5 µM stock) | 0.6 | 0.6 |
| Reverse primer (5 µM stock) | 0.6 | 0.6 |
| Nuclease free $H_2O$ | 2.8 | 1.8 |
| Template cDNA (10 ng/µl) | 1 | - |
| Template cDNA (5 ng/µl) | - | 2 |
| Total | 10 | 10 |

The general qPCR cycling conditions were: 95°C/10min; 40 cycles of cDNA melting at 95°C/15s and annealing and extension at 60°C/60s. A melt curve was performed after the 40 cycles between 65-95°C, increment 0.5°C, 5s. A diagrammatic representation of this process can be found in Figure 2.7.

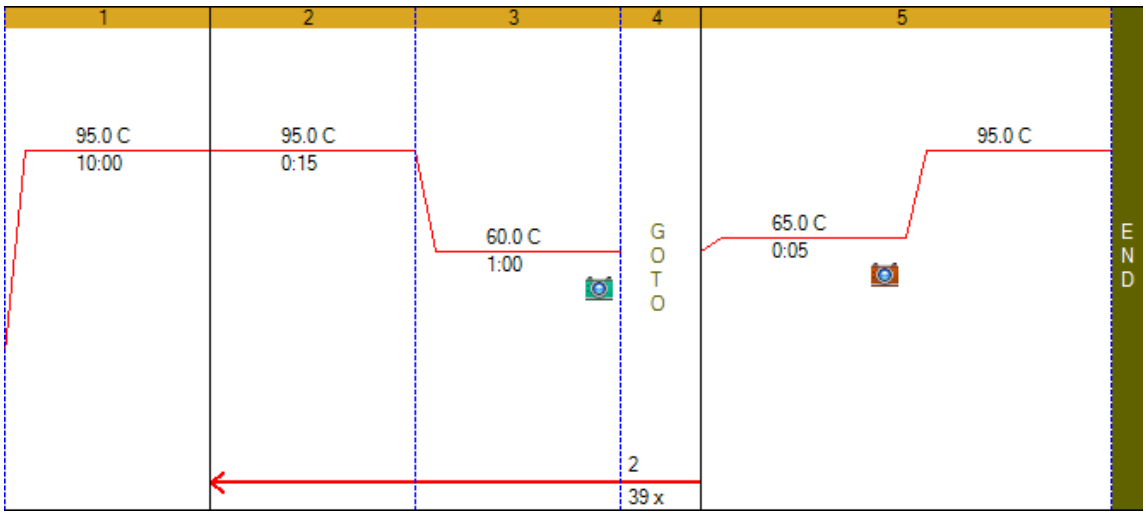



*Figure 2.7. General qPCR running protocol.*

*The protocol contains 40 cycles with a denaturation temperature of 95°C and an annealing temperature of 60°C, as optimised for the designed primers. Following amplification, a melt curve was performed between 65-95°C with an increment of 0.5°C every 5 seconds. Copied from the actual protocol run in BioRad CFX Manager software.*

#### 2.10.2.1 Primer QC

Each set of primer pairs (the forward and reverse primer) for each of the selected genes were tested to ensure suitability for downstream qPCR application. This involved harvesting RNA from BxPC-3 and SW620 cells, reverse transcription and running the cDNA with each primer pair and with a no template control (NTC) using nuclease free water instead of cDNA. The qPCR product was then run on a 2% agarose gel and the bands assessed. To pass the quality assessment, the qPCR products needed to have amplification of a single, correctly sized product on the agarose gel.

#### 2.10.2.2 Agarose Gels

2% agarose gels were poured by combining 2 g agarose powder with 100 ml 1xTAE buffer and heating in the microwave for approximately 1 min or until a rolling boil was reached. The agarose solution was allowed to cool briefly, then 10 µl SYBR green I was added and gently swirled. Gels were poured into Fisher Scientific Mini Plus submarine gel tanks (Thermo Fisher Scientific, Massachusetts, USA), a 16-well gel comb inserted and allowed to set. The set gels were covered with 1xTAE buffer until completely submerged then the combs were carefully removed. The DNA ladder used for size comparison of qPCR products was the Thermo Scientific™ GeneRuler Low Range DNA Ladder (Thermo Fisher Scientific, Massachusetts, USA) which allows size estimation of small sized (down to 25 bp) DNA fragments. 3 µl of DNA ladder was used per lane. Each 10 µl qPCR products had 2 µl of Thermo Scientific™ 6X TriTrack DNA loading dye (Thermo Fisher Scientific, Massachusetts, USA) added. After gentle mixing, 5 µl of sample/dye were loaded per sample. Samples were loaded and gels were run at 80V for 90 minutes or until approximately 80% down the gel. Gels were visualised on the Bio-Rad ChemiDoc™ Touch Imaging System (Bio-Rad Laboratories, California, USA) using the SYBR green setting under nucleic acid gel visualisation.

#### 2.10.2.3 Melt Curves

Melt curves were performed for each sample following the qPCR amplification stage. The details of the melt curve can be found in Figure 2.7.

#### 2.10.2.4 Temperature Gradient Analysis of GAPDH

To assess the optimum temperature for annealing of GAPDH primers, temperatures of between 55.3-67.3°C were selected and a gradient protocol run. qPCR cycling conditions were: 95°C/10min; 40 cycles of cDNA melting at 95°C/15s and annealing and extension at the relevant temperature for 60s. A melt curve was performed after the 40 cycles between 65-95°C,

increment 0.5°C, 5s. A diagrammatic representation of the protocol and the plate layout detailing the temperatures studied can be found in Figure 2.8.

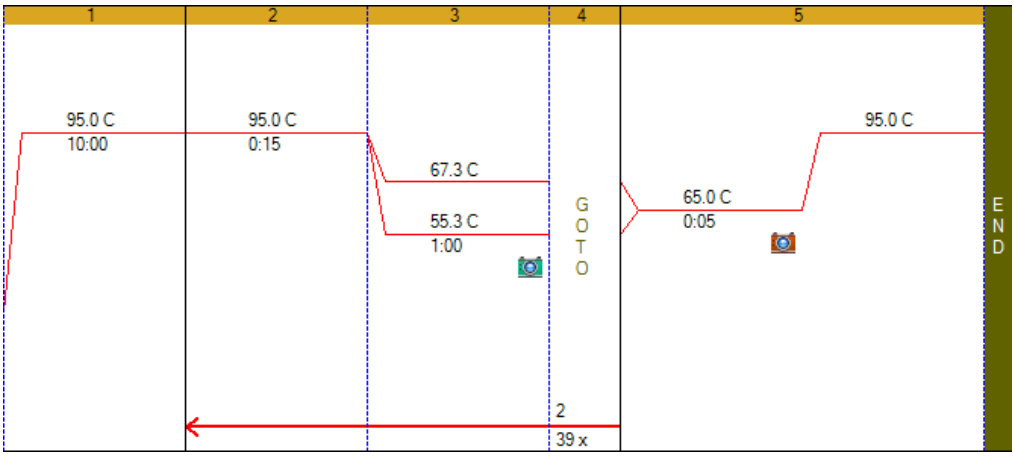



*Figure 2.8. The qPCR cycling conditions for the GAPDH temperature gradient analysis.*

### *2.10.2.5 Calibration Curves/Amplification Efficiency Assessment*

#### 2.10.2.5.1 Comparing the amplification efficiency of RNA harvested from 24- versus 96-well plates

SW620 cells were plated into 24-well Costar® (Corning, New York, USA) and 96-well Falcon® (Corning, New York, USA) plates at a concentration of $2x10^6$ and $2x10^5$ cells per well respectively and incubated overnight at 37°C, 5% $CO_2$. The cells from each sample were harvested using 350 µl of Buffer RLT supplied with the QIAGEN RNeasy Mini Kit (Qiagen, Hilden, Germany). Due to volume restrictions of the wells, the 96-well sample were harvested by adding 175 µl to each well, transferring to an Eppendorf tube, then adding a further 175 µl to the well and transferring to the same Eppendorf tube. This process is shown diagrammatically in Figure 2.9a. RNA samples were purified, then reverse transcribed using random primers.

For each sample, a ten-fold serial dilution starting at 70 ng/µl was performed to give a 5 $log_{10}$ serial dilution series. 1 µl of each sample was plated into each qPCR reaction according to Table 2.12 (Reaction 1) to give a cDNA range from 7 pg- 70 ng. Each sample was run in triplicate with additional no reverse transcription control (non-transcribed RNA only), RNA diluted to 10 µg/ml and no template control (water only, no cDNA, also in triplicate). The general qPCR protocol was run for each plate.

#### 2.10.2.5.2 Comparing the amplification efficiency of RNA harvested from 96- versus 384-well versus resazurin treated 384-well plates

A similar amplification efficiency assay was performed using BxPC-3 cells and on a larger scale. This time the samples being compared were from 96- and 384-well plates with an additional set of samples for BxPC-3 cells plated in 384-well format, previously treated with resazurin. The

Falcon® flat bottom TC-treated 96-well plate (Corning, New York, USA) was seeded with $1x10^5$ cells per well. The cell culture treated, flat bottom 384-well plate (Greiner Bio-One, Kremsmünster, Austria) (including resazurin treated samples) was seeded at $2.5x10^4$ cells per well. Both plates were incubated overnight at 37°C, 5% $CO_2$. To half of the 384-well samples, 25 µl of 160 µM resazurin solution was added and the cells incubated for 2 h, then removed and discarded. The cells were washed twice with PBS to remove any excess resazurin. All samples were harvested in 350 µl buffer RLT supplied with the QIAGEN RNeasy Mini Kit (Qiagen, Hilden, Germany). For each of the three 96-well sample, 175 µl buffer RLT was added to each well then transferred to an Eppendorf tube. A further 175 µl buffer RLT was added to each well and transferred to the same Eppendorf tube. For the 384-well, 60 µl of buffer RLT was added per well and three wells were transferred to an Eppendorf tube per biological replicates (3 biological replicate tubes each containing 3x384-wells). To this tube, a further 170 µl buffer RLT was added to give a total of 350 µl per biological repeat. Resazurin treated samples were harvested in the same manner as 384-well samples. This process is shown diagrammatically in Figure 2.9b.

The RNA was purified and quality assessed according to section 2.10.1.2 and reverse transcription of the purified DNA samples was performed according to section 2.10.1.3. For each sample, 13 µl of RNA was reverse transcribed due to low sample concentrations. The cDNA concentration and quality were assessed according to section 2.10.1.3.

For each biological repeat, a ten-fold serial dilution starting at 70 ng/µl was performed to give a 5 $\log_{10}$ serial dilution series. 1 µl of each sample was plated into each qPCR reaction according to Table 2.12 (Reaction 1) to give a cDNA range from 7 pg- 70 ng. Each biological replicate was run in triplicate with additional no reverse transcription control (non-transcribed RNA only) and no template control (water only, no cDNA, also in triplicate). The general qPCR protocol was run for each plate. The three sample conditions (96-well, 384-well and resazurin treated BxPC-3 cells) were tested with three technical repeats of three biological replicates with 18S and TGF-B primer pairs.

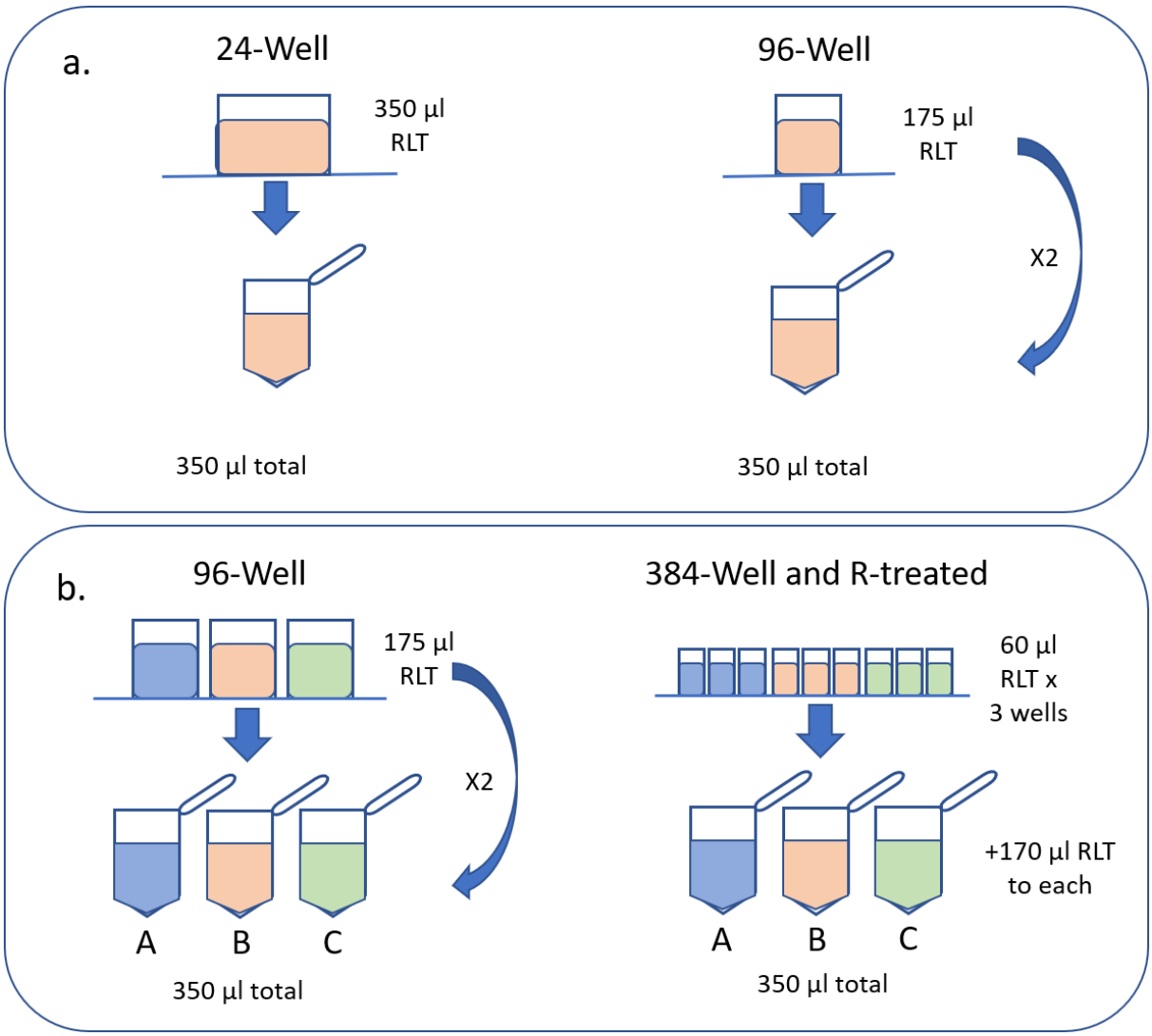



*Figure 2.9. The RNA harvesting process from 24, 96 and 384 well plates.*

*Diagrammatic representation from a. 24- and 96-well plates for the 24- to 96-well optimisation stage and b. 96-well, 384-well and R-treated plates with biological triplicates for the 96-to 384-well and R-treated samples optimisation stage. Each well size ended up with 350 µl sample harvested in buffer RLT ready for RNA purification.*

#### *2.10.2.6 Assessing the amplification efficiencies of GAPDH and Tubulin in BxPC-3 cells*

The 6 hour time-point BxPC-3 untreated control cDNA acquired from the 2D fraction testing assay (section 2.10.3 below) was 10-fold serially diluted from 700 ng/µl to 7 pg/µl. The amplification efficiencies of GAPDH and Tubulin primers were assessed between 700 ng to 7 pg per reaction. The samples were plated in triplicate with volumes according to Table 2.12 (Reaction 2) plus no reverse transcription (RNA only) and no template (water only) controls in triplicate.

#### *2.10.2.7 qPCR Data Processing*

qPCR data were processed using Bio-Rad CFX Manager 3.1 (Bio-Rad Laboratories, California, USA). The mean and standard deviations of the Cq values were calculated from the raw data collected. Agarose gel images were processed using Bio-Rad Image Lab™ software version 6.0.0

build 25 (Bio-Rad Laboratories, California, USA). Calibration curves were plotted from raw data using Graphpad Prism 9.0.0 (GraphPad Software, San Diego, USA). $R^2$ values were taken from the line of best fit and amplification efficiencies calculated from the gradient of the same line.

Double-delta Cq ($2^{-\Delta\Delta Cq}$) analysis was performed manually using Microsoft® Excel® for Office 365 (Microsoft Corporation, Washington, USA) with GAPDH as the reference gene (Livak and Schmittgen, 2001). Automatic analysis on Bio-Rad CFX Manager 3.1 (Bio-Rad Laboratories, California, USA) was used to help assess appropriateness and stability of reference genes.

### 2.10.3 2D Venom Fraction qPCR Testing

BxPC-3 and SW620 cells were plated into separate Costar® 24-well plates (Corning Inc., New York, USA) at a density of $1x10^6$ cells per well and incubated overnight at 37°C, 5% $CO_2$. Fraction samples (N.nub_i17r1, N.nct_i18r2 and N.pal_i17r2) were added to the cells at a concentration of 40 µg/ml. Following treatment, the fractions were removed and discarded and replaced with appropriate cell media. The 6h samples were harvested 4h later (6 h since first exposure) and the 24h samples were harvested 18 h later (24 h since first exposure). All three fraction samples treated at 40 µg/ml plus the untreated control were harvested by adding 350 µl buffer RLT then transferring to a 1.5 ml microcentrifuge tube. The RNA was purified and quality assessed according to section 2.10.1.2. RNA was reverse transcribed to cDNA and quality assessed according to section 2.10.1.3. qPCR was run according to section 2.10.2 using all the primers detailed in Table 2.11.

A general plate layout was used to make qPCR plating and data analysis more efficient and reduce the risk for erroneous pipetting. This plate layout can be found in Figure 2.10. Six genes were compared per plate including GAPDH and Actin which were used as inter-plate calibrator genes. GAPDH acted as the reference gene for every plate so that the expression of each test gene could be compared to the GAPDH expression levels.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | GAPDH | | | Test Gene 1 | | | Test Gene 3 | | | GAPDH | | |
| B | GAPDH | | | Test Gene 1 | | | Test Gene 3 | | | Actin | | |
| C | GAPDH | | | Test Gene 1 | | | Test Gene 3 | | | Test Gene 1 | | |
| D | GAPDH | | | Test Gene 1 | | | Test Gene 3 | | | Test Gene 2 | | |
| E | Actin | | | Test Gene 2 | | | Test Gene 4 | | | Test Gene 3 | | |
| F | Actin | | | Test Gene 2 | | | Test Gene 4 | | | Test Gene 4 | | |
| G | Actin | | | Test Gene 2 | | | Test Gene 4 | | | | | |
| H | Actin | | | Test Gene 2 | | | Test Gene 4 | | | | | |

*Figure 2.10. The general layout for the 2D Fraction qPCR Assay.*

*The cDNA used came from the following treatments: orange = untreated control, blue = N.nub_i17r2 treated, yellow = N.pal_i17r2 treated, green = N.nct_i18r2 treated. Grey = no template controls (water only).*

Data were collected automatically during the qPCR cycling process using Bio-Rad CFX Manager 3.1 (Bio-Rad Laboratories, California, USA).

Storage conditions for the 2D fraction exposed nucleic acid samples were as follows: the RNA samples were frozen at -20°C immediately after RNA purification on the same day as sample harvesting. The reverse transcription of all samples was performed together, 3 (24h samples) or 4 (6h samples) days after RNA purification. The remaining RNA was immediately stored at -80°C. 5 µg of each cDNA sample was diluted to 5 ng/µl ready for qPCR reactions, aliquoted and stored at -20°C. The remaining concentrated cDNA was stored at -80°C.

# Chapter 3 - Screening of Whole Cobra Venoms to Investigate their Effect on Pancreatic and Colorectal Cancer

## 3.1 Introduction

### 3.1.1 Why Cobra Venoms?

Hundreds of thousands of venomous species possess venom which contain bioactive molecules which could become potential cancer treating drug entities. It is therefore important to be selective when deciding which venoms to study. Although some venom labs screen venom from varied species to maximise molecule diversity (Mccullough *et al.*, 2020), this is most effective in target-based screening where the mechanism of the target is already known. This allows identification of diverse molecules acting on the same target which may be deconvoluted with follow up experiments.

As this study was a phenotypic screen without a specific target, this approach may have found venoms having varied targets and mechanisms of activity which would be difficult to effectively deconvolute later. Cobra venoms were chosen to narrow down the spectrum of different bioactive molecule classes, but still provide diversity within each bioactive molecule class. This allowed more effective structure activity relationship studies to be performed as cobra venoms contain many similarly structured proteins with specific amino acid alterations leading to a variety of toxicities.

### 3.1.2 The Use of The Resazurin Reduction Assay as a HTS Tool

Resazurin is a redox dye which can be used to assess the viability and metabolic activity of cultured cells. Also known as alamarBlue®, PrestoBlue®, AB Assay, Vybrant, UptiBlue and CellTiter-Blue®, resazurin has been used to study cell viability and compound toxicity for a range of different applications over the past five decades. It was originally used to detect contamination of dairy products (Erb and Ehlers, 1950) but has more recently been utilised for a range of applications including cell response to radiation (Anoopkumar-Dukie *et al.*, 2005), quantitative analysis of migration and invasion of carcinoma cells (Al-Nasiry *et al.*, 2007), semen fertilising capability (Zrimšek *et al.*, 2004, Erb and Ehlers, 1950), investigation of environmental toxins (Zare *et al.*, 2015, Mcnicholl *et al.*, 2007) and has been use to screen a range of antibiotics (Sarker *et al.*, 2007), amongst other applications. Resazurin even has application in assessing the effect of drug treatments on viruses when cultured with host cells (Che *et al.*, 2009).

In cells which are viable and actively metabolising, resazurin is reduced to resorufin and eventually dihydroresorufin (Figure 3.1). The total amount of conversion can be measured colourmetrically or fluorometrically as the weakly fluorescent resazurin is converted to the pink coloured, highly fluorescent compound resorufin.

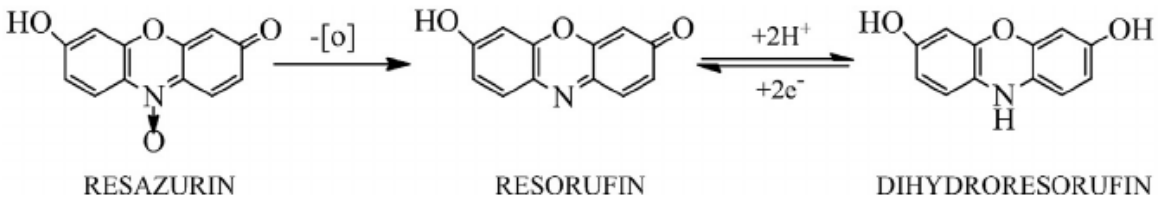



*Figure 3.1. Resazurin reduction reaction.*

*Resazurin may be irreversibly reduced to resorufin and may then be reversibly reduced further to dihydroresorufin. Figure from (Chen et al., 2015).*

The resazurin conversion rate is directly proportional to the number of viable cells present, allowing accurate estimation of the viability of cells in the assay (Shiloh *et al.*, 1997, O'brien *et al.*, 2000). The exact mechanism of resazurin reduction within living cells is not yet fully understood. Some researchers conclude that resazurin is metabolised by mitochondrial enzymes within the cell (Zhang *et al.*, 2004), whilst others postulate that it could be due to oxygen consumption during metabolism (O'brien *et al.*, 2000). Furthermore, once converted to resorufin, further reduction may occur, converting resorufin to dihydroresorufin, a non-fluorescent compound. This could lead to a plateau or even decrease in overall fluorescence signal generated by the cells, despite active metabolism still taking place (Gong *et al.*, 2020).

Resazurin and its products are relatively non-toxic to cells and remain stable in cell media, making it an ideal reagent for measurements over many time points and for retrieval of cells for other assays if required (O'brien *et al.*, 2000). It should be noted, however, that at high concentrations and/or prolonged exposure, resazurin has been found to be toxic to cells (Pace and Burg, 2015, Gong *et al.*, 2020). Assays utilising resazurin can compare the ability of cells treated with a test compound to reduce resazurin against the control cells. This can be used to measure cell toxicity of drug compounds, poisons, environmental toxins, etc.

Resazurin has been referred to in the literature by a range of different brand names including alamarBlue® (Rampersad, 2012), PrestoBlue® (Emter and Natsch, 2015) and CellTiter-Blue®(Promega, 2016). A study by Borra and colleagues (Borra *et al.*, 2009) compared resazurin and alamarBlue® using a colorimetric method and concluded that they were the same compound and that the branded alamarBlue® contained approximately 560 µM resazurin. O'Brien *et al*. (2000) performed a similar study using mass spectrometry and NMR spectroscopy

and also concluded that alamarBlue® and resazurin are the same compound. They found that alamarBlue® contained approximately 440 µM resazurin.

Resazurin (or alamarBlue or PrestoBlue) has been shown to be comparable and equivalent to 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), another dye commonly used to measure cell viability (Emter and Natsch, 2015, Hamid *et al.*, 2004).

The results of a resazurin assay may be interpreted visually (Sarker *et al.*, 2007), using absorbance data (Zrimšek *et al.*, 2004), fluorescence data or RGB signals generated by a digital CCD camera (Borra *et al.*, 2009). The absorbance and fluorescence spectra for resazurin and resorufin are displayed in Figure 3.2.

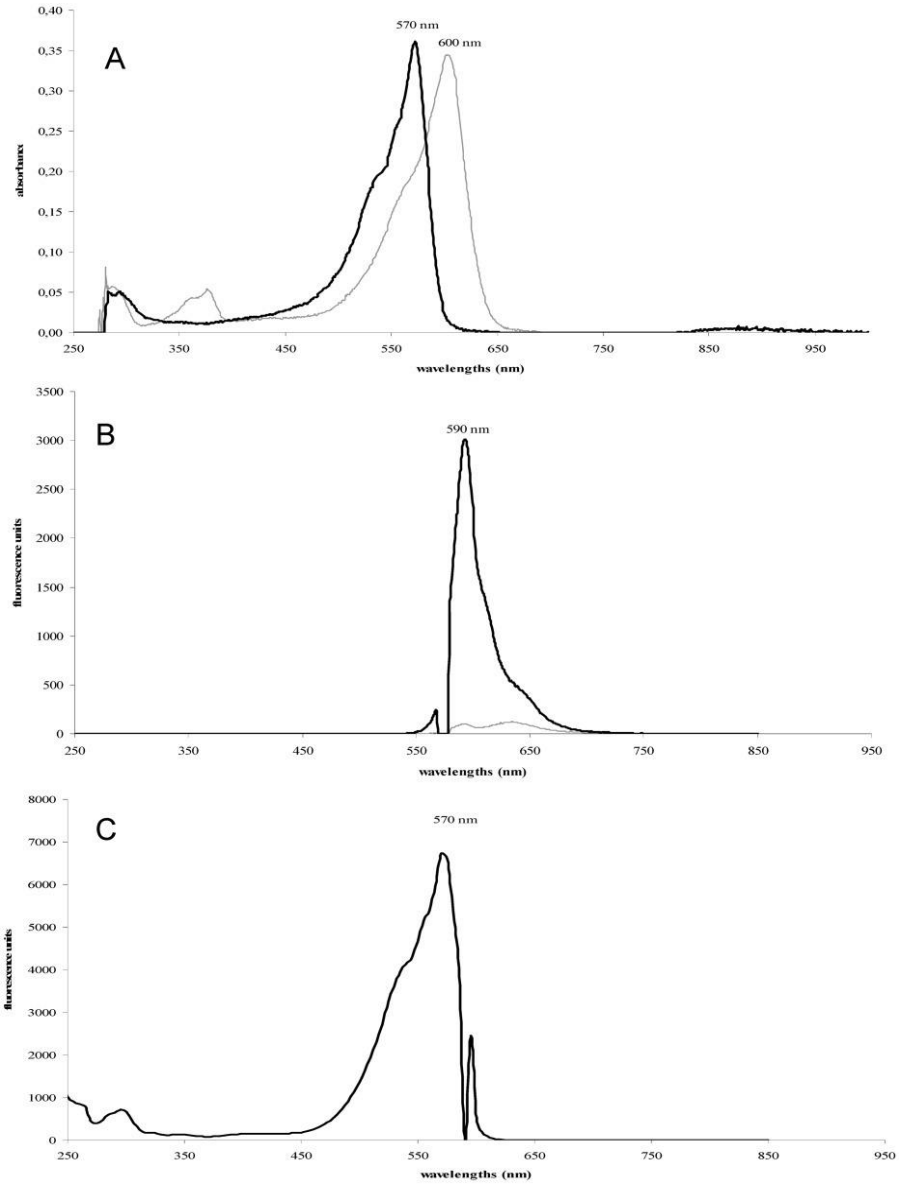



*Figure 3.2. Resazurin and resorufin absorbance and fluorescence spectra.*

*Absorbance spectra of resazurin (600 nm) and resorufin (570 nm). (b) Emission fluorescence spectra with 570 nm excitation wavelength (c) Excitation spectra. Lines show resazurin (grey) and resorufin (black). Permission obtained for use from (Perrot et al., 2003).*

Resazurin may be useful for high throughput screening (HTS) applications and has successfully been used during screening in 384 or 1536 well microtiter plates (Lim *et al.*, 2016, Shum *et al.*, 2008). In 384 well plates, a Z' of 0.79 in black plates and 0.82 in clear plates were achieved (Lim *et al.*, 2016). HTS in 1536-well microtiter plates achieved a Z' of 0.89 (Shum *et al.*, 2008). It should be noted that fluorescent compounds, as well as those with anti-oxidant properties can interfere with the assay and lead to false negative results (Shenoy *et al.*, 2017). Resazurin assays have improved performance when the experimental parameters are optimised for each cell line of interest (Larsson *et al.*, 2020).

Resazurin has been used for several venom and poison studies previously. This includes investigation of venom induced cell death pathways from *Chironex fleckeri* (box jellyfish) venom (Lau *et al.*, 2019), cytotoxicity investigation of toad poison on leukaemia cells (Abdelfatah *et al.*, 2019, Ferreira *et al.*, 2013), investigation of ion channel toxins in *Centruroides limpidus* scorpion venom (Cota-Arce *et al.*, 2020) and tetracycline as a prevention of dermonecrosis caused by *Loxoceles* spider venom (Paixao-Cavalcante *et al.*, 2007).

Overall, resazurin provides a cost effective and robust method of measuring cell viability which has already been used to study the effects of venom.

### 3.1.3 Chapter 3 Aims

The aims of this chapter were to:

- optimise a resazurin-based cytotoxicity assay to effectively screen a panel of cobra venoms against BxPC-3 pancreatic cancer and SW620 colorectal cancer cells
- screen a panel of 19 cobra venoms using the optimised assay
- select the most active cobra venoms to take forwards to HPLC separation

## 3.2 Results

### 3.2.1 Optimising the resazurin assay to allow screening of cobra venoms against cancer cells

The following sections detail the optimisation of the resazurin assay to allow screening of venoms against cancer cell lines. The sections are broken down as follows:

3.2.1.1 Optimising the concentration of resazurin for use in the venom screens

3.2.1.2 Resazurin deposits and potential toxicity

3.2.1.3 Choosing an appropriate positive control

3.2.1.4 Optimisation of cell seeding densities and fluorescence excitation wavelengths for cell assays

3.2.1.5 Limitations of the resazurin assay over prolonged periods of time

3.2.1.6 Investigation of assay robustness by Z'

#### *3.2.1.1 Optimising the concentration of resazurin for use in the venom screens*

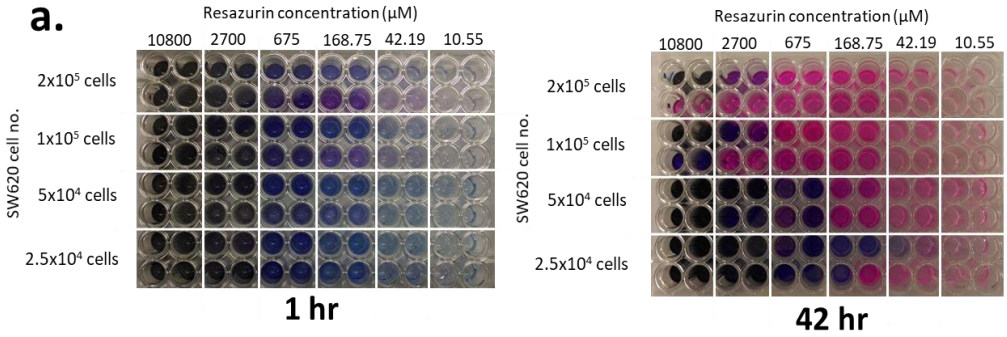
Before screening could be undertaken, the optimal concentration of resazurin to be used needed to be determined. Concentrations of resazurin reported in the literature range from 2.7 mM used for bacterial assays (Sarker *et al.*, 2007) down to 44 µM for both adherent and non-adherent mammalian cells (Anoopkumar-Dukie *et al.*, 2005). Commercially available, branded resazurin solutions range between 440-560 µM (Borra *et al.*, 2009, O'brien *et al.*, 2000). These are usually used as a 10% solution, giving a standard final resazurin assay concentration of 44-56 µM.

Resazurin concentrations of 10.8 mM, 2.7 mM, 675 µM, 168.75 µM, 42.19 µM and 10.55 µM were tested to determine the appropriate resazurin dose for screening assays. SW620 cell seeding densities of $2x10^5$, $1x10^5$, $5x10^4$ and $2.5x10^4$ cells/well were plated to give a matrix of different cell numbers and resazurin concentrations to give a more detailed picture of the interaction between these two factors. 12.5% DMSO was used as a positive control with normal cell media as a negative control. The plates were measured at various time points from 1 hour up to 48 hours following resazurin addition. The aim of the optimisation was to see which combination of cell number and resazurin concentration lead to the largest window between positive and negative control values.

The heatmap (Figure 3.3) shows the signal to background (S/B) ratio for each of the combinations of resazurin concentration and SW620 cell number. This was calculated by dividing the negative control (untreated cells) by the positive control (cells treated with 12.5% DMSO) to show how many times larger the signal is compared to the background of the assay. A value of

1 indicates that the signal is equal to the background. The largest S/B ratio was 5.09, achieved by the highest cell number (2x10$^5$ cells/well), with the highest resazurin concentration (10.8 mM), at the 42-hour post-resazurin dose time point. The signals in this section were measured for absorbance at 595 nm and fluorescence values were investigated, increasing the S/B ratio to around 12.

**a.**




**b.**

2x10$^5$ cells/well

| | Resazurin concentration (μM) | | | | | |
|---|---|---|---|---|---|---|
| | 10800 | 2700 | 675 | 168 | 42.19 | 10.55 |
| **1h** | 1.08 | 1.08 | 1.44 | 1.33 | 1.30 | 1.16 |
| **18h** | 2.48 | 2.44 | 1.77 | 1.10 | 1.21 | 1.35 |
| **24h** | 3.62 | 1.76 | 1.25 | 1.03 | 1.07 | 1.11 |
| **42h** | 5.09 | 1.88 | 1.10 | 1.14 | 1.00 | 0.89 |
| **48h** | 4.07 | 2.03 | 1.27 | 1.28 | 1.40 | 1.43 |

1x10$^5$ cells/well

| | Resazurin concentration (μM) | | | | | |
|---|---|---|---|---|---|---|
| | 10800 | 2700 | 675 | 168 | 42.19 | 10.55 |
| **1h** | 0.98 | 0.99 | 1.27 | 1.39 | 1.17 | 1.33 |
| **18h** | 1.07 | 4.20 | 2.16 | 1.09 | 0.84 | 1.12 |
| **24h** | 1.44 | 4.27 | 2.14 | 1.23 | 0.98 | 1.10 |
| **42h** | 2.10 | 3.36 | 1.91 | 1.33 | 1.06 | 0.94 |
| **48h** | 2.13 | 3.41 | 1.98 | 1.36 | 1.14 | 1.18 |

5x10$^4$ cells/well

| | Resazurin concentration (μM) | | | | | |
|---|---|---|---|---|---|---|
| | 10800 | 2700 | 675 | 168 | 42.19 | 10.55 |
| **1h** | 1.00 | 0.98 | 0.95 | 1.06 | 1.04 | 1.26 |
| **18h** | 1.00 | 1.09 | 1.28 | 1.36 | 1.04 | 1.25 |
| **24h** | 1.01 | 1.24 | 1.40 | 1.40 | 1.16 | 1.18 |
| **42h** | 1.02 | 1.50 | 1.64 | 1.44 | 1.25 | 1.16 |
| **48h** | 1.01 | 1.53 | 1.77 | 1.46 | 1.22 | 1.24 |

2.5x10$^4$ cells/well

| | Resazurin concentration (μM) | | | | | |
|---|---|---|---|---|---|---|
| | 10800 | 2700 | 675 | 168 | 42.19 | 10.55 |
| **1h** | 0.92 | 0.93 | 1.01 | 1.06 | 0.96 | 1.08 |
| **18h** | 0.91 | 0.94 | 1.08 | 1.59 | 1.46 | 1.27 |
| **24h** | 0.92 | 0.95 | 1.10 | 1.59 | 1.43 | 1.26 |
| **42h** | 0.91 | 0.96 | 1.19 | 1.77 | 1.40 | 1.26 |
| **48h** | 0.92 | 0.97 | 1.26 | 1.67 | 1.32 | 1.09 |

*Figure 3.3. Optimisation of resazurin concentration.*

*The optimisation was performed using SW620 cells taking absorbance readings at 595 nm. (a) Representative plate photos at 1 hr and 42 hours post resazurin dosing. Cell seeding density is shown down the side of the plates and resazurin concentration (μM) is shown along the top. (b) Heatmaps showing the S/B ratio between positive and negative control wells at different cell seeding densities. Hours of resazurin exposure are shown down the side and resazurin concentration (μM) is shown along the top.*

### 3.2.1.2 *Resazurin deposits and potential toxicity*

The cells exposed to the various concentrations of resazurin and observed using an inverted microscope. Cells exposed to a resazurin concentration of 675 µM or greater began to have black deposits appear in the wells as shown in Figure 3.4. In addition, the cells exposed to the higher resazurin concentrations (≥675 µM) appeared smaller and more rounded than usual. The 'normal' morphology can be observed in the 168 µM sample shown in Figure 3.4 with the cells becoming smaller and more rounded as they are exposed to higher doses of resazurin.

It was decided that this effect would be examined further and that doses between 160 and 660 µM would be investigated to decide upon a suitable dose (Figure 3.5). The appearance of the black dots was proportional to the number of cells present in the well and the highest cell number (1- and $2x10^5$ cells/well) had toxicity up until a dose of 360 µM and 260 µM respectively. This contrasts with the lowest cell numbers (2.5- and $5x10^4$ cells/well) which only had black dots visible from a higher dose of 460 µM resazurin exposure. The only dose of resazurin with no visible black dots, cell rounding or size reduction for all cell numbers was 160 µM and so this dose was selected to take forwards.

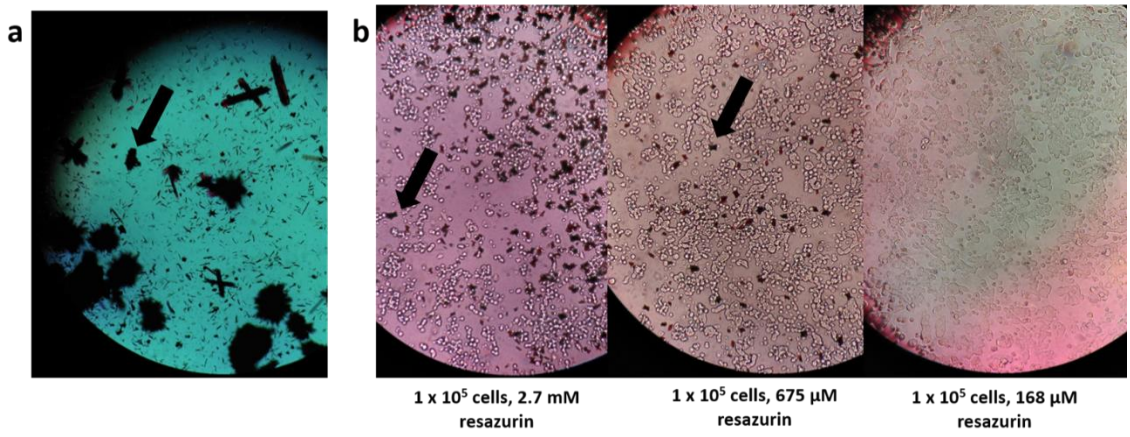



*Figure 3.4. Resazurin deposits observed in samples.*

*Observed in (a) column 1 of the compound plate (no cells present, 21.6 mM resazurin concentration) and (b) $1x10^5$ SW620 cells exposed to various doses of resazurin. Black deposits (indicated by a black arrow) can clearly be seen in the compound plate and in the cell containing wells at 2.7 mM and 675 µM concentrations of resazurin but appear to be absent (or unobservably low) in the 168 µM sample. Cells in the wells exposed to 2.7 mM and 675 µM resazurin concentrations also appear smaller, rounder and sparser than the cells exposed to 168 µM resazurin, despite being plated at the same time.*

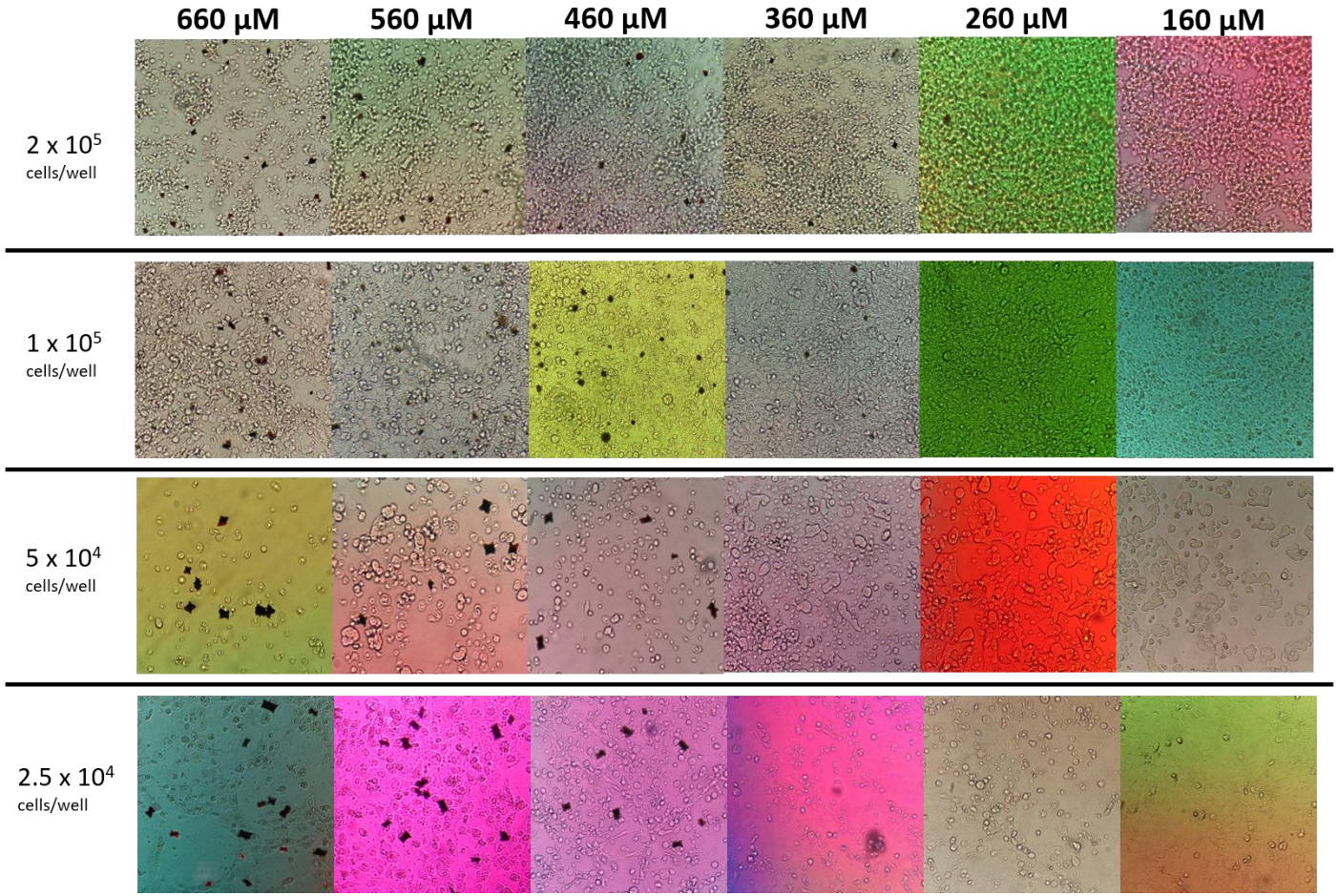



*Figure 3.5. Resazurin deposits observed even at lower concentrations.*

*Microscopy photos of different numbers of SW620 cells exposed to descending concentrations of resazurin between 160-660 µM. Black deposits can still be seen between 460-660 µM in the 2.5 x$10^4$ and 5x$10^4$ cell concentrations, 360-660µM in the 1x$10^5$ cell concentrations and 260-660 µM in the 2x$10^5$ cell concentration. Cells in wells with black deposits also appear smaller, rounder and sparser than cells in wells without the presence of black deposits.*

### *3.2.1.3 Choosing an appropriate positive control*

During the resazurin concentration optimisation in section 3.2.1.1, 12.5% DMSO was used as a positive control to inhibit cell growth. It was found, however, that even the high 12.5% DMSO still allowed cells at high concentration to continue growing and to metabolise the resazurin dye to resorufin.

*Naja nigricollis* venom was tested to see whether it could act as a better positive control for the screening assays. This venom was investigated as this had been previously shown to be cytotoxic to many cell types (Conlon *et al.*, 2020). A dose response assay was performed according to section 2.2.2 to decide a suitable concentration of *Naja* nigricollis venom to use as a positive control. Calculating from the wider venom concentration range, *Naja nigricollis* venom in water and PBS gave $IC_{50}$ values of approximately 15 µg/ml and 8 µg/ml respectively. A narrower concentration range was also tested, although this resulted in concentration response curves which gave extremely steep curves (nHill > 5 in all cases) where it was not possible to accurately calculate an $IC_{50}$ value. A concentration of 100 µg/ml was chosen as it showed a full inhibitory affect against the tested SW620 cell line and was approximately 10X the calculated $IC_{50}$ value. It was decided that the venom would be suspended in PBS as this is isotonic for the cells and did not significantly affect the $IC_{50}$ of the venom.

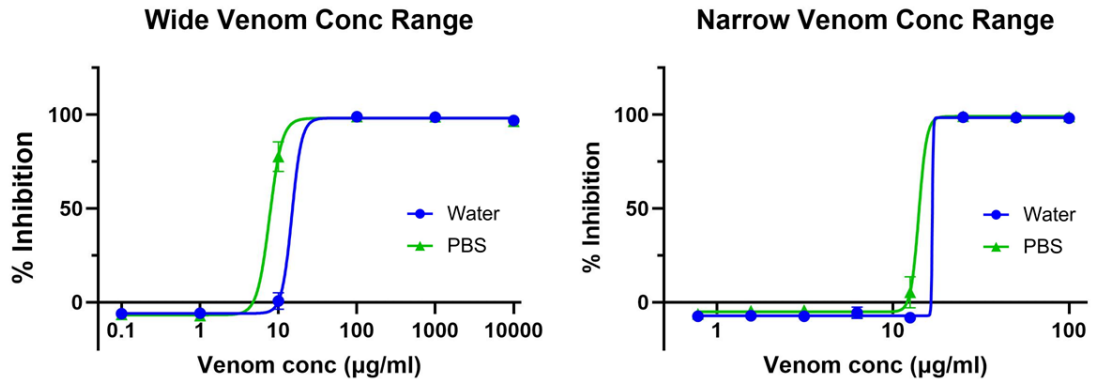



*Figure 3.6. Naja nigricollis dose response curve.*

*'Wide venom conc range' shows a dose response curve from 0.1-1000 µg/ml and 'narrow venom conc range' shows a dose response curve from 0.78-100 µg/ml on the SW620 cell line. Error bars show the SEM based on n=3.*

To determine the minimum required incubation time of the positive control, 100 µg/ml of *Naja nigricollis* venom was added to wells of SW620 and BxPC-3 cells and incubated for up to 6 hours according to section 2.2.2. Figure 3.7 shows the S/B ratio between the venom treated and the untreated wells. After 60-120 minutes of venom exposure the graphs begin to plateau, and no further significant improvements to S/B ratio are observed. Therefore, a venom incubation time of 120 minutes was used for future experiments.

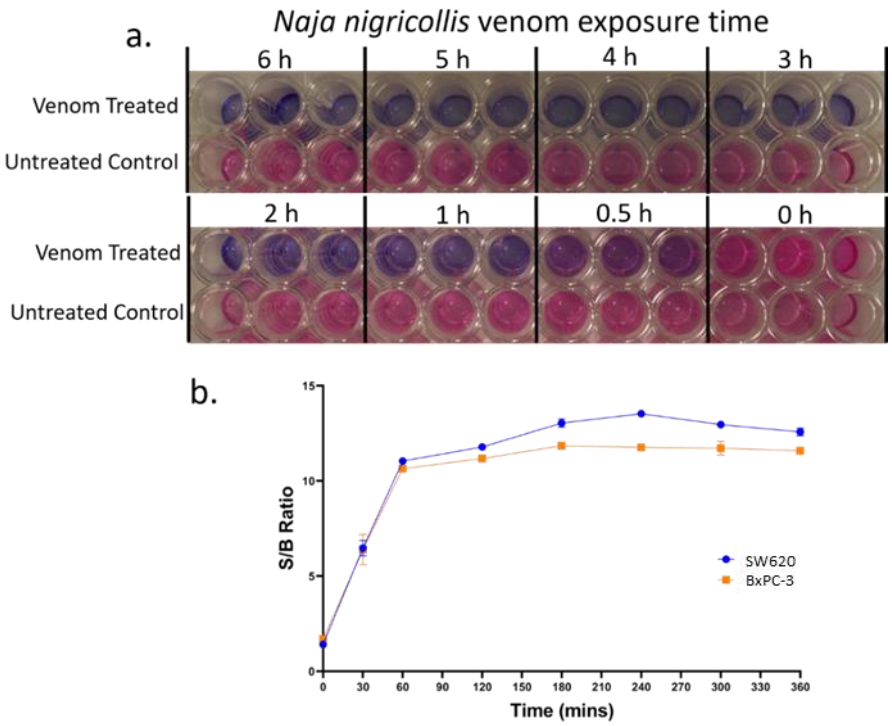



*Figure 3.7. The effect of cell exposure to Naja nigricollis venom over time.*

*(a) Plate photos showing the colour differences of the venom treated and untreated SW620 cells exposed to resazurin. Blue = full growth inhibition, purple = partial growth inhibition, pink = no growth inhibition (b) Graph to show a linear increase in S/B ratio up to between 60-120 minutes of venom exposure in SW620 and BxPC-3 cell lines followed by a plateau in S/B ratio. Error bars show SEM based on n=3.*

### *3.2.1.4 Optimisation of cell seeding densities and fluorescence excitation wavelengths for cell assays*

#### 3.2.1.4.1 Optimisation for 96-well format

Once the resazurin concentration had been optimised, the number of cells to be plated per 96-well plate well was optimised for BxPC-3 and re-optimised for the previously tested SW620 cells. Fluorescence was also investigated as a potentially better means of measuring resazurin changes in response to cell metabolism. To establish the optimum number of cells to be plated per well of each 96-well plates for use during the resazurin assay, a variety of cell numbers were tested to see which provided the biggest window between cells treated with the positive and negative control. *Naja nigricollis* venom at a concentration of 100 µg/ml was used as a positive control and cell media appropriate for each cell line was used as a negative control to maintain normal cell growth and metabolism for the cells.

Cells were treated either with or without venom and the S/B ratio was calculated for different time points. For fluorescence readings, the S/B ratio was calculated using fluorescence values from different excitation wavelengths (355, 485 and 544 nm) with an emission wavelength of 590 nm. For absorbance readings, an absorbance wavelength of 595 nm was used.

The S/B ratio for BxPC-3 cells at different excitation wavelengths and at 595 nm absorbance are shown in Figure 3.8. This assay revealed that the optimum number of cells to seed for the BxPC-3 cell line until 120 minutes is $1x10^5$ cells per well. For the SW620 cell line, the optimum number of cells to seed for the assay until 150 minutes is $2x10^5$ cells/well. The results are displayed in Figure 3.9. There did not appear to be much of a difference between the fold-change calculated from an excitation wavelength of 544 nm compared to 485 nm for either cell line. All wavelengths were therefore taken forward for Z' assay robustness testing.

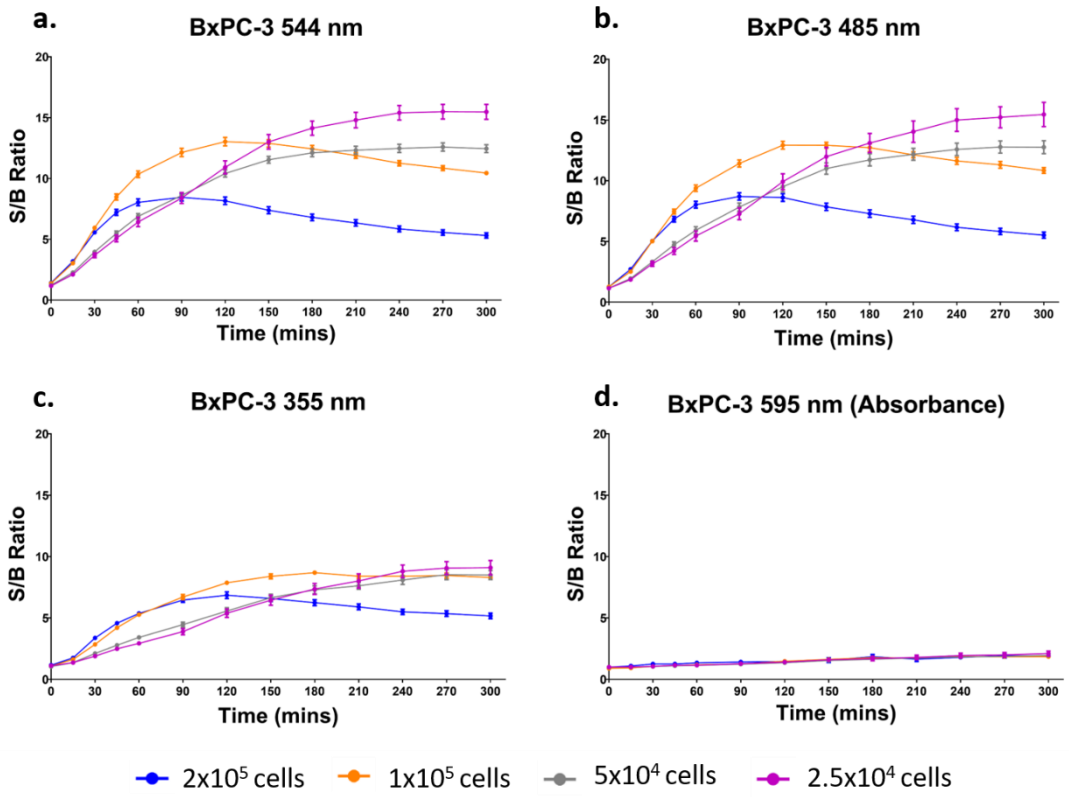



*Figure 3.8. BxPC-3 96-well cell number optimisation.*

*Graphs showing the S/B ratio between values of BxPC-3 cells exposed to the positive control (Naja nigricollis venom) and the negative control (appropriate cell media). Fluorescence values measured with an emission wavelength of 590 nm and excitation wavelength of (a) 355 nm (b) 485 nm (c) 544 nm. (d) Absorbance values measured at 595 nm.*

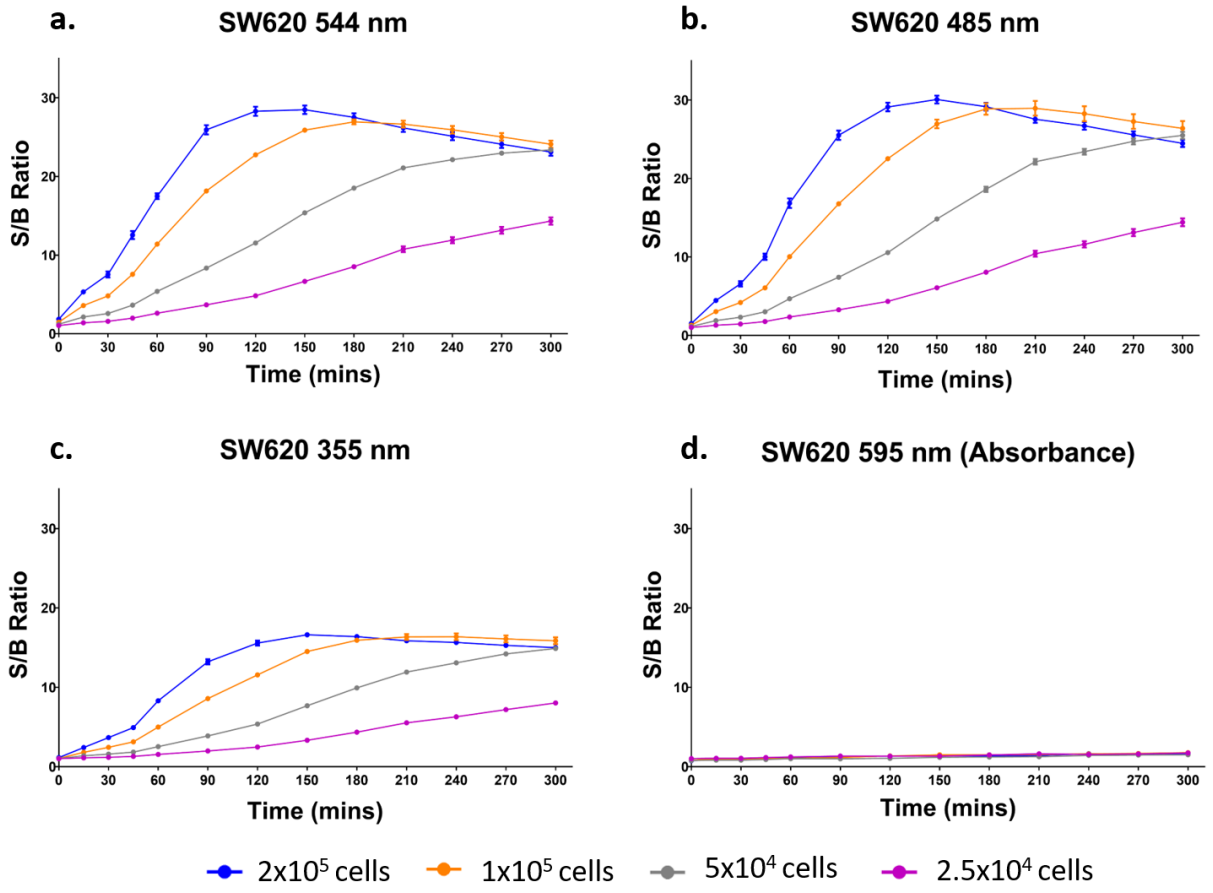



*Figure 3.9. SW620 96-well cell number optimisation.*

*Graphs showing the S/B ratio between values of SW620 cells exposed to the positive control (Naja nigricollis venom) and the negative control (appropriate cell media). Fluorescence values measured with an emission wavelength of 590 nm and excitation wavelength of (a) 355 nm (b) 485 nm (c) 544 nm. (d) Absorbance values measured at 595 nm.*

#### 3.2.1.4.2 Optimisation for 384-well format

The assay was also optimised for 384 well format with the results presented in Figure 3.10. Different cell seeding densities were selected for each of the cell lines due to the difference in optimum cell number determined in the 96-well version of the assay.

By the standard 120 minute assay time point, the highest two cell numbers of BxPC-3 (2.5 x $10^5$ and 3.3 x $10^5$) had almost the same S/B ratio and thus the lower cell number of the two (2.5 x $10^5$ cells/well) was chosen to make the assay more economical. For the SW620 cells, the highest cell number tested (5 x $10^5$ cells) gave a much larger S/B ratio than the next highest number (3.3 x $10^5$) at the standard 120 minutes assay time. Therefore, the highest cell number (5 x $10^5$cells/well) was selected for future assays for this cell line.

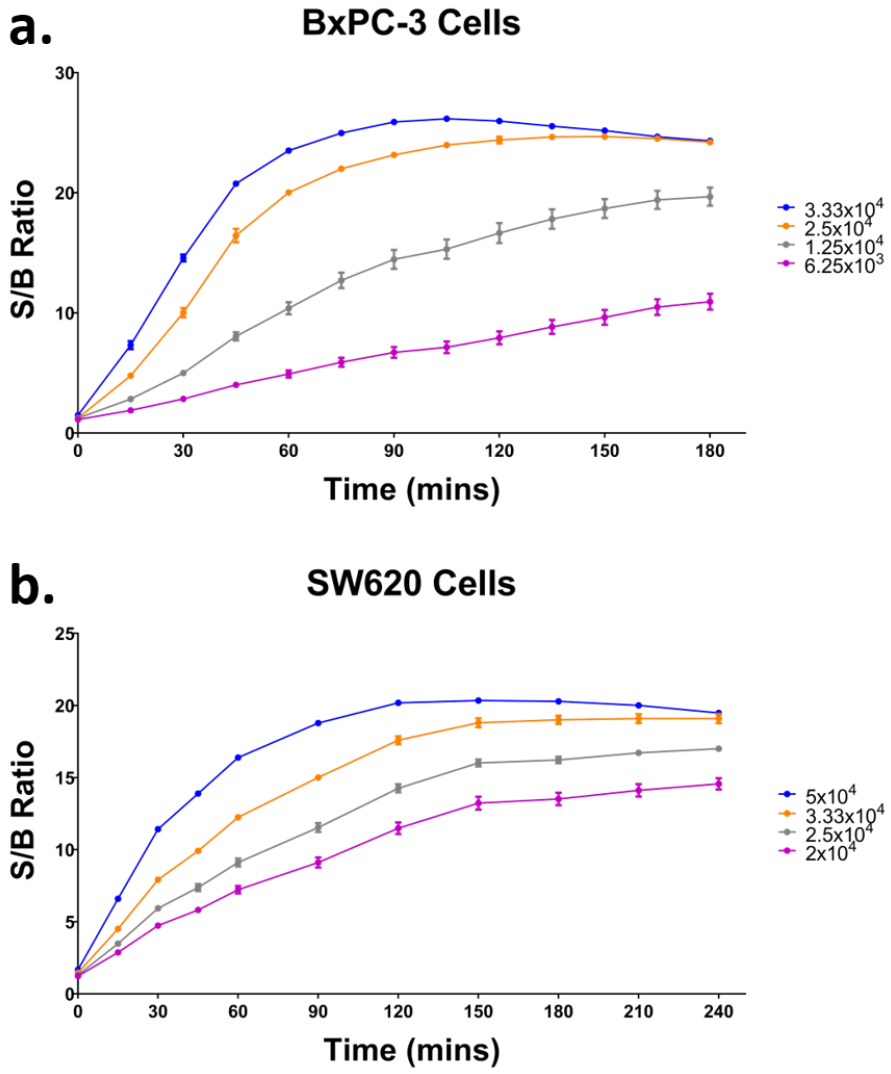



*Figure 3.10. 384-well cell number optimisation.*

*Graphs to show the S/B ratio between the positive control (cells treated with Naja nigricollis venom) compared to the negative control (cells treated with appropriate cell media) over time for (a) BxPC-3 and (b) SW620 cells.*

### *3.2.1.5 Limitations of the resazurin assay over prolonged periods of time*

In the previous sections, the fluorescence values of the negative control are divided by the positive control in order to give the S/B ratio. In the graphs displayed in Figure 3.11, the raw, relative fluorescence units (RFU) are plotted against time to show that after an initial log phase, the cells then appear to undergo a lag phase where the fluorescence values plateau. . This effect is particularly pronounced in the highest cell number wells as their combined metabolic rate is higher than that of the lower cell number wells. At this resazurin concentration, the graphs begin to plateau after around 120 minutes. Therefore, when using this concentration of resazurin, the assay should not be run for much beyond 120 minutes.

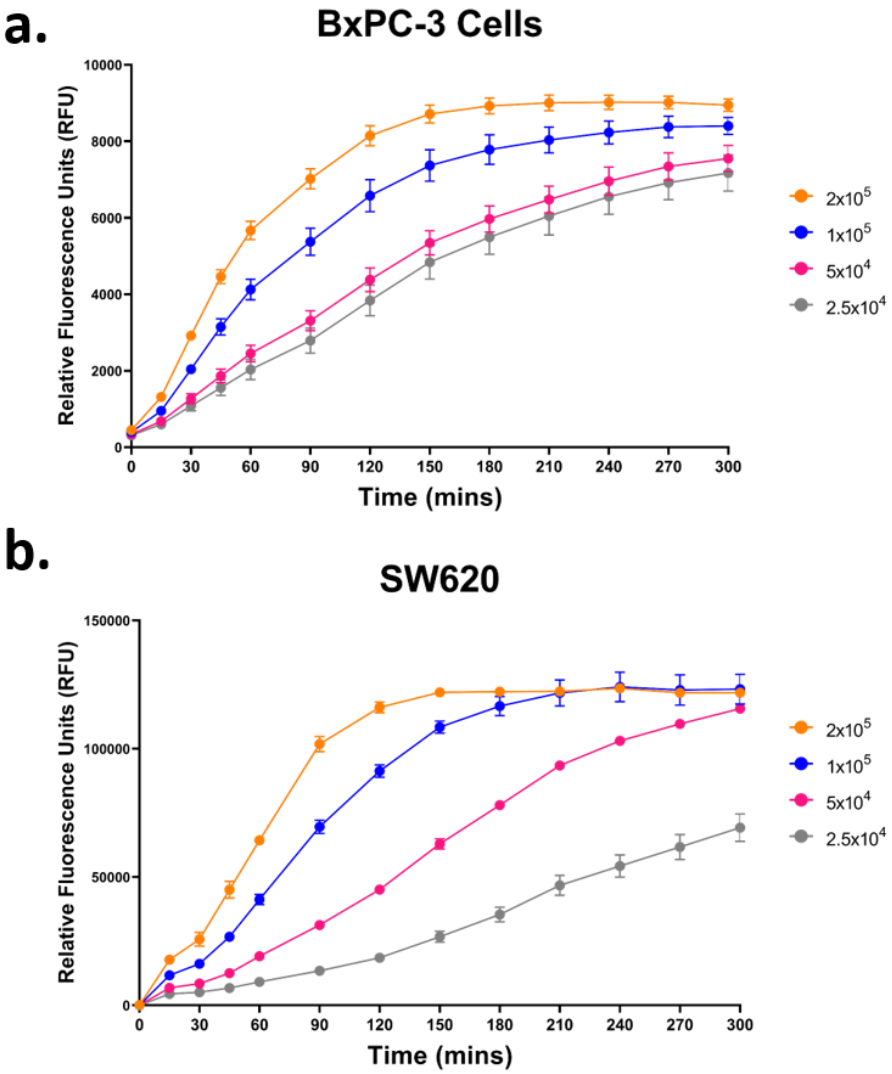



*Figure 3.11. The effect of resazurin fluorescence readings over time.*

*Graphs illustrating the raw, relative fluorescence values (RFU) over time for (a) BxPC-3 cells and (b) SW620 cells. These graphs illustrate that at higher cell numbers, the fluorescence values eventually plateau as the free resazurin is depleted.*

### *3.2.1.6 Investigation of assay robustness by Z'*

#### 96-well Format

The robustness of the optimised assays was tested to determine the Z'. Z' is a statistical test which investigates the robustness of an assay by taking into consideration the sample means and standard deviations of each sample. The Z' values were calculated for each different time point at 595 nm absorbance and at the three different fluorescence excitation wavelengths and are compared in Figure 3.12.

The absorbance readings failed to provide a satisfactory value for Z' (<0.5 = poor HTS assay) and therefore absorbance was not used for this assay. It was unclear from the initial data in section 3.2.1.4 whether an excitation wavelength of 485 or 544 nm was preferable for either cell line, but upon performing the Z' assay, the 544 nm excitation wavelength outperformed the 485 nm wavelength. It should be noted, however, that all three wavelengths produced satisfactory values (>0.5) for Z' at all times tested and so any of the three wavelengths would have been suitable for the assay.

In both cell lines, the 544 nm excitation wavelength produced the highest Z' compared to the other two wavelengths and the optimal time point was after 2 hours of resazurin exposure in both cases. These were therefore the parameters chosen to take forward for the cobra venom screen.

#### 384-well format

After optimising cell plating densities for 384-well format, the robustness of assay was investigated by repeating the Z' assay as described above. Results can be found in Figure 3.12. Despite slight increases in Z' for both cell lines after 120 minutes, a compromise was reached on Z' versus time taken for assay. Since both cell lines should have the same length of exposure to venom fractions during assays, it was decided that the 120-minute time point (a Z' pass for both cell lines) would be used for future assays.

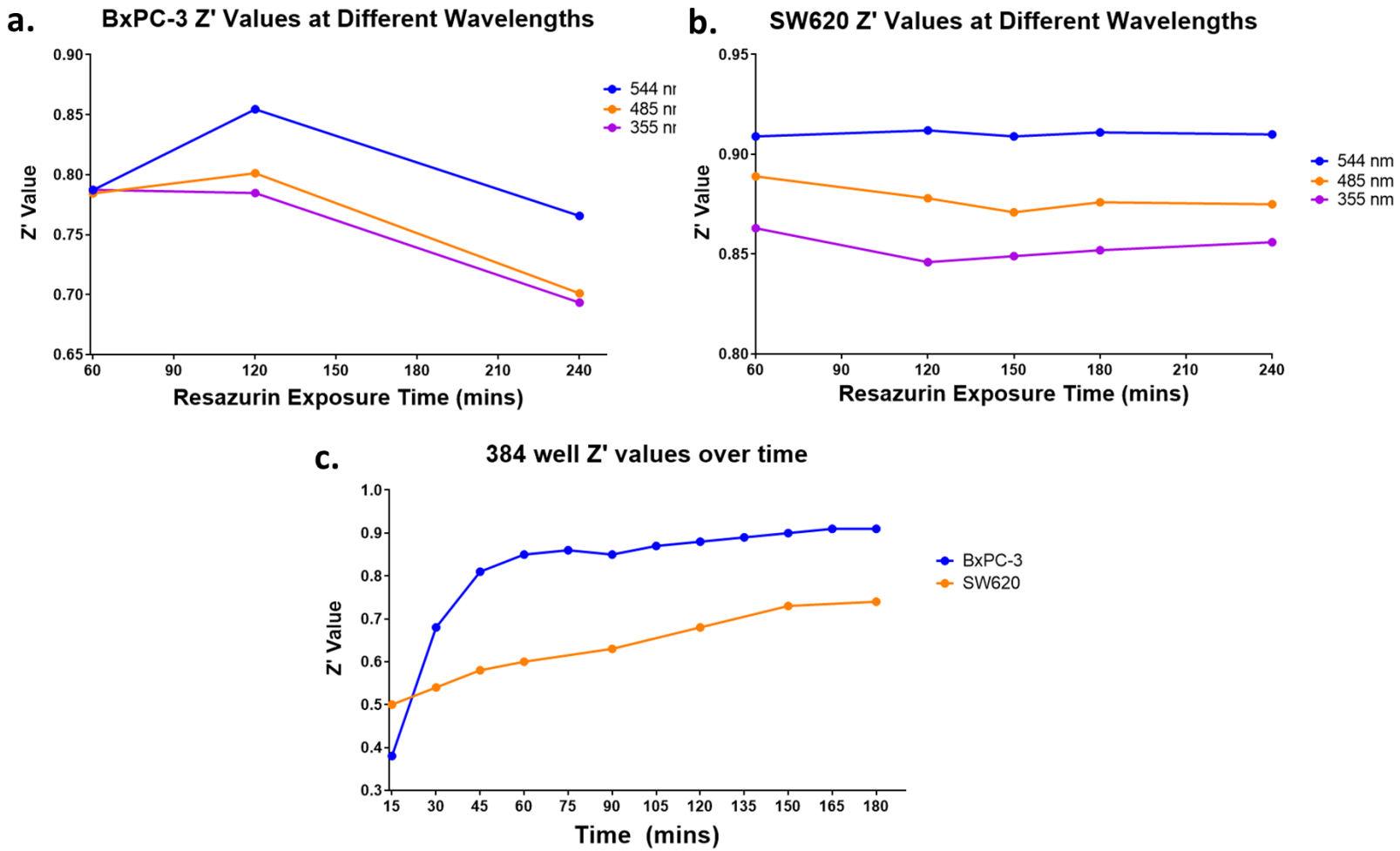



*Figure 3.12. Z' values at different wavelengths and over time.*

*Graphs to show the changes in Z' over time at different wavelengths for the (a) BxPC-3 and (b) SW620 cell lines. (c) Graph to show Z' over time for BxPC-3 cells (2.5x10$^4$ cells/well) and SW620 cells (5x10$^4$ cells/well) (at 544 nm)*

### 3.2.2 Investigating Inhibition of SW620 and BxPC-3 Cell Viability by Whole Cobra Venom Using Optimised Resazurin Screening Assay

Following resazurin assay optimisation, the cobra venom screening assay was used to screen a range of 19 different whole cobra venoms. The details of the screened venoms can be found in Table 3.1.

*Table 3.1. Details of the cobras used in the cobra venom screen.*

*Details for each of the cobras screened including the species, common name, the 'type' of cobra (spitting or non-spitting) and the location each snake may be found (Africa or Asia). Positive control (+ con) was a 100 µg/ml concentration of Naja nigricollis venom. Negative (untreated) control (- con) was cell media appropriate for the cell line being grown.*

| CODE | Species | Common name | Type | Region |
|---|---|---|---|---|
| **N.mos** | *Naja mossambica* | Mozambique spitting cobra | Spitting | Africa |
| **N.nct** | *Naja nigricincta nigricincta* | Zebra Spitting Cobra | Spitting | Africa |
| **N.nig** | *Naja nigricollis* | Black Necked Spitting Cobra | Spitting | Africa |
| **N.nub** | *Naja nubiae* | Nubian Spitting Cobra | Spitting | Africa |
| **N.pal** | *Naja pallida* | Red Spitting Cobra | Spitting | Africa |
| **N.sam** | *Naja samarensis* | Siamese Samar Cobra | Spitting | Asia |
| **N.sia** | *Naja siamensis* | Indo Chinese Spitting Cobra | Spitting | Asia |
| **N.spu** | *Naja sputatrix* | Javan Spitting Cobra | Spitting | Asia |
| **N.aan** | *Naja annulifera* | Snouted Cobra | Non-spitting | Africa |
| **N.anc** | *Naja anchietae* | Anchieta's Cobra | Non-spitting | Africa |
| **N.haj** | *Naja haje* | Egyptian Cobra | Non-spitting | Africa |
| **N.hle** | *Naja haje legionis* | Moroccan Cobra | Non-spitting | Africa |
| **N.mel** | *Naja melanoleuca* | Forest Cobra | Non-spitting | Africa |
| **N.niv** | *Naja nivea* | Cape Cobra | Non-spitting | Africa |
| **N.atr** | *Naja atra* | Chinese Cobra | Non-spitting | Asia |
| **N.kao** | *Naja kaouthia* | Monocled Cobra | Non-spitting | Asia |
| **N.naj** | *Naja naja* | Spectacled Cobra | Non-spitting | Asia |
| **N.nka** | *Naja naja karchachensis* | Pakistan Black Cobra | Non-spitting | Asia |
| **O.han** | *Ophiophagus hannah* | King Cobra | Non-spitting | Asia |
| **+ Con** | *Naja nigricollis* | Black Necked Spitting Cobra | Spitting | Africa |
| **- Con** | N/a | N/a | N/a | N/a |

Venoms were screened at 100 µg/ml and 10 µg/ml concentrations in quadruplicate across two 96-well plates (duplicate on each plate). Plate images are shown in Figure 3.13. Using the fluorescence plate readings, the percentage inhibition of venom treated cells was calculated for each cell line relative to the negative (untreated) control, as shown in other venom HTS screen papers (Teixeira *et al.*, 2016).

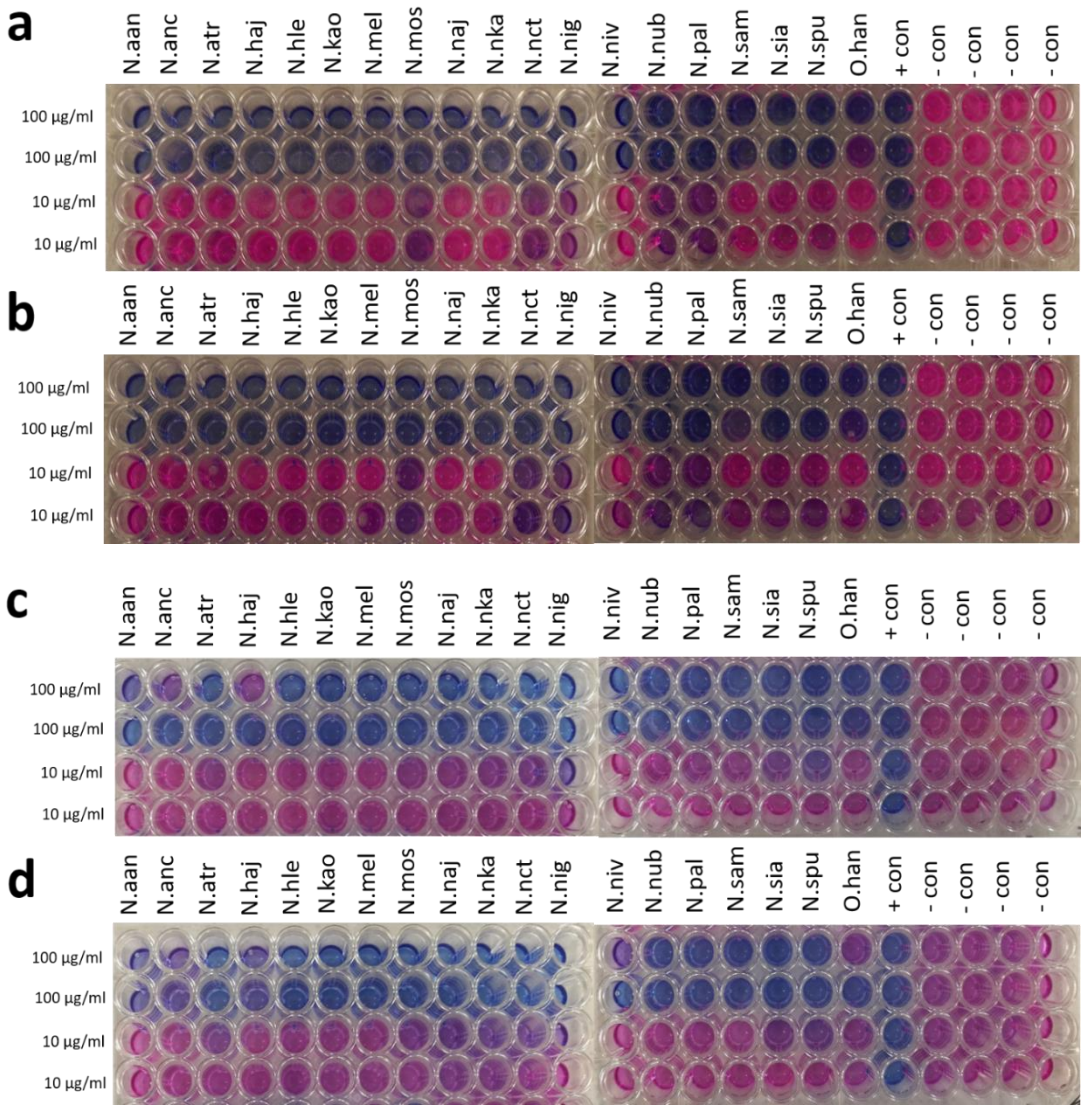



*Figure 3.13. Plate images taken after 2 hours of resazurin exposure.*

*(a) SW620 plate 1 (b) SW620 plate 2 (c) BxPC-3 plate 1 (d) BxPC-3 plate 2. Each plate contained the optimised number of cells and was treated with either 100 µg/ml or 10 µg/ml of each cobra venom for 2h, then treated with resazurin dye. Pink wells = healthy, metabolising cells, purple/blue wells = inhibited cells. Positive control (+ con) = cells treated with 100 µg/ml Naja nigricollis whole venom. Negative control (- con) = cells treated with appropriate cell media.*

Histograms showing the distribution of cell inhibition when each cell line was exposed to the different concentration of each of the cobra venoms are shown in Figure 3.14. All the distributions were skewed to varying degrees and none of the data sets showed a Gaussian distribution. The quantile-quantile plot (QQ-plot, Appendix II) illustrates that the data are not normally distributed. Anderson-Darling, D'Agostino & Pearson, Shapiro-Wilk and Kolmogorov-Smirnov normality tests were run on the screen data, and all showed the data was not normally

distributed in either cell line, at any concentration (P<0.0001 in all cases). Summary data can be found in Appendix II.

Graphical representations of the results of the screens can be found in Figure 3.15 (100 µg/ml) and Figure 3.16 (10 µg/ml). All cobra venoms tested at 100 µg/ml concentration caused at least 40% inhibition in both cell lines. To determine which venom samples caused the most significant inhibition, a Kruskal-Wallis statistical analysis with Dunn's multiple comparisons post-hoc test was run for each cell line. A summary of the statistical data is shown in Table 3.2.

When tested at 10 µg/ml concentration, some selectivity was observed with African spitting cobras causing at least 40% inhibition in the SW620 cell line only. No other cobra type caused more than 40% inhibition and no cobras caused more than 40% inhibition in the BxPC-3 cell line. Again, Kruskal-Wallis statistical analysis with Dunn's multiple comparisons post-hoc test was run for each cell line. A summary of the statistical data is shown in Table 3.2.

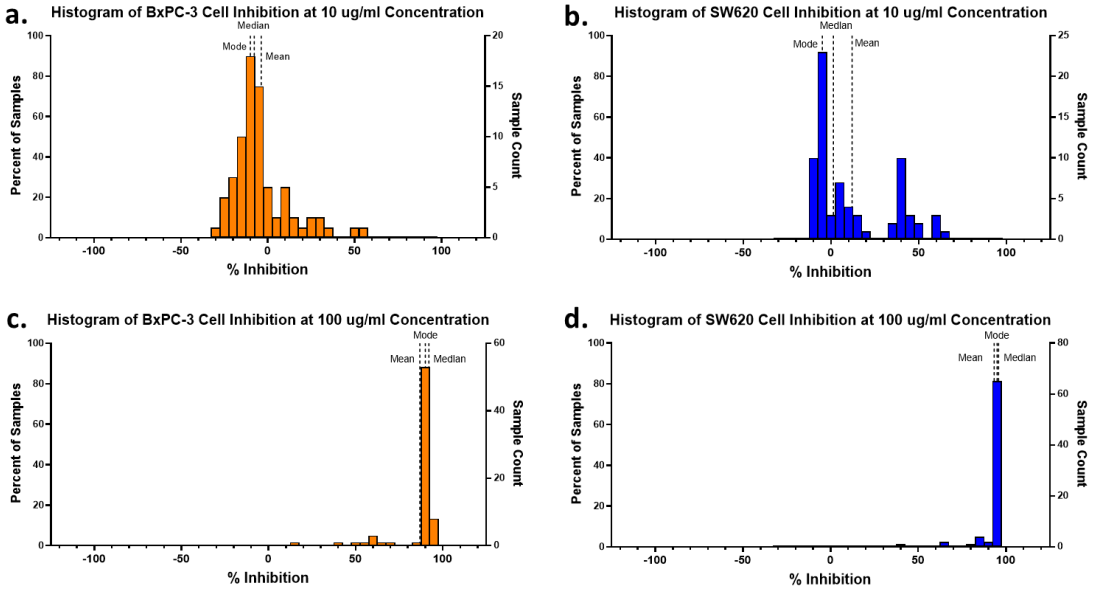



*Figure 3.14. Histograms showing the distribution of cell inhibition following exposure to venoms at different concentrations.*

*(a) BxPX-3 cells at 10 µg/ml (b) SW620 cells at 10 µg/ml (c) BxPC-3 cells at 100 µg/ml (d) SW620 cells at 100 µg/ml concentration. Mean, median and mode are indicated for each data set.*

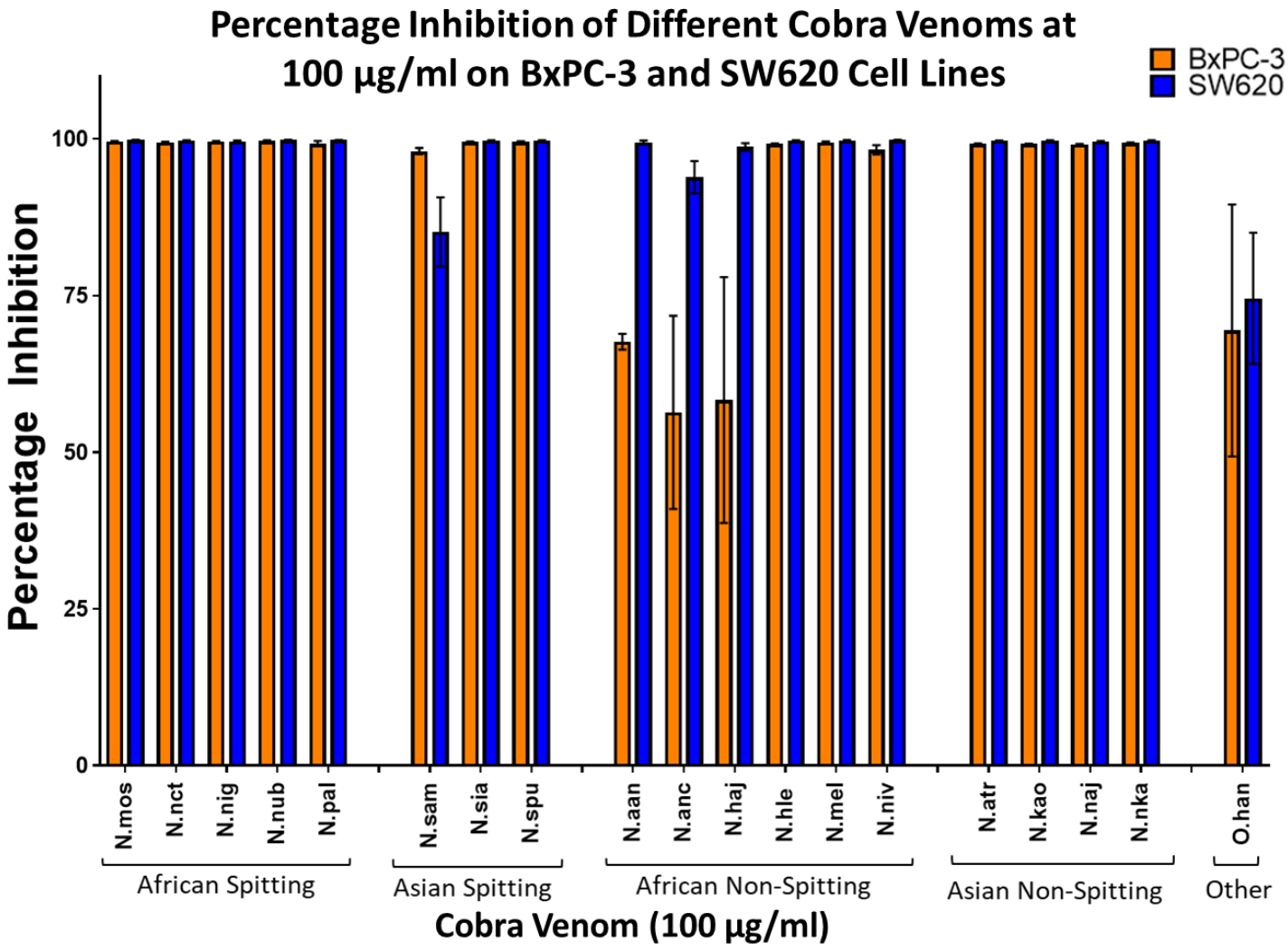



*Figure 3.15. Percentage inhibition of the cobra venoms tested at 100 µg/ml concentration.*

*BxPC-3 cells are shown in blue and SW620 cells are shown in orange. Cobras are grouped according to their type (spitting/non-spitting) and location of origin (Africa/Asia). Three letter species codes indicate which cobra is responsible for each percentage inhibition. Full species codes can be found in Table 3.1. Error bars represent SEM based on n=4.*

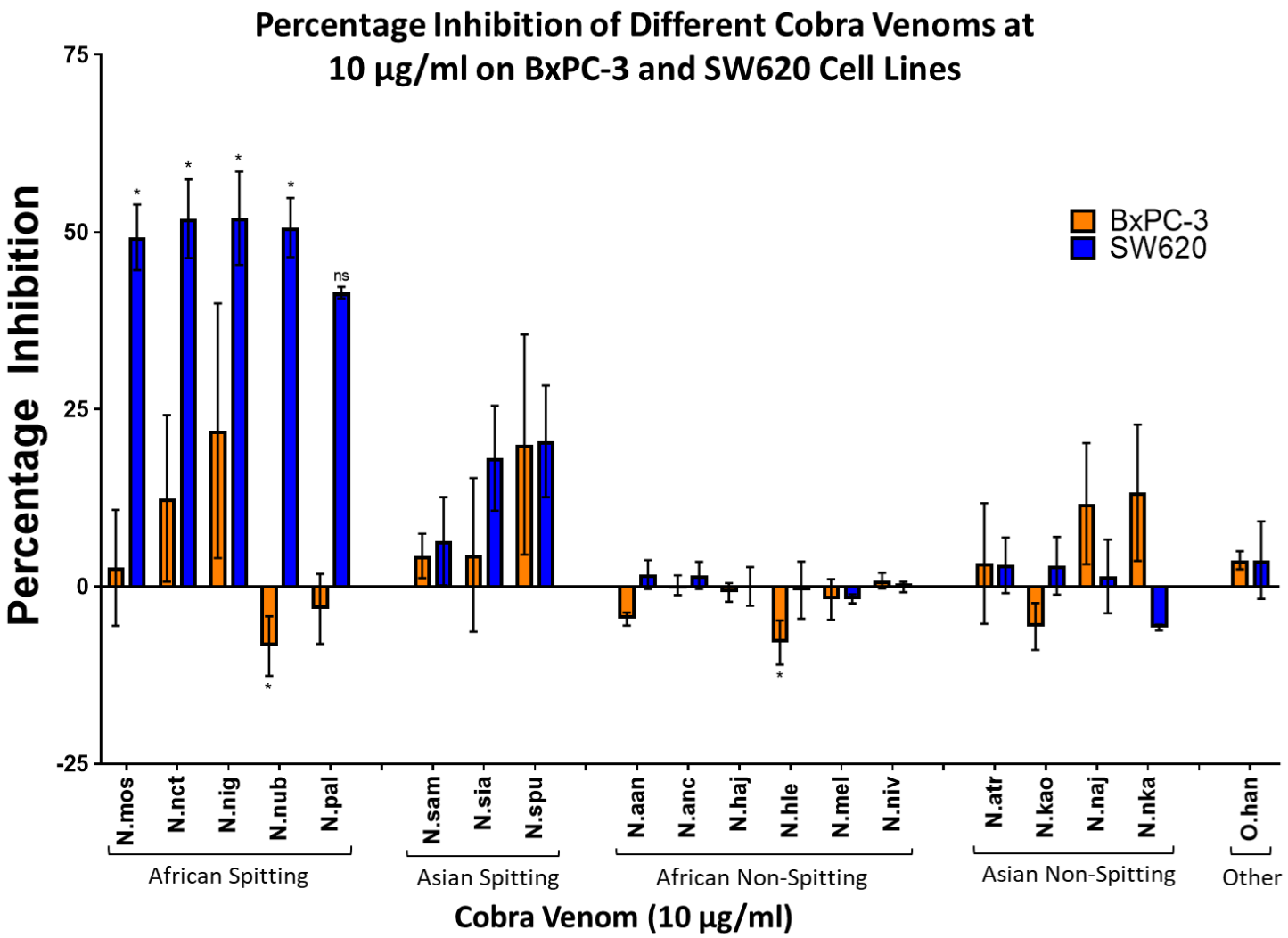



*Figure 3.16. Percentage inhibition of the cobra venoms tested at 10 µg/ml concentration.*

*BxPC-3 cells are shown in blue and SW620 cells are shown in orange. Cobras are grouped according to their type (spitting/non-spitting) and location of origin (Africa/Asia). Three letter species codes indicate which cobra is responsible for each percentage inhibition. Full species codes can be found in Table 3.1. Error bars represent SEM based on n=4. * represents statistical significance at P<0.05 level (ns = not significant at the P<0.05 level). Full statistical analysis is shown in Table 3.2.*

*Table 3.2. Details of the statistical significance levels of the cobra species tested.*

*Table includes species code, Latin name, common name, continent of origin and spitting/non-spitting characteristic. Statistical test was Kruskal-Wallis with Dunn's multiple comparison post-hoc test when compared to the negative (untreated) control wells. * = p<0.05, **= p<0.01, *** = p<0.001, **** = p<0.0001, ns = not significant (p>0.05). Shades of green indicate a significant increase in fluorescence signal suggesting cell growth inhibition whilst shades of orange indicate a significant decrease in fluorescent signal, suggesting cell growth stimulation. Positive (+ve) control= 100 µg/ml N.nig venom, negative (-ve) control = appropriate complete cell media.*

| Code | Species | Region | Type | SW620 100 µg/ml Significance | SW620 10 µg/ml Significance | BxPC-3 100 µg/ml Significance | BxPC-3 10 µg/ml Significance |
|---|---|---|---|---|---|---|---|
| N.mos | *Naja mossambica* | Africa | Spitting | *** | * | *** | ns |
| N.nct | *Naja nigricincta nigricincta* | Africa | Spitting | * | * | ** | ns |
| N.nig | *Naja nigricollis* | Africa | Spitting | *** | * | *** | ns |
| N.nub | *Naja nubiae* | Africa | Spitting | *** | * | *** | * |
| N.pal | *Naja pallida* | Africa | Spitting | *** | ns | ** | ns |
| N.sam | *Naja samarensis* | Asia | Spitting | ns | ns | ns | ns |
| N.sia | *Naja siamensis* | Asia | Spitting | * | ns | *** | ns |
| N.spu | *Naja sputatrix* | Asia | Spitting | * | ns | ns | ns |
| N.aan | *Naja annulifera* | Africa | Non-spitting | * | ns | ns | ns |
| N.anc | *Naja anchietae* | Africa | Non-spitting | ns | ns | ns | ns |
| N.haj | *Naja haje* | Africa | Non-spitting | ns | ns | ns | ns |
| N.hle | *Naja haje legionis* | Africa | Non-spitting | * | ns | ns | * |
| N.mel | *Naja melanoleuca* | Africa | Non-spitting | ** | ns | ** | ns |
| N.niv | *Naja nivea* | Africa | Non-spitting | ** | ns | ns | ns |
| N.atr | *Naja atra* | Asia | Non-spitting | ns | ns | ns | ns |
| N.kao | *Naja kaouthia* | Asia | Non-spitting | * | ns | ns | ns |
| N.naj | *Naja naja* | Asia | Non-spitting | ns | ns | ns | ns |
| N.nka | *Naja naja karchachensis* | Asia | Non-spitting | * | ns | * | ns |
| O.han | *Ophiophagus hannah* | Asia | Non-spitting | ns | ns | ns | ns |
| + con | *Naja nigricollis* | Africa | Spitting | **** | *** | **** | ** |
| - con | N/a | N/a | N/a | N/a | N/a | N/a | N/a |

## 3.3 Discussion

### 3.3.1 Optimisation of Resazurin Assay for Investigation of Cobra Venom Toxicity

A range of resazurin doses were selected from the literature and investigated against the SW620 cell line. The S/B ratio – the difference between the positive control and negative control values – was assessed to ensure the difference between the positive and negative control wells was as large as possible, leading to a more robust assay allowing hits to be confidently identified. HTS assays should have low variability and a high S/B ratio in order to minimize the risk of false positives and false negatives (Macarron and Hertzberg, 2011).

The upper and lower limits of resazurin dose that were tested were based on the concentrations used for a bacterial screening assay by Sarker *et al*. (2007) (2.7 mM concentration used) and a mammalian cell screen by Anoopkumar-Dukie *et al*., (2005) (44 µM concentration used). Based on these two widely spaced reference concentrations, a range was determined, and the most appropriate concentration chosen.

Interestingly, the S/B ratio for this highest resazurin/highest cell number combination decreases at the 48-hour time point. This is likely because prolonged exposure to resazurin could lead to depletion of the remaining available stocks of resazurin, causing an artificial decline in the conversion rate from resazurin to resorufin. Thus the accuracy of the correlation between the cell viability and fluorescence value is lost and a lower S/B ratio is observed than is actually the case (Uzarski *et al.*, 2017). Furthermore, once converted to resorufin, further reduction may occur, converting resorufin to dihydroresorufin, a non-fluorescent compound. This could reduce the fluorescence signal generated by the negative control cells whilst the background fluorescence in the negative control wells is still climbing, reducing the effective ratio between the two (Gong *et al.*, 2020).

Although the combination of factors detailed above appeared to be the best combination to use for the cobra screen, higher resazurin concentrations appeared to be toxic to the cells. Toxicity included cells appearing sparser and more rounded with the appearance of black dots next to the cells. It was unclear exactly what these black dots were, however, they increased in number alongside increasing resazurin concentration. It is thought that the black deposits were either resazurin or the accumulation of resorufin.

This toxic effect is not well documented in the literature, however Pace and Burg (2015) describe a similar toxic effect to multiple cell lines in culture when exposed to resazurin at higher concentrations or for prolonged periods of time. It is unclear from both the Pace and Burg study and this one whether it is resazurin or its reduced state resorufin that is causing this toxic effect but it should be borne in mind when designing resazurin based assays, particularly if cell

recovery post assay is desired. Gong et al. (2020) make a similar observation to that in the presented study there is a correlation between both resazurin concentration and resazurin exposure time with cell viability. They suggest that concentrations of 2 mM for 1 hour, 1 mM for 2 hours, 0.4 mM and 0.2 mM for 8 hours or 0.1 mM for 24 hours cause a significant decrease in cell viability. This suggests a strong correlation between all three factors (resazurin concentration, exposure time and resulting cell viability) which is also suggested by the presented study.

The manufacturer of alamarBlue (Invitrogen, California, USA) suggests using a 1 in 10 dilution of their product (giving a 44-56 µM final concentration according to the literature (Borra *et al.*, 2009, O'brien *et al.*, 2000)) and incubating for between 1-4 hours. They also suggest that the fluorescence or absorbance signal is stable for approximately 7 hours and so plates should be read within this time. The resazurin can then be washed off cells and they will continue to proliferate normally. One study concluded that a 10% alamarBlue solution applied to cells for 3 hours does not lead to significant toxicity even after resazurin was removed and cells cultured for a further two days (Pace and Burg, 2015). Other studies recommend that even at lower concentrations, long term incubation should be avoided if further work with the treated cells is desired, although these were in cell lines of non-cancerous origin (Gong *et al.*, 2020). In the presented study, further optimisation led to a final concentration of 160 µM resazurin with a final plate read after 120 minutes. Although a slightly higher concentration than presented by the supporting literature, these conditions were deemed to be suitable for the cell lines tested since the cells were of a cancerous origin and discarded immediately after the assay and not cultured or tested further.

It was decided that *Naja nigricollis* venom would be a more appropriate positive control for the assay than the DMSO that was being used previously. The results indicated the shortest venom incubation time needed to achieve the good S/B ratio needed for a positive control in a screening assay. This incubation time is shorter than used in other studies which tend to incubate venom with cells for between 24-72 hours (Ebrahim *et al.*, 2016, Teixeira *et al.*, 2016). However, the presented study indicated that any time points over 120 minutes do not appear to improve the S/B ratio, and thus this time point was used for all subsequent venom incubation.

The standard fluorescence settings used in other resazurin-based studies involve an excitation wavelengths of 530 nm and emission wavelength of 590 nm (O'brien *et al.*, 2000, Lim *et al.*, 2016, Hamid *et al.*, 2004). However, the plate reader used in the presented study did not have filters available for an excitation wavelength of 530 nm. Therefore, the three closest wavelengths (355 nm, 485 nm and 544 nm) were tested to establish the optimum settings to

use for the study. The emission wavelength was set to the standard 590 nm. In addition to fluorescence, absorbance at 595 nm was tested to compare absorbance against fluorescence readings. Reading the plate for fluorescence gave a much larger window than using absorbance readings with the optimal excitation and emission wavelengths being determined by Z' assay.

Optimum cell number was independently optimised for each of the tested cell lines and for 96- and 384-well plate formats. The cell densities selected for this assay (2 x $10^5$ cells/well for SW620 and 1 x $10^5$ cells/well for BxPC-3 cells) were higher than those used in other studies (7.5 x $10^3$ cells/well (Larsson *et al.*, 2020) 1.5x$10^4$ cells/well (O'brien *et al.*, 2000)) but these seeding densities provided the largest S/B ratios and best Z' of those tested.

After comparing S/B ratio, the raw fluorescence values were plotted to assess the linearity of the relationship between cell metabolism and raw fluorescence. After an initial log phase, the raw fluorescence values entered a lag phase whereby the fluorescence values plateaued. This effect was more prominent in wells with more cells compared to those with fewer cells. This is thought to be due to resazurin being a limited reagent in the reaction and there being a lack of 'free' resazurin available to be converted to fluorescent resorufin (Uzarski *et al.*, 2017). Furthermore, once converted to resorufin, further reduction may occur converting resorufin to non-fluorescent dihydroresorufin. This could reduce the fluorescence signal being generated whilst the fluorescence is still simultaneously climbing from conversion from resazurin to resorufin, thus maintaining a net constant fluorescence value (Gong *et al.*, 2020). For this reason, the assay plates were read within the time frame in which there was a linear relationship between cell viability and raw fluorescence value.

At both cell densities, the assay was deemed to be robust, as measured by Z' assay and therefore has the potential to be used in robotics platforms for ultra-high throughput applications. The assay has already been miniaturised to 384 well (Lim *et al.*, 2016, Shum *et al.*, 2008, Che *et al.*, 2009) and 1536 well formats (Shum *et al.*, 2008) by other research teams. Resazurin provides an excellent, affordable, non-destructive alternative to MTT (Hamid *et al.*, 2004, Emter and Natsch, 2015) or other similar cell viability assays without the need for radioactivity.

Because the resazurin assay is designed to be binary, any inhibitors which show only partial inhibition at the chosen dose may suffer from having a large variability between replicates. This variability may be due to true biological differences or may be caused by random errors or 'noise' such as pipetting errors, inconsistently mixed samples prior to application, partial sedimentation of samples, robotic malfunction or evaporation of solvent leading to differences in compound

concentration (Malo *et al.*, 2006). There may also be large plate-to-plate variability, although this can be mitigated by the inclusion of appropriate controls on each plate.

### 3.3.2 Cobra Venom Screening

The results from the cobra screen showed that although all of the venoms tested inhibited both cell lines, the venoms causing the most significant inhibition were the African spitting cobra venom group in both cell lines at the higher concentration and selectively in the SW620 cell line in the lower concentration. At the lower venom dose, the *Naja nubiae* and *Naja haje legionis* venoms appear to be causing a small but significant level of stimulation rather than inhibition. This could be due to activation of a stress response leading to a protective response with the activation of pro-survival signalling cascades at the lower dose which is replaced by pre-death signalling in the higher venom dose treated cells (Fulda *et al.*, 2010). This effect could also be due to the presence of growth factors such as VEGF in snake venom which could contribute to short-term growth stimulation when the venom toxicity is low (Ferreira *et al.*, 2021, Chapeaurouge *et al.*, 2018).

A relationship appears to exist between cytotoxicity and spitting trait in African spitting species with spitting species causing high cytotoxicity and non-spitting species causing a much reduced cytotoxic effect (Panagides *et al.*, 2017). No such relationship was found between Asian spitting and non-spitting species, however (Panagides *et al.*, 2017). African spitting cobras also show the highest potency for activating sensory neurons and causing a pain response in sensory neuron cell line models when compared to the other cobra types (Kazandjian *et al.*, 2021).

Comparing the results of a chicken embryo test between spitting and non-spitting cobra venoms, a strong effect of spitting cobra venom on ocular vasculature was identified, inducing vasodilation and reduction of blood flow in the eye at high doses (Pollakova *et al.*, 2021)

In this study, the presence of more potent cytotoxic compounds in the African spitting cobras could explain why the venoms tested from this group appeared to have significant inhibitory effect on both cell lines, but particularly the SW620 cell line.

The spitting trait evolved convergently in both African and Asian species. The main selective pressure for these evolution events is assumed to be as defence against predation. The divergence of hominin species from other homonids and their occupation of African savanna habitats 7 million years ago (MYA) could have contributed to the evolution of more cytotoxic venom components in *Afronaja* species approximately 6.7 MYA (Pozzi *et al.*, 2014, Kazandjian *et al.*, 2021). This theory is supported by the convergent evolution of the spitting trait in some Asian species of cobra corresponding to the movement of *Homo erectus* through Asia

approximately 2.5 MYA (Prat, 2018). Exposure to the same selective pressure for an additional 4.5 MY may explain why African spitting cobras evolved this trait considerably earlier than their Asian counterparts and why they evolved more potent cytotoxic compounds in their venom.

Venoms, regardless of spitting/non-spitting and geographical considerations, display Newtonian behaviour and show no significant difference in flow properties or physical properties including pH, viscosity and protein concentration (Avella *et al.*, 2021).

Venoms combine hundreds to thousands of diverse compounds and biomolecules. Generally speaking, false negatives are less likely to occur when compounds are pooled together (Malo *et al.*, 2006). It is possible, however, that multiple compounds are present which have opposing effects on growth or proliferation. It is therefore important to separate the venom components out from the mixture as early as possible to gain better insight into which constituents have activity. All venoms from the 'African spitting' group of cobras were taken forwards as active samples to be separated by high performance liquid chromatography (HPLC) and screened again (Chapter 4).

# Chapter 4 - Optimisation of High Performance Liquid Chromatography for Venom Screening against Cancer Cells

## 4.1 Introduction

Venoms are complex mixtures of hundreds to thousands of different types of bioactive compounds including proteins, peptides, salts, neurotransmitters and organic components (Casewell *et al.*, 2013, Prashanth *et al.*, 2017). When analysing mixtures such as whole venoms, it can be difficult to know exactly what the effects are as there could be many opposing reactions occuring simultaneously. For example, one component in a venom could be increasing blood pressure whilst another is reducing it (Sitprija and Sitprija, 2012). The presence of opposing effects could mask some potentially interesting effects and make them difficult to identify. It is therefore important that individual venom components are separated as early as possible.

Whole venoms are not usually suitable as potential medicines as they contain such a diversity of different components, many of which could lead to a variety of unintended side effects that may be detrimental to the patient. If the compound(s) that is/are eliciting the intended effect can be separated from the mixture, that single entity may be further studied and optimised, away from the other components which do not contribute to the intended effect.

There are two main methods used for utilising venoms in drug discovery. This includes fractionating venoms and screening them against a biological target of interest, or alternatively by performing 'venomics' where each venom component is individually identified and characterised (Oldrati *et al.*, 2016, Vetter *et al.*, 2011). Both of these methods have their advantages and disadvantages and using a combination of both is the optimal approach (Prashanth *et al.*, 2017).

### 4.1.1 High Performance Liquid Chromatography

Chromatography is a widely utilised analytical chemistry separation method which uses the intrinsic properties of compounds in a mixture to move through the system at different rates in order to separate them (Mcmurry, 2011). High Performance Liquid Chromatography (HPLC) is an effective method, able to separate mixtures containing many similar compounds with good resolution. A stationary phase provides a surface for compounds to interact with in some manner (dependant on the type of HPLC) and a mobile phase that flows through the system, moving the components through the system at a rate dependant on the type of HPLC and properties of each compound (Meyer, 2013).

There are many different types of HPLC which use the intrinsic properties of chemical entities to separate them from a mixture. “Normal phase” HPLC uses a relatively polar stationary phase and relatively non-polar mobile phase. “Reverse phase” (RP) HPLC is, as the name suggests, the same as “normal phase”, only reversed - relatively non-polar stationary and polar mobile phases (Meyer, 2013). Reverse phase HPLC is currently the preferred method of separation of peptides prior to mass spectrometry analysis (Xie *et al.*, 2012, Zhang *et al.*, 2010).

Size exclusion chromatography (SEC, sometimes known as gel filtration chromatography) uses a gel matrix to separate molecules from a mixture based on their size. Pores within the matrix of varying sizes will either include or exclude the molecules, depending on the size of the pore and the size of the molecule. Large molecules are excluded from the pores and flow through the column fastest and are eluted first. The smaller the molecule, the more pores the molecules may diffuse into and the longer they take to pass through the column, eluting later. Size exclusion chromatography therefore separates molecules from a mixture and elutes them in decreasing molecular weight order (Hagel, 2001, Andrews, 1965).

Ion exchange chromatography consists of either anions (such as $SO_3^-$) or cations (such as $N(CH_3)_3^+$) covalently bonded to the stationary phase with a liquid mobile phase. Ions in the solute of the opposite charge are attracted to the stationary phase and are slowly eluted according to their charge and the pH or charge of the mobile phase (Harris, 2010).

The characteristics utilised for separation of proteins by reverse phase and size exclusion chromatography are reasonably fixed and not altered by the buffers used in the separation process. However, in ion exchange chromatography, the overall charge of proteins can be readily manipulated by pH due to their amphoteric characteristics. Protein recovery, retention time and peak resolution can be altered by differing pH and salt concentrations in the mobile phase buffers during ion exchange chromatography (Kopaciewicz and Regnier, 1983).

Using multiple types of HPLC has several advantages which are described by Stoll and Carr (2017). If the mixture is complex, one dimensional HPLC is unlikely to be able to provide adequate separation into single components, leading to poor chromatographical resolution and contamination due to multiple components ‘bleeding’ into each sample. In addition, components which are very similar in the property selected for separation (e.g. similar size for SEC) may not be adequately separated in a single dimension of separation. A simplified diagram of the two-dimensional separation process is shown in Figure 4.1.

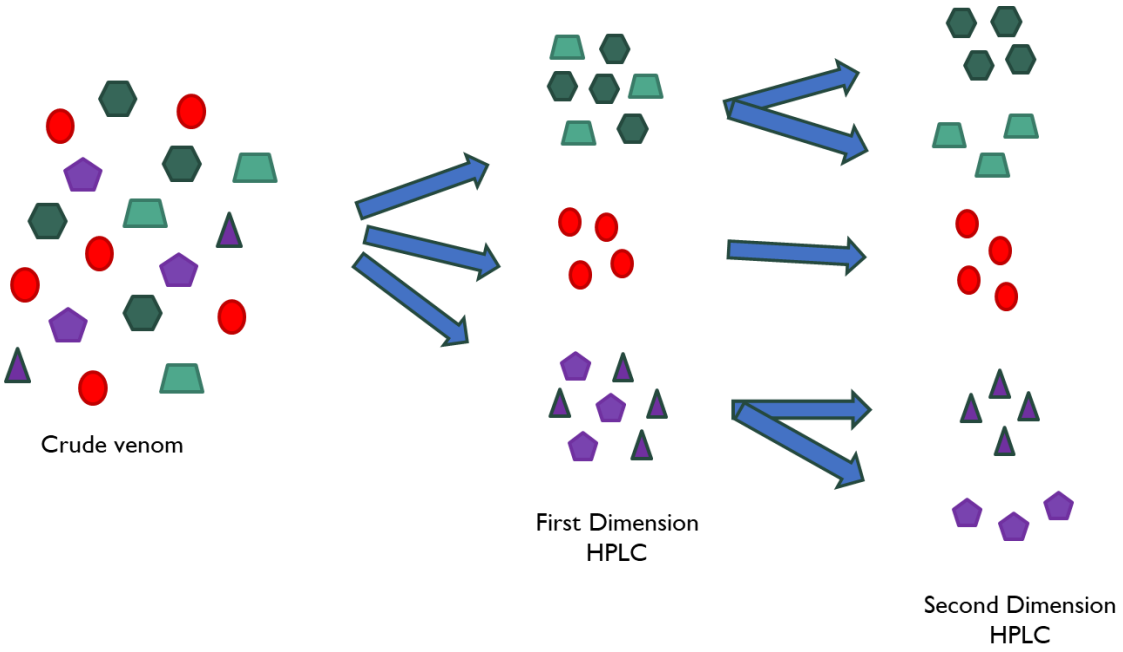



*Figure 4.1. A simplified diagram showing the two-dimension venom separation process.*

*Whole venom is represented on the left as a mixture of components with different properties - in this case shape and colour. The first dimension of HPLC separates based on one property - colour - however, this process cannot always separate colours that are very similar. The second dimension separates based on a different property - shape- and helps to further separate the different components. Real properties for separation could include size, ionic charge, lipophilicity, etc.*

Two-dimensional (2D) separations generally take longer then would a one-dimensional separation, however there can be a stark improvement in separation quality for more complicated mixtures. To obtain better resolution of 1D HPLC for complex mixtures, the run time must be extended to the point of impracticality, whilst running two separate dimensions becomes much more time and resource efficient (Stoll and Carr, 2017). The difference between 1D HPLC separation using columns with different peak separation capacities compared to 2D separation is illustrated in Figure 4.2. 2D HPLC separation reaches a level of separation of compounds unobtainable by 1D HPLC efforts, even when more complicated mixtures are used.

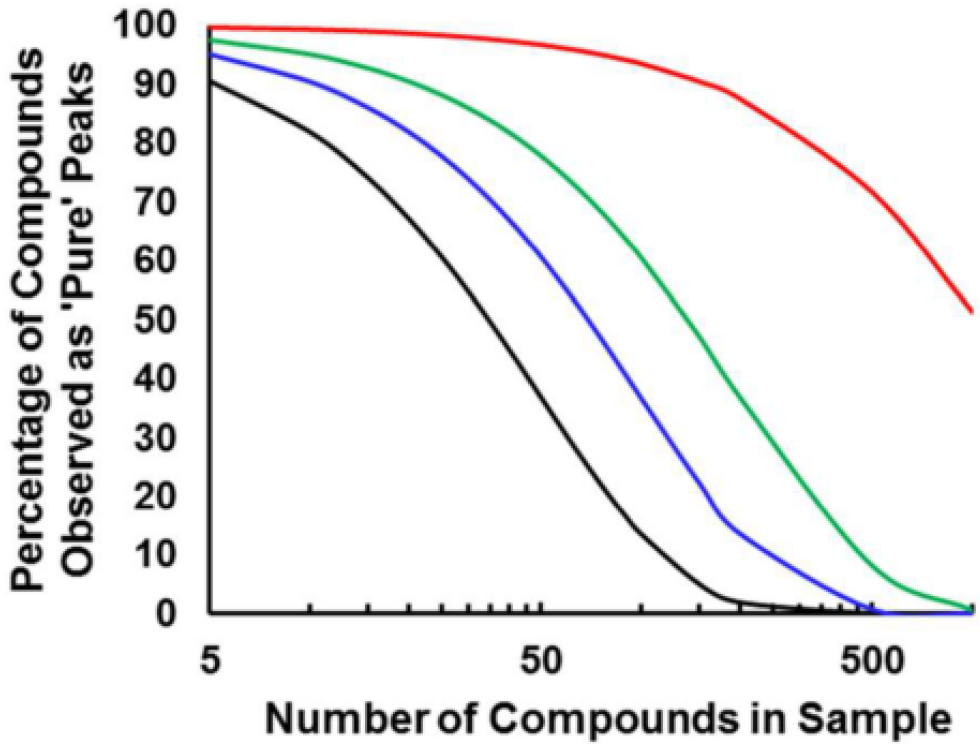



*Figure 4.2. 1D vs 2D HPLC separation capacities.*

*Logarithmic graph illustrating the percentage of pure compounds obtained from a mixture of different number of compounds following 1D HPLC with columns with separation capacities of 100 (black), 200 (blue) and 400 (green) compared with an effective peak capacity of 3000, not obtainable by 1D HPLC but readily reached by 2D HPLC separation (red). Figure reprinted with permission from Stoll, D. R. & Carr, P. W. 2017. Two-Dimensional Liquid Chromatography: A State of the Art Tutorial. Anal. Chem., 89, 519-531. Copyright (2017) American Chemical Society.*

Samples are diluted with each dimension and some losses may be experienced due to less than 100% recovery of samples (Pirok *et al.*, 2019). This effect may be exacerbated if there is a prolonged second dimension phase as first dimension samples will then be left incubating for a long period of time, potentially increasing protein degradation. This effect may be overcome by running a slow, controlled first dimension phase followed by a rapid, efficient second dimension process, minimising the length of time the first dimension fractions are left vulnerable to degradation (Pirok and Schoenmakers, 2018).

There could be compatibility issues with buffers from different phases of HPLC as some rely on ionic interactions or extreme pH conditions (Pirok *et al.*, 2019). Care should be taken to ensure buffers are compatible both with the samples to be tested but also all columns involved in the separation process.

### 4.1.2 Separation of Peptides and Proteins by RP HPLC

Peptide molecules have a diversity of functional groups including acidic, basic, hydrophobic and hydrophilic functional groups. Proteins and peptides may be separated from a mixture using the properties of these functional groups including the size, charge and hydrophobicity of the amino acids that make them up. The reproducibility of sample separation is important to ensure consistency between runs. Peptide retention standards may be used to ensure quality and reproducibility of HPLC runs (Klaassen *et al.*, 2019).

Different amino acids have different hydrophobicity values ranging from tryptophan (most hydrophobic) to lysine (most hydrophilic) (Guo *et al.*, 1986). As such, each amino acid has a particular retention time on a reverse phase HPLC column, dependant on the pH of the buffers (Guo *et al.*, 1986). Using this principle, amino acids may be separated from a mixture by RP HPLC and, assuming parameters are kept consistent, elution order should be predictable and consistent on subsequent repeat runs.

The differentiation between what constitutes a protein or a peptide is reasonably arbitrary, although some sources claim that peptidic molecules larger than 3-5 kDa should be classified as proteins (Tibbitts *et al.*, 2016). As molecules get larger, the two- and three-dimensional structures often become more complex. This may include having hydrophobic regions of larger proteins completely internalised when in an aqueous buffer. Features such as hydrogen bonding and disulphide bridges can have a large effect on protein conformation, and this can have a profound effect on HPLC retention time that can be difficult to predict (Mant and Hodges, 2017). Peptide bonds absorb light at around 210-220 nm whilst aromatic side chains of proteins absorb at 250-290 nm. Most proteins or polypeptides will contain tyrosine and so absorption of light may be measured in this range in order to estimate their concentration in solution (Mant and Hodges, 2017).

A combination of cation exchange followed by RP HPLC is a common technique used for peptide separation to improve separation quality and chromatographic resolution (Phillips *et al.*, 2010). More recently, deep learning algorithms have been applied to attempt more accurate peptide retention time prediction in liquid chromatography (Bouwmeester *et al.*, 2019, Ma *et al.*, 2018).

### 4.1.3 Separation of Venom Samples by HPLC

There are numerous research groups investigating the effect of animal venoms and many of these groups use HPLC to separate venoms to study individual components. It is common in venom research to use reverse phase HPLC with buffers containing trifluoroacetic acid (TFA) and an increasing gradient of acetonitrile (ACN). Fractions are usually collected at either 215 or 280 nm. Table 4.1 details the HPLC conditions that have been used for a variety of different

venomous species. For better separation, some groups use more than one dimension of HPLC. This includes SEC and IE chromatography, usually combined with RP HPLC (Nascimento *et al.*, 2006, Gao *et al.*, 2008, Bonfim *et al.*, 2001).

*Table 4.1. Reverse phase HPLC details for venom separation in the literature.*

*HPLC has been used to separate a variety of different venomous animals utilising trifluoroacetic acid (TFA) and acetonitrile (ACN) buffers.*

| **Species** | Type | TFA (%) | ACN (%) | Flow rate (ml/min) | Time (mins) | Wavelength (nm) | Reference |
|---|---|---|---|---|---|---|---|
| ***Eristocophis macmahoni*** | Snake (Viper) | 0.05 | 0-60 | 1 | Not specified | 280 | (Ali *et al.*, 1999) |
| ***Cerastes sp*. and *Macrovipera sp.*** | Snake (Vipers) | 0.05 | 5-70 | 1 | 108-160 | 215 | (Bazaa *et al.*, 2005) |
| ***Bothrops jararacussu*** | Snake (Pit viper) | 0.1 | 0-66.5 | 2 | 60 | 280 | (Bonfim *et al.*, 2001) |
| ***Naja sp.*** | Snake (Cobras) | 0.1 | 5-95 | 1 | 60 | 280 | (Nawarak *et al.*, 2003) |
| ***Bitis arietans*** | Snake (Puff Adder) | 0.1 | 5-70 | 1 | 145 | 215 | (Juàrez *et al.*, 2006) |
| ***Sistrurus sp.*** | Snake (Rattlesnake) | 0.1 | 5-70 | 1 | 165 | 215 | (Sanz *et al.*, 2006) |
| ***Oreumenes decoratus*** | Wasp | 0.1 | 5-65 | 2.5 | 30 | 215 | (Konno *et al.*, 2007) |
| ***Buthus martensii Karsch*** | Scorpion | 0.1 | 0-80 | 1 | 30 | 280 | (Gao *et al.*, 2008) |

### 4.1.4 Nano Differential Scanning Fluorimetry (NanoDSF)

Protein stability can be measured using various techniques. Fluorescent or radioactive labelling is required by some biomolecule exploration methods in order to generate a signal (Chen *et al.*, 2000). Some methods require immobilisation of protein surfaces, such as surface plasmon resonance (Myszka, 1997). This can interfere with readings and some methods require optimisation and time consuming and expensive clean-up operations (Seidel *et al.*, 2012). A good alternative method is isothermal titration calorimetry (ITC) which is also label-free and can determine the thermodynamics of protein interactions in solution. However, this method requires large quantities of proteins which are not always available in venom studies (Falconer and Collins, 2011).

Nano differential scanning fluorimetry (nanoDSF) technology uses the intrinsic fluorescence present in amino acids such as tryptophan and tyrosine to allow protein stability to be investigated without the requirement labels or immobilisation (Maschberger *et al.*, 2015). Changes in fluorescent emission properties upon heating can be measured and a graph

illustrating protein unfolding may be generated (Maschberger *et al.*, 2015). This method may be used with any buffer, including those containing detergents, and at any protein concentration.

## 4.2 Chapter 4 Aims

The aims of this chapter were to:

- optimise the HPLC separation process
- separate active cobra venoms from Chapter 3 including venoms from *Naja mossambica, Naja nigricincta, Naja nigricollis, Naja nubiae* and *Naja pallida* and screen the resulting fractions
- assess stability of samples using nanoDSF
- select venom fractions that cause selective inhibition against BxPC-3 cells over SW620 cells

A simplified diagrammatic representation of the final optimised process described in this chapter is shown in *Figure 4.3*.

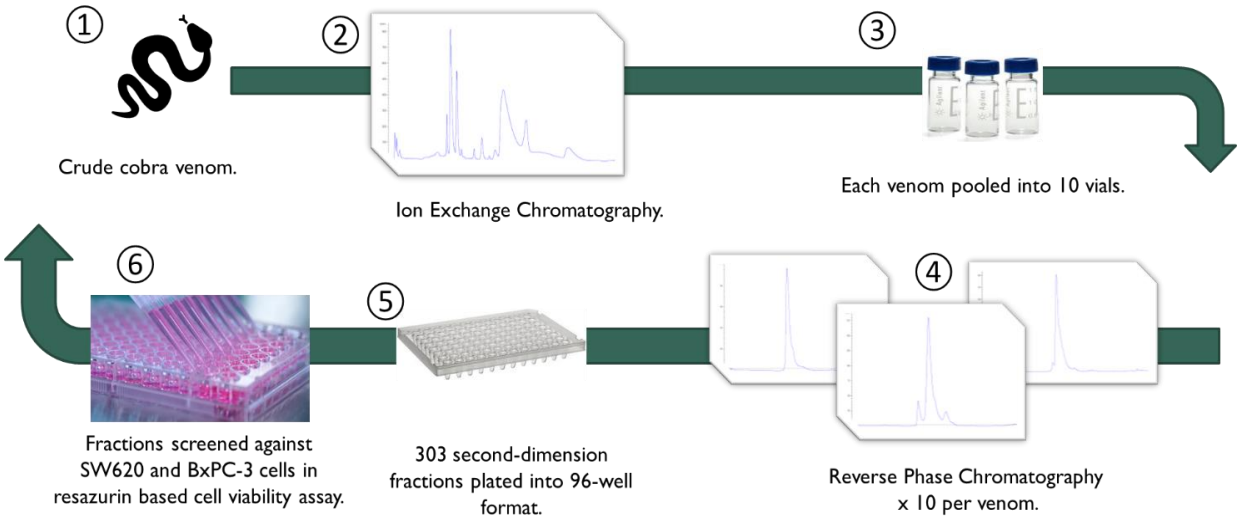



*Figure 4.3. Simplified diagrammatic representation of the final optimised cobra venom separation processes involved in chapter 4.*

*Each whole venom was separated by ion exchange chromatography, then re-pooled into 10 samples, separated by RP chromatography, then screened against SW620 and BxPC-3 cells in a cell viability assay.*

## 4.3 Results

### 4.3.1 First Dimension Reverse Phase HPLC

The five African spitting cobra venoms identified as 'hits' in Chapter 3 – *Naja mossambica, Naja nigricincta, Naja nigricollis, Naja nubiae* and *Naja pallida* – were separated into their individual components. Reverse phase HPLC was performed according to section 2.5.1.

The venom samples were run using 50 µg venom protein to obtain sharp peaks and aid good separation. The analytical fractions were not collected or tested. The analytical traces are shown run at 215 nm (Figure 4.4) and 280 nm (Figure 4.5). Later, preparative HPLC with up to 10 mg venom protein was run and the resulting fractions were collected and screened against the cancer cell lines. The preparative first dimension HPLC traces are shown in Figure 4.6.

### 4.3.2 Screening the First Dimension RP HPLC Fractions Against Cancer Cells

Following the preparation of the fractions during first dimension RP HPLC, the fractions were each screened against BxPC-3 and SW620 cells to determine which of the fractions retained cytotoxic activity. Before data analysis could be performed, the data needed to be normalised.

#### *4.3.2.1 Data Normalisation*

Raw fluorescence data collected from the plate reader required normalisation so that any plate-to-plate variability was minimised. Three types of normalisation method (percentage of control, normalised percentage inhibition (NPI) and Z score) were compared to assess which would be the most appropriate to use. The resulting graphs from the three different methods can be seen in Figure 4.7. The percentage of control and NPI graphs ended up looking almost identical to one another whilst the Z score appeared slightly different. Unlike the other two methods, the Z score has a statistics basis and does not rely on controls to deliver its values. Sometimes, however, the Z score is not appropriate for data analysis due to the presence of too many or too few hits. Figure 4.8 illustrates an example of when Z score would not be the most appropriate method to use.

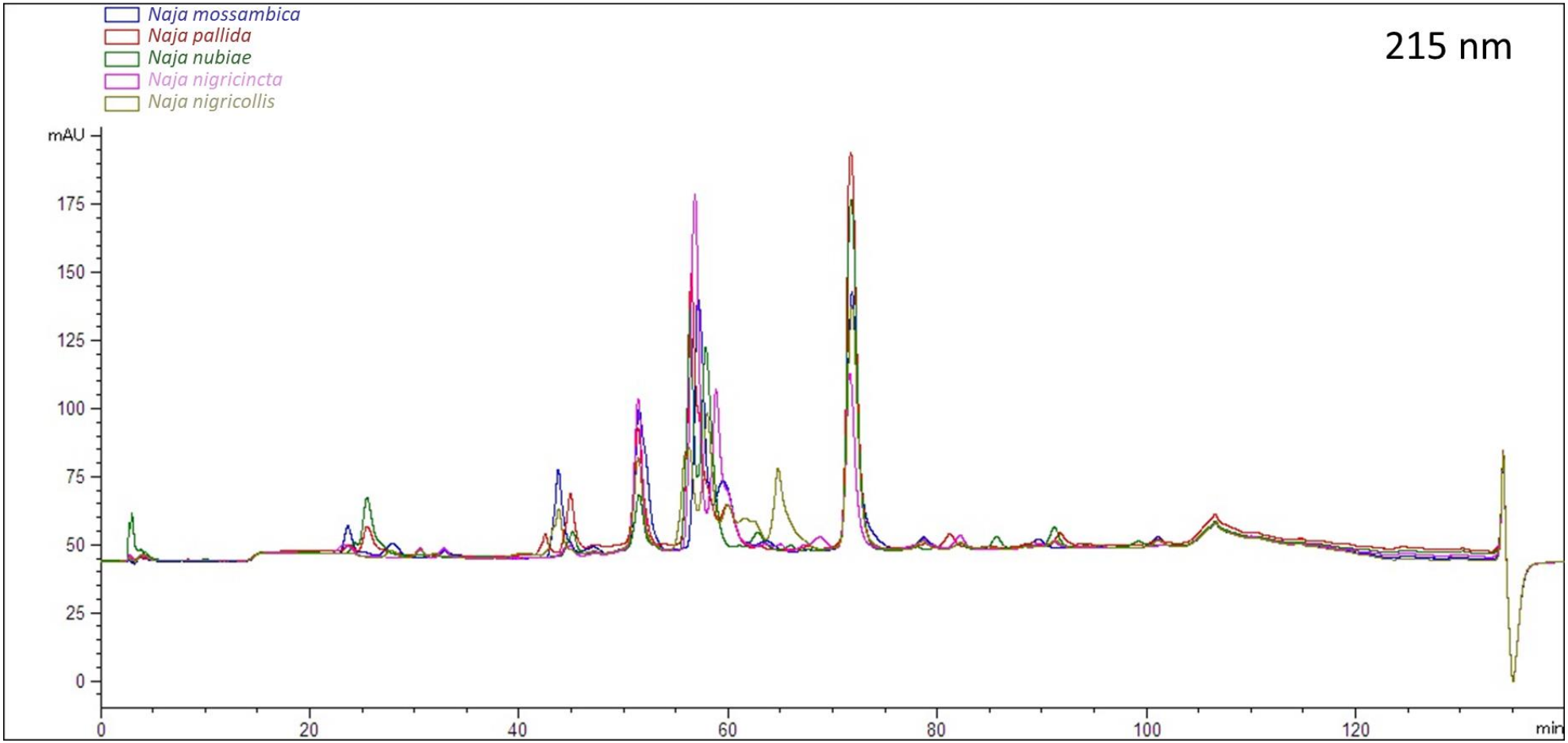



*Figure 4.4 Analytical HPLC traces for the five African Spitting cobras (215 nm)*

*The trace shows the retention time (mins) plotted against the detector response (mAU) detected at 215 nm wavelength.*

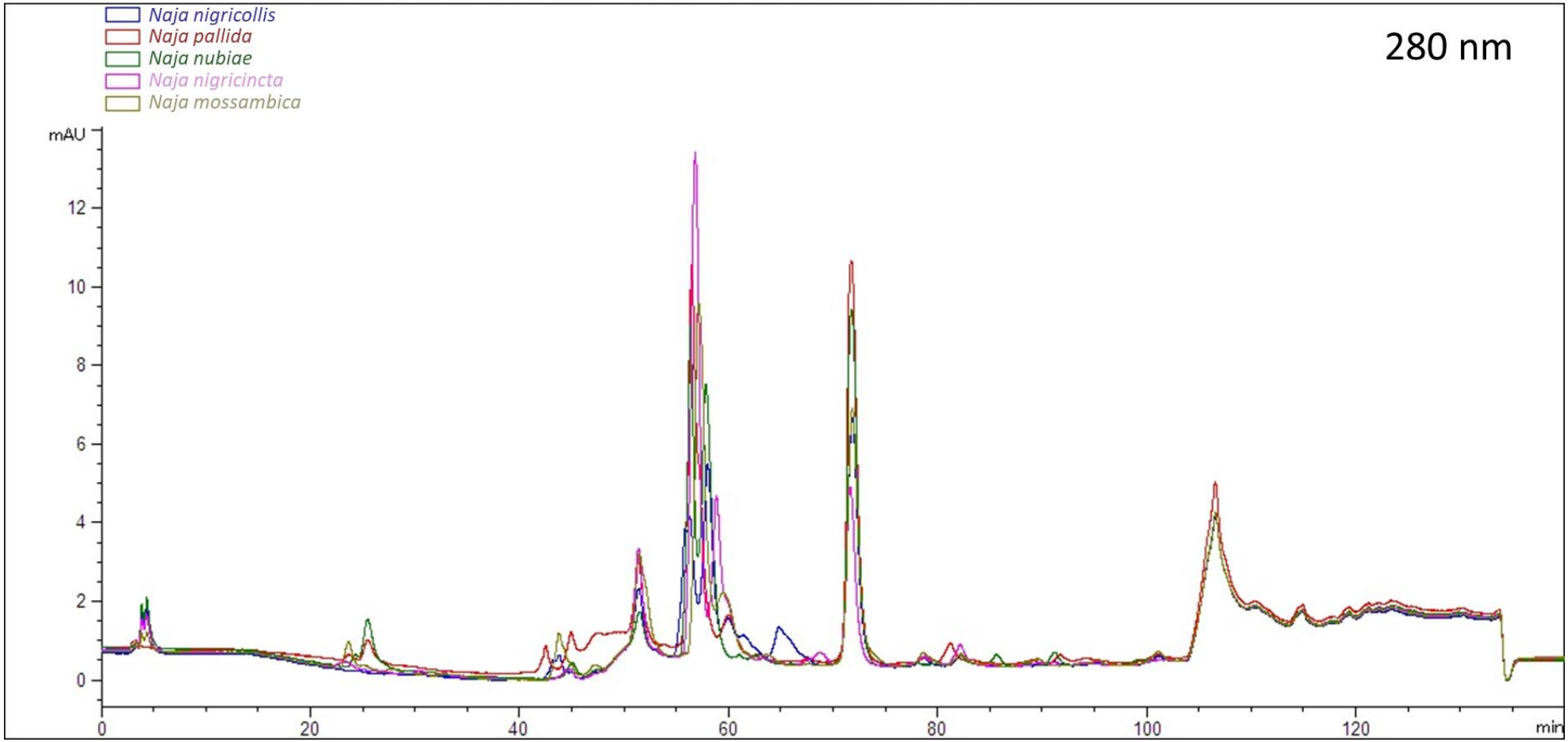



*Figure 4.5. Analytical HPLC traces for the five African Spitting cobras (280 nm).*

*The trace shows the retention time (mins) plotted against the detector response (mAU) detected at 280 nm wavelength.*

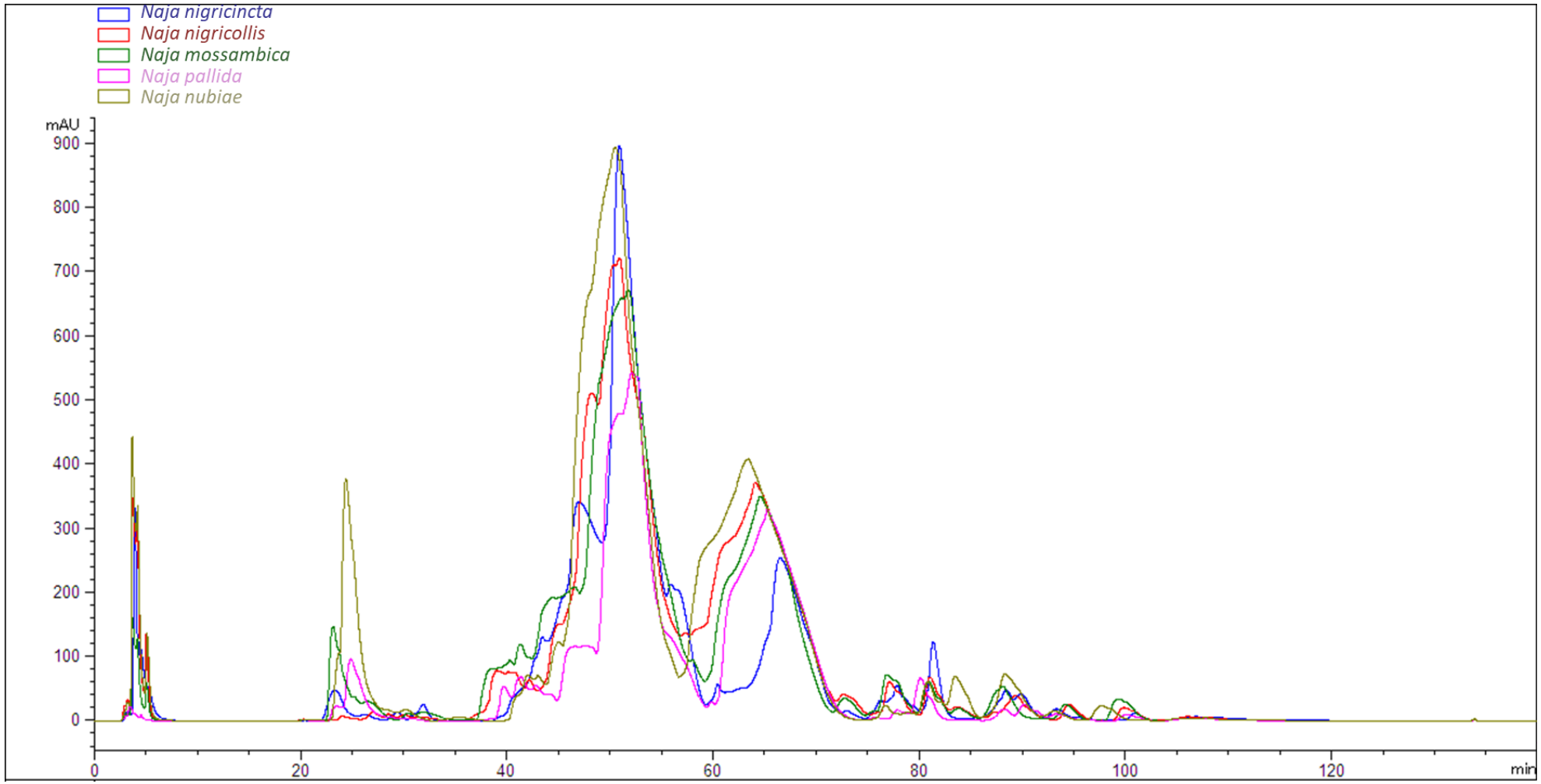



*Figure 4.6. Preparative HPLC traces for the five African Spitting cobras.*

*The traces show the retention time (mins) plotted against the detector response (mAU). All preparative HPLC fractions were collected at 280 nm.*

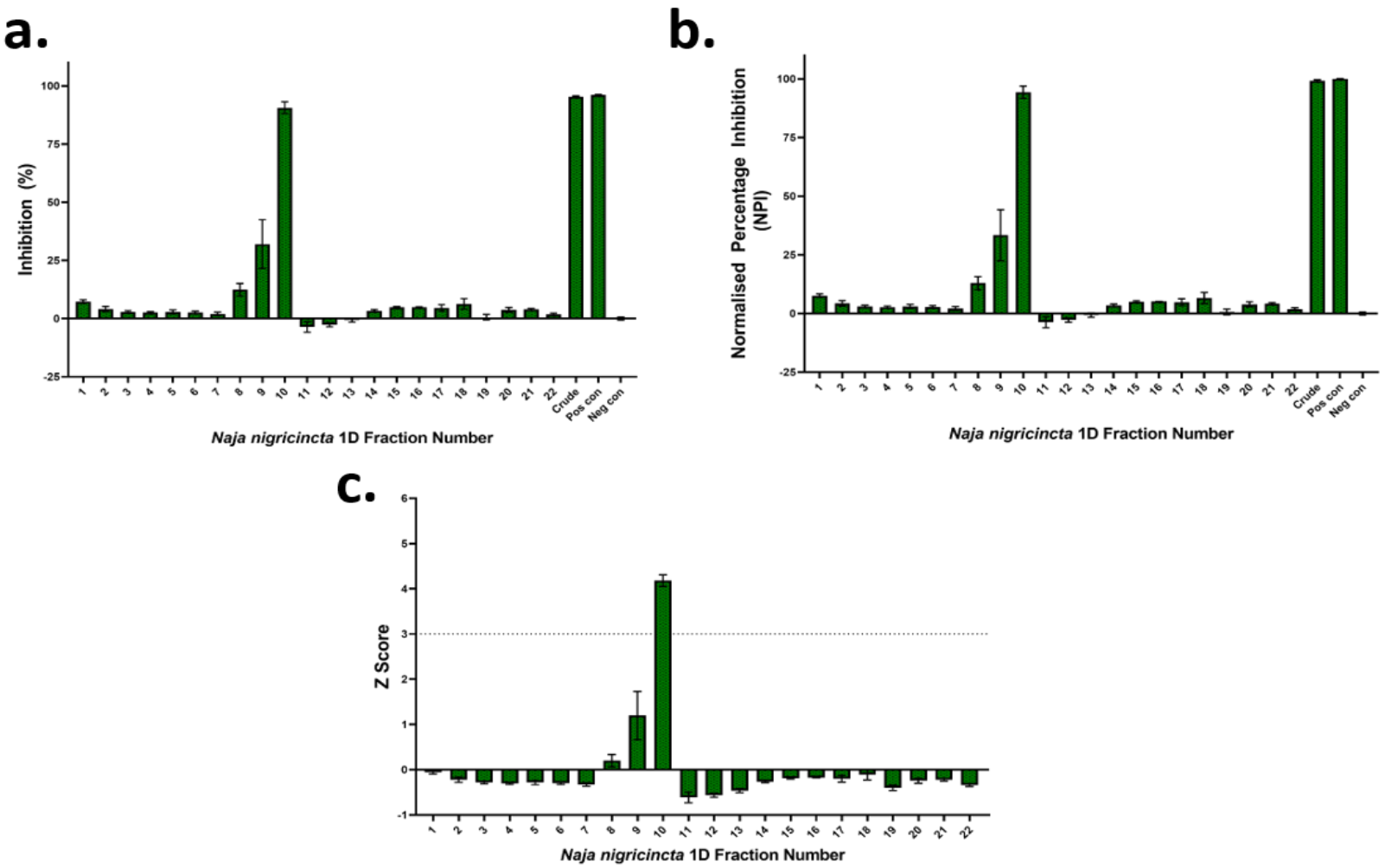



*Figure 4.7. Comparison of different normalisation methods.*

*(a) percentage of control, (b) normalised percentage inhibition (NPI), (c) Z score.*

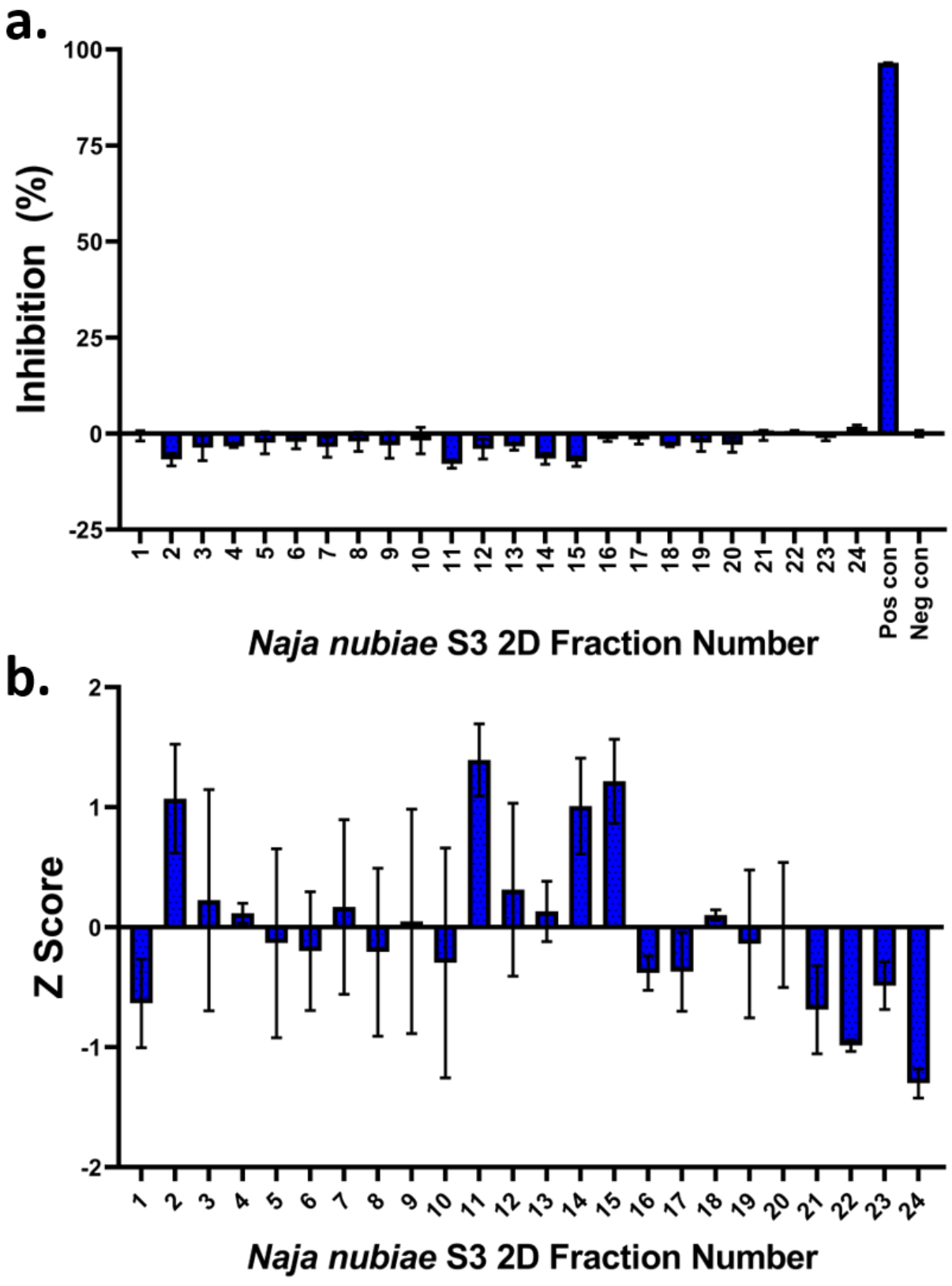



*Figure 4.8. Example of when Z score is not an appropriate method for analysis.*

*Because there are no distinct hits (as shown by percentage inhibition in (a), the baseline for the Z score is such that very small changes in cell viability lead to seemingly large changes in Z score (b).*

#### *4.3.2.2 Data Analysis & Statistics*

The raw fluorescence values for each first-dimension cobra venom fraction screen were normalised using the 'Z score' method. For this particular data set, a 'Z score' of ±3 was set as an appropriate threshold according to Brideau *et al*. (2003). The alternative would have been to rank the results and select an appropriate proportion to take forwards for further work (Brideau *et al.*, 2003). The results are shown in Figure 4.9. Fractions above the threshold line at y=3 were deemed to be significant based on Z score. There was one fraction from each of the five cobras tested that reached significance – *Naja nigricincta* RP fraction 10, *Naja nubiae* RP fraction 10, *Naja pallida* RP fraction 10, *Naja mossambica* RP fraction 13 and *Naja nigricollis* RP fraction 11. These results are included in the summary in Table 4.2.

Statistical significance was also tested in each of the different species cobra venom screens using Kruskal-Wallis (KW) testing. Each of the five sets of cobra venom fractions tested gave a significant result ($P<0.0001$) in KW testing. Dunn's multiple comparisons test was run as a non-parametric post-hoc test to identify which of the fractions was having the significant effect. Since the Z score was used for the statistical analysis and thus there were not control wells with which to compare, the active fractions were deemed to be those causing a statistically significant effect against at least one other fraction. The fractions deemed to be statistically significant are listed in Table 4.2. Since active fractions are essentially outliers compared to the bulk of the data, an outlier test (ROUT) was also performed to verify any findings from the Kruskal-Wallis with Dunn's multiple comparisons test. The results of the tests are shown in Appendix II and summarised in Table 4.2. For *Naja mossambica*, all three tests performed suggested that fraction 13 was the only active fraction. For *Naja nigricincta* and *Naja nigricollis* venoms, two fractions were identified by at least one test (fractions 9 and 10; 11 and 12 respectively) and therefore both fractions were taken forwards. *Naja nubiae* and *Naja pallida* both had their fraction 10 identified as significant and so these fractions were taken forwards.

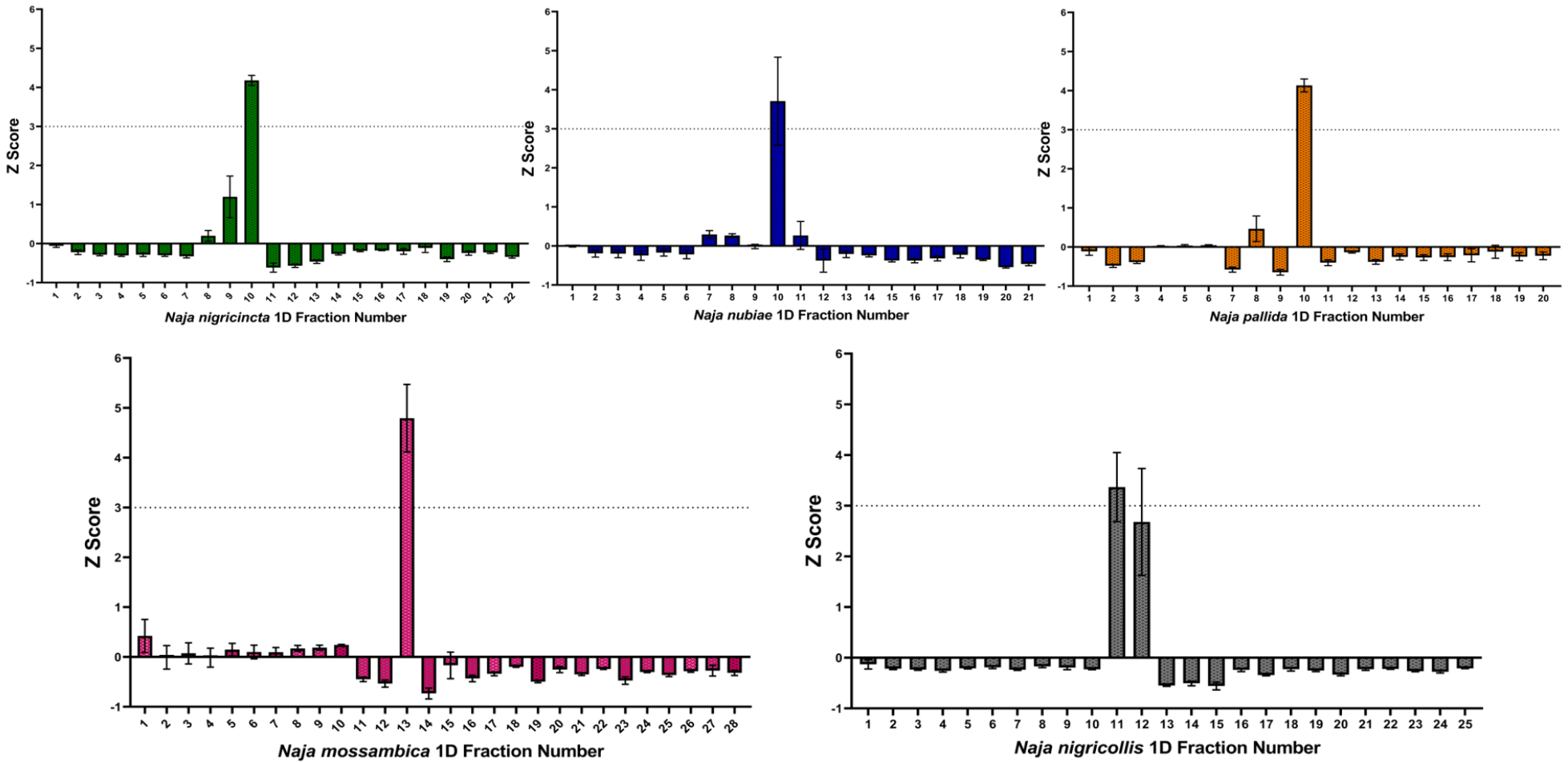



*Figure 4.9. Screening results of the first-dimension RP HPLC cobra screening fractions.*

*Fraction number plotted against Z score in each case. Threshold set to Z score ≥3. Error bars are SEM based on n=3.*

*Table 4.2. Summary of first dimension cobra fraction screening.*

*For each species, the fractions identified by each test (Z score (≥3), Kruskal-Wallis (KW) followed by Dunn's multiple comparisons test (Dunn) and ROUT outlier test) are displayed. Statistical testing performed on the Z scores for each set of data, n = 3, * significant at p<0.05 (Dunn's multiple comparisons test only).*

| Species & Test | Significant fraction(s) | Significance level | Compared to fraction… |
|---|---|---|---|
| ***Naja mossambica*** | | | |
| Z score | 13 | | |
| KW & Dunn | 13 | * | 14 |
| ROUT outlier test | 13 | | |
| ***Naja nigricincta*** | | | |
| Z score | 10 | | |
| KW & Dunn | 10 | * | 11 & 12 |
| ROUT outlier test | 9, 10 | | |
| ***Naja nigricollis*** | | | |
| Z score | 11 | | |
| KW & Dunn | 11, 12 | *, * | 13 & 15, 15 |
| ROUT outlier test | 11, 12 | | |
| ***Naja nubiae*** | | | |
| Z score | 10 | | |
| KW & Dunn | 10 | * | 20 |
| ROUT outlier test | 10 | | |
| ***Naja pallida*** | | | |
| Z score | 10 | | |
| KW & Dunn | 10 | * | 7 & 9 |
| ROUT outlier test | 10 | | |

### 4.3.3 Second Dimension HPLC

#### *4.3.3.1 Second Dimension SEC Separation and Screening*

Before separating the active venoms further, the buffers to be used for the second-dimension size exclusion chromatography (SEC) were optimised. Results of the optimisation are in Figure 4.10. The first run (Figure 4.10a) showed undefined peaks that did not go back to the baseline and a large peak that did not fully elute before the end of the run. To try and capture this cut-off peak, the run time was increased from 20 to 40 minutes and a reduced downslope threshold was used (Figure 4.10b). This still resulted in peaks that did not go back down to baseline and had a large, undefined peak from 23 minutes. A mixture of IE buffer A and sodium chloride containing buffer B (see Table 2.2 for buffer compositions) resulted in much sharper peaks that went back to baseline and a more compact, well defined final peak.

Second dimension HPLC (SEC) was performed according to section 2.5.2 on the active 1D RP fractions identified in section 4.3.2 (*Naja mossambica* fraction 13, *Naja nigricincta* fractions 9 and 10, *Naja nigricollis* fractions 11 and 12, *Naja nubiae* fraction 10 and *Naja pallida* fraction 10) to separate the venom components further. These fractions were then screened against SW620 cells to see if any of the fractions caused inhibition in this cell line. The results are displayed in in Figure 4.11 and show that some of the *Naja nigricincta* fractions caused inhibition against SW620 cells, with N.nct_r10s8 reaching significance according to the Z score (≥3). However, a Kruskal-Wallis test was performed with Dunn's multiple comparisons post-test and none of the samples showed statistical significance against the negative control sample. Results of normality tests, QQ-plot and Kruskal-Wallis with Dunn's multiple comparisons tests are shown in Appendix II.

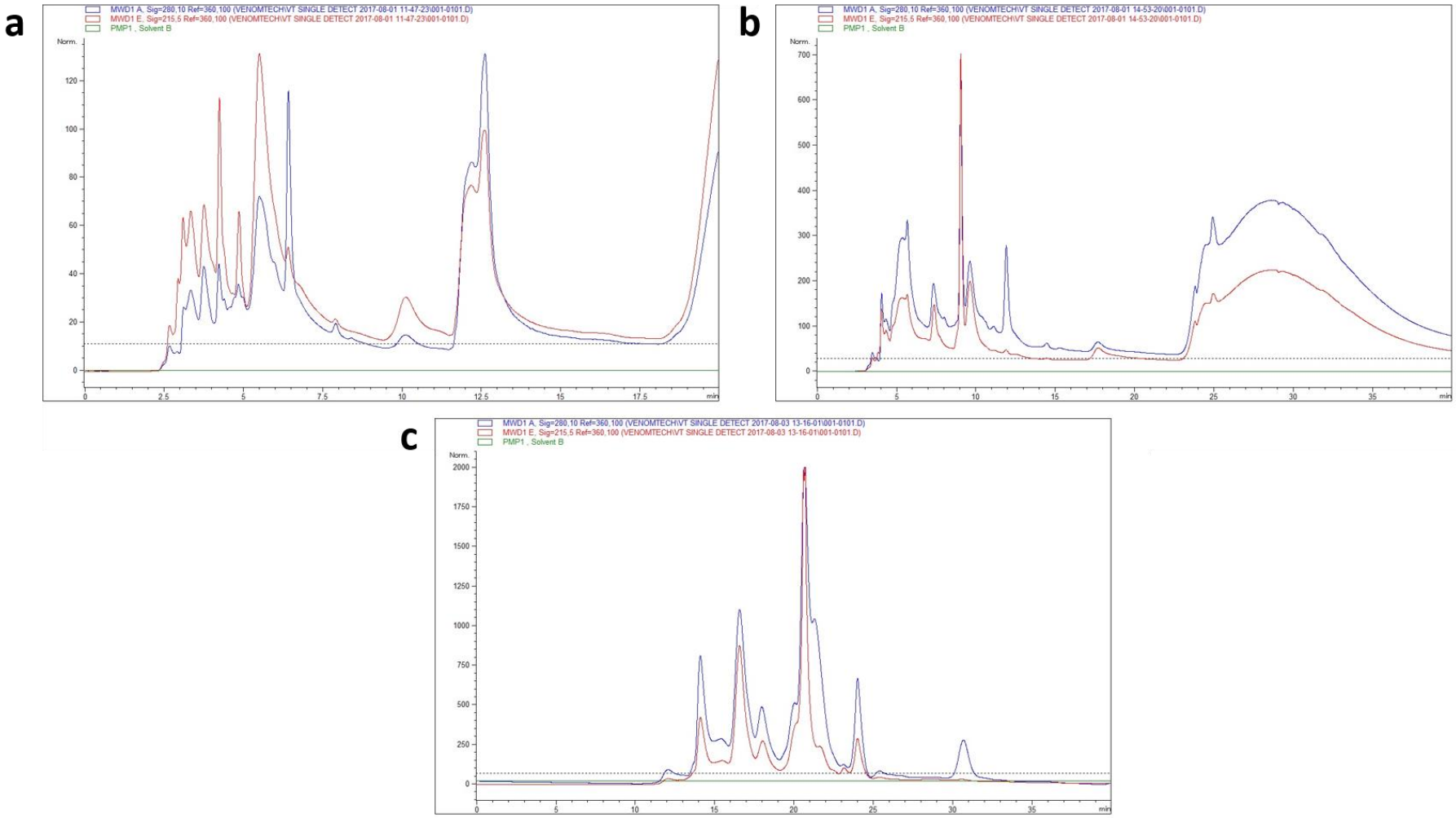



*Figure 4.10. Optimisation of the ion exchange buffers for use in SEC HPLC.*

*(a) 100% IE buffer A, 20-minute run. (b) 100% IE buffer A, 40-minute run, downslope threshold modified from 0.5 U/s to 0.2 U/s. (c) Buffer composition changed to 80% IE buffer A, 20% IE buffer B, 40 minute run. Blue chromatograms = signal read at 280 nm, red chromatograms = signal read at 215 nm.*

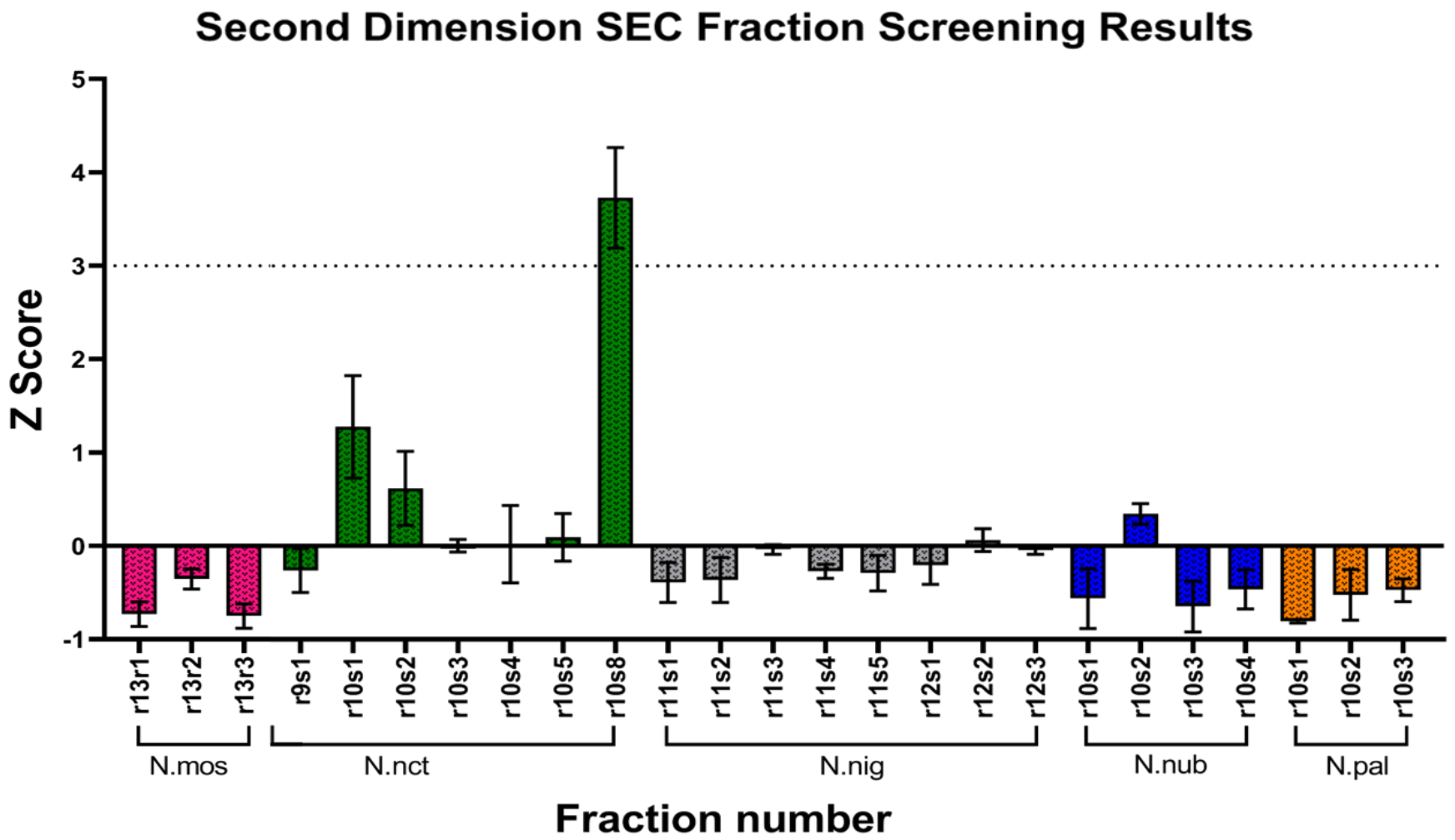



*Figure 4.11. Z scores for each of the 2D SEC fractions (using IE buffers) screened against SW620 cells.*

*Fractions are grouped according to which cobra the venom originated from. Z scores of more than 3 (indicated by the dotted line) are considered significant. Error bars represent SEM based on n=3.*

#### 4.3.3.2 Assessing the Effect of Salt Within HPLC Buffers on Cell Viability

The buffers used for second dimension SEC HPLC contained sodium dihydrogen orthophosphate dihydrate ($NaH_2PO_4$), sodium phosphate dibasic dihydrate ($Na_2HPO_4$) and 0.2 M sodium chloride (NaCl). Although these salts were in reasonably low concentrations during HPLC separation, the resulting fractions were then lyophilised to produce samples up to 20-fold more concentrated than the initial buffers.

An assay was run to discover whether these higher salt concentrations present in some samples could be causing a significant reduction in cell viability by themselves. A range of different concentrations of salt buffer were screened against SW620 cells according to section 2.5.2. The results of the assay can be found in Figure 4.12. Statistics were performed according to section 2.4. A Welch's ANOVA test found that there was a statistically significant difference between the samples ($P<0.001$). A Dunnett's post-hoc test indicated that the 10- and 20-fold concentration of the ion exchange buffers did cause a small but statistically significant reduction in cell viability compared to the negative control ($P<0.05$). It was therefore decided that the screening results described in section 4.3.3 were not suitable for analysis and a different separation approach was taken. Results of normality tests, QQ-plot and Welch's ANOVA with Dunnett's post-hoc tests are shown in Appendix II.

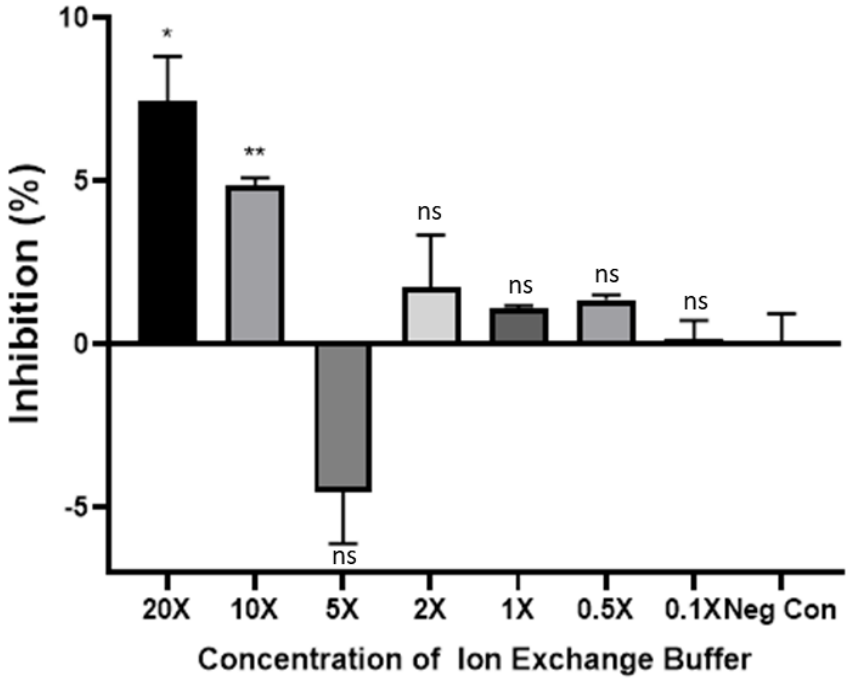



*Figure 4.12. Percentage inhibition of SW620 cells following exposure to the concentrated ion exchange HPLC buffer.*

*Error bars are SEM based on n=3. Significance levels determined by Welch's ANOVA with Dunnett's post-hoc test. (* = p<0.05, ** = p<0.001, ns = not significant (p>0.05)).*

#### 4.3.3.3 Re-screening 2D SEC Fractions

SEC separation was performed on the first dimension RP fractions (selected in section 4.3.2). The fractions were screened, and the results are shown in Figure 4.13. None of the screened fractions retained activity following second dimension separation.

New buffers were investigated excluding phosphate salts to avoid the problem described in section 4.3.3.2 above. The results of the buffer optimisation can be found in Figure 4.14. It was found that a 10-minute SEC run using buffer C provided the best resolution and sharpest peaks, so these parameters were used for HPLC separation.

#### 4.3.3.4 Reversing the dimensions to try to find 'hits'

The HPLC phases were reversed such that SEC was performed first and then 'hit' fractions were further fractionated by RP HPLC. *Naja nubiae* venom was screened against SW620 cells first to see the effect of this phase reversal. Although none of the 1D fractions had a Z score above 3, the two fractions with the highest Z scores were taken forward to second dimension fraction screening. The results are shown in Figure 4.15. The second-dimension fractions were also screened against SW620 cells. The data were tested for statistical significance using Kruskall-Wallis testing with Dunn's post-hoc test, however, none caused statistically significant inhibition of cell growth ($P>0.05$ in all cases). The Z score is shown for the first-dimension fractions; however Z score was inappropriate for the second dimension fractions and so percentage inhibition is shown instead. Results for both dimensions of screening are shown in Figure 4.15.

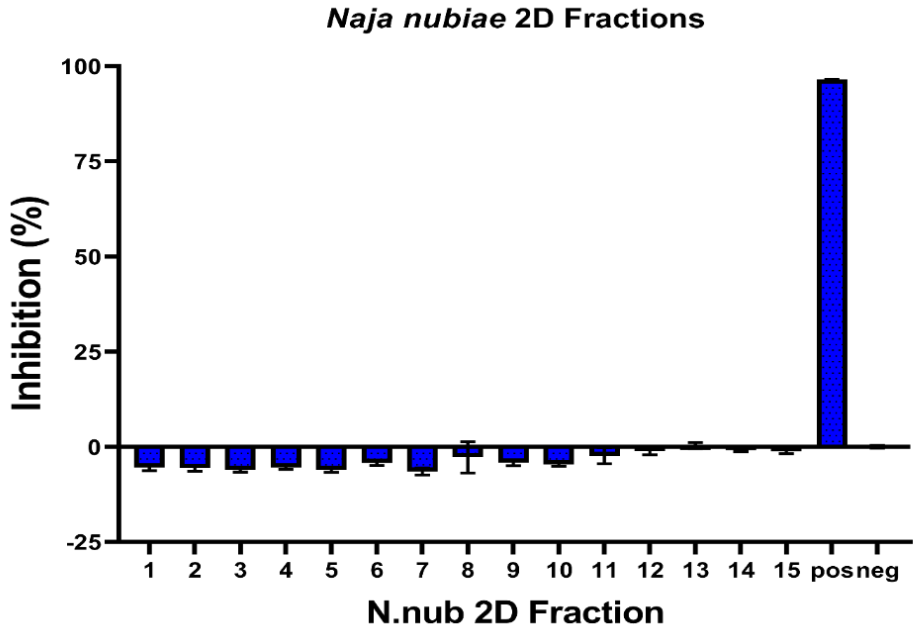



*Figure 4.13. Results of screening second dimension SEC N.nub fractions collected using modified RP buffers against SW620 cells.*

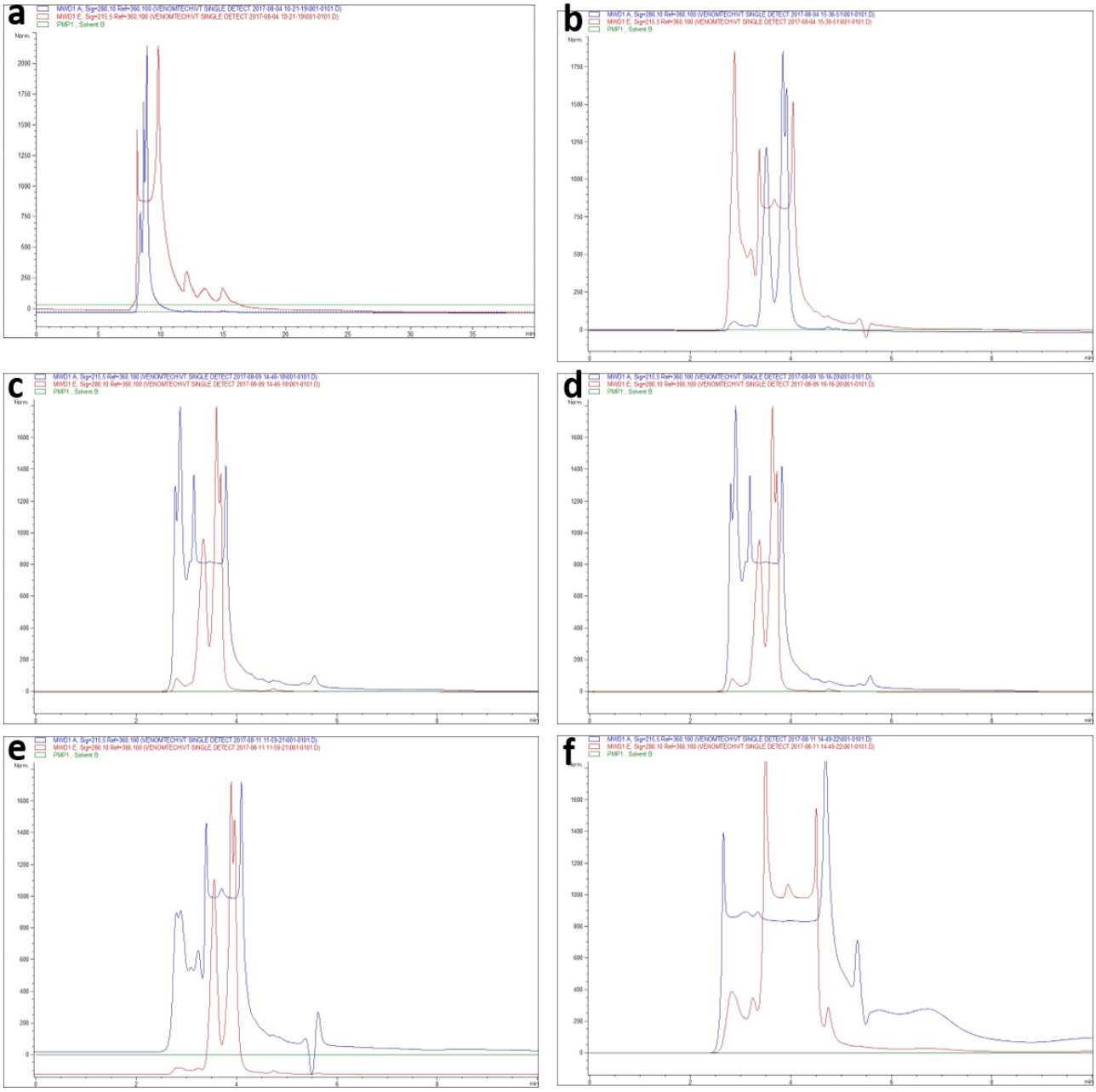



*Figure 4.14. Traces showing optimisation of SEC using reverse phase buffers.*

*0.5 mg protein loaded unless otherwise stated. (a) 62.5% RP Buffer A, 37.5% RP Buffer B, 1ml/min flow rate, 280 nm collection, 40 min run. (b) 100% RP Buffer C (30% ACN, 0.2% TFA), 1 ml/min flow rate, 280 nm collection, 10 min run. (c) 100% RP Buffer D (30% ACN, 0.1% TFA), 1 ml/ min flow rate, 215 nm collection, 10 min run. (d) 100% RP Buffer D, 0.5 ml/min flow rate, 215 nm collection, 10 min run. (e) 100% Buffer C, 1 ml/min flow rate, 215 nm collection, 10 min run. (f) Full run (5 mg protein) 100% Buffer C, 1 ml/min flow rate, 215 nm collection, 10 min run.*

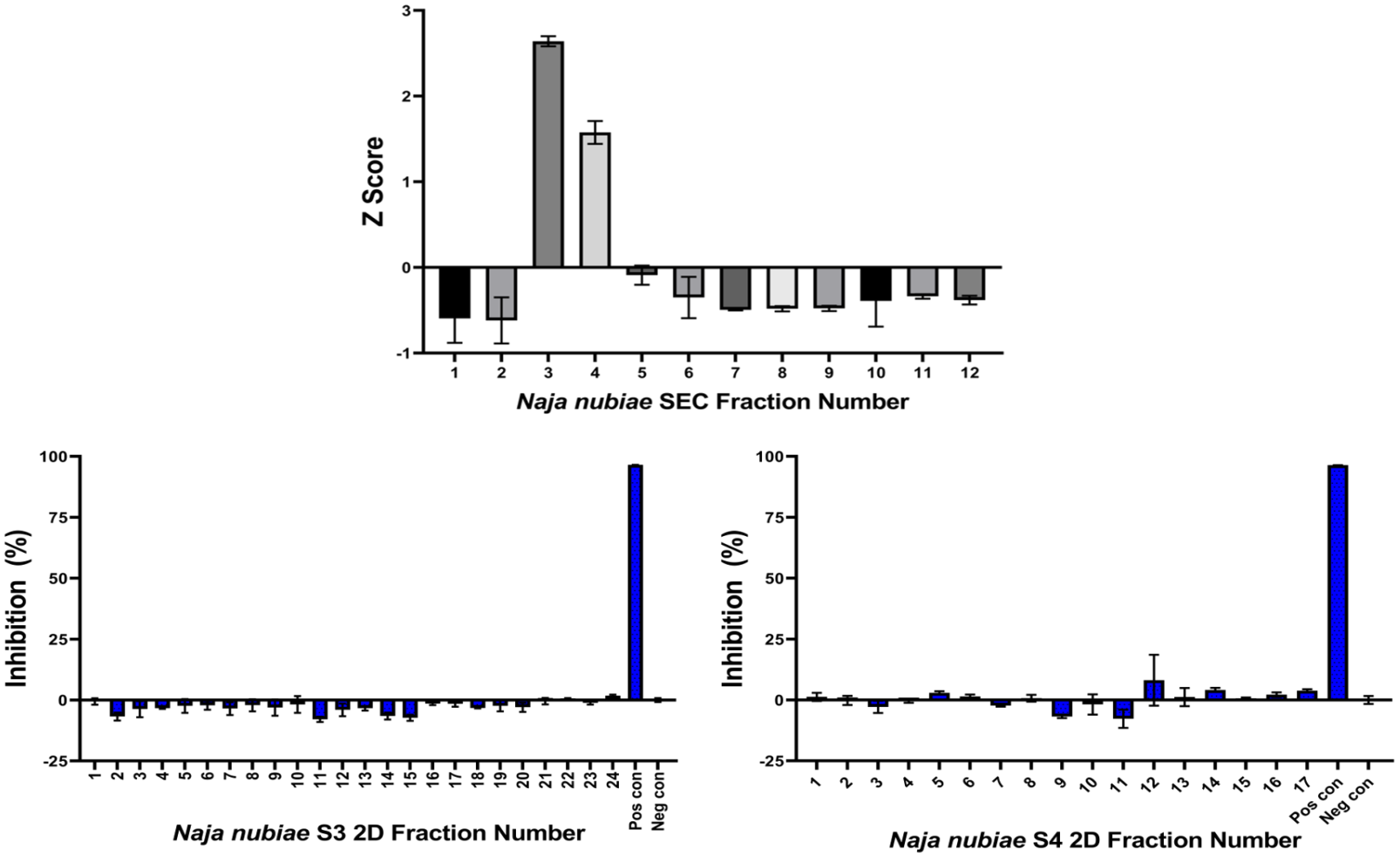



*Figure 4.15. First dimension SEC followed by second dimension RP HPLC.*

*(a) Z score for each of the* Naja nubiae *1D SEC fractions. Percentage inhibition of each of the Naja nubiae SEC (b) SEC fraction 3 fractions and (c) SEC fraction 4 fractions.*

### 4.3.4 NanoDSF

To investigate the effect on venom protein stability when exposed to the buffers used for optimised SEC, the NanoTemper Prometheus (NanoTemper Technologies GmbH, Munich, Germany) was used. *Naja nigricollis* venom was exposed to a range of different diluents (distilled water, 100% DMSO, 100% isopropanol alcohol (IPA) and the optimised ion exchange buffers C and D) and then the average melting temperatures measured, as described in section 2.7. The results of the analysis are shown in Figure 4.16. RP buffer B was also tested, however, due to the volatile nature of this buffer, no melt curve was observed.

Statistical analyses were performed according to section 2.4. Normality tests and QQ plot are shown in Appendix II. Welch's ANOVA with Dunnett's post-hoc tests were used to assess the statistical significance of the different diluents on melt temperature compared to crude, undiluted venom. Details of these tests are provided in Appendix II. These tests indicated there was a significant difference in melt temperature between all the diluents when compared to the crude venom, except venom diluted in distilled water ($p<0.01$ for venom diluted in DMSO and IPA, $p<0.0001$ for venom diluted in RP buffers C and D and $p>0.05$ for distilled water). Figure 4.16c shows that the ion exchange buffers C and D make the largest difference to the melt temperature of the venom, bringing the average melt temperature down from 84.6°C to 51.6°C and 48.8°C for buffers C and D respectively.

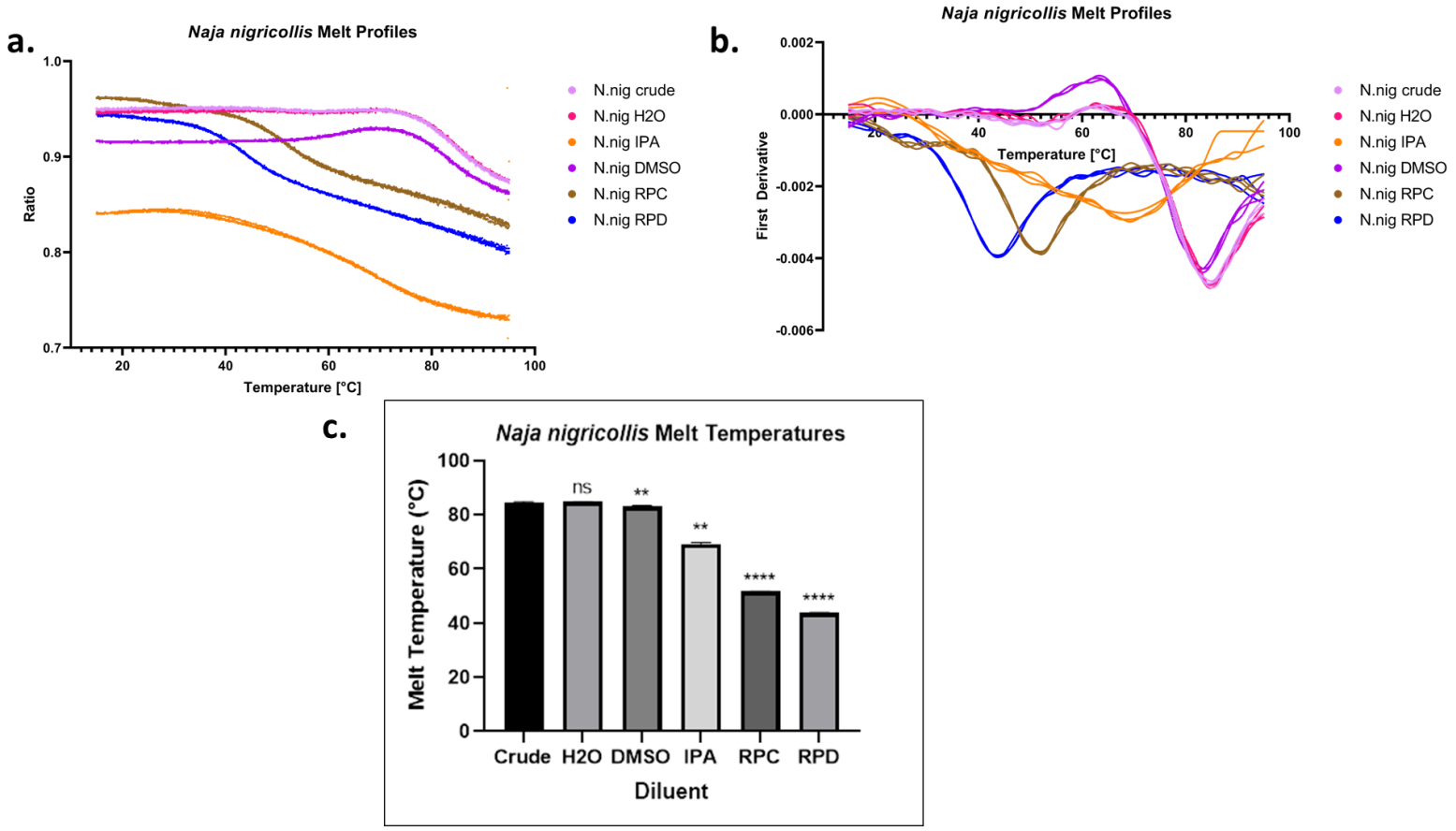



*Figure 4.16. The change in melting temperatures of* Naja nigricollis *venom when exposed to different diluents.*

*(DMSO = 100% dimethyl sulfoxide, IPA = 100% isopropanol alcohol, RPC/D = reverse phase buffer C/D), measured using the NanoTemper Prometheus. (a) Melt profiles (ratio), (b) Melt profiles (First derivative of ratios), (c) Average melting temperatures where error bars are standard deviation based on n=3. Significance displayed as result of Welch's ANOVA with Dunnett's post-hoc test comparing each sample to the crude, undiluted sample, ns= not significant (p>0.05), ** p<0.01, **** p<0.0001.*

### 4.3.5 Performing 2D HPLC on the African Spitting Cobra Venoms

Two-dimensional HPLC was performed to separate the five African spitting cobras - *Naja mossambica* (N.mos), *Naja nigricincta* (N.nct), *Naja nigricollis* (N.nig), *Naja nubiae* (N.nub) and *Naja pallida* (N.pal) - which were identified as active in Chapter 3. Venoms were first separated by ion exchange chromatography according to section 2.5.6, then pooled into 10 samples for each venom and separated further by second dimension RP HPLC according to section 2.5.7. A total of 303 fraction samples were collected from the five species and were lyophilised according to section 2.5.1.3. HPLC traces for each venom may be found in Appendix III.

### 4.3.6 Screening 2D HPLC Fractions Against Bxpc-3 and SW620 Cells to Assess Cell Viability

The cell viability of BxPC-3 and SW620 cells when exposed to 20 µg/ml of each 2D venom fraction were assessed. Images of the plates were taken over the time course and representative plate photos can be seen in Figure 4.17.

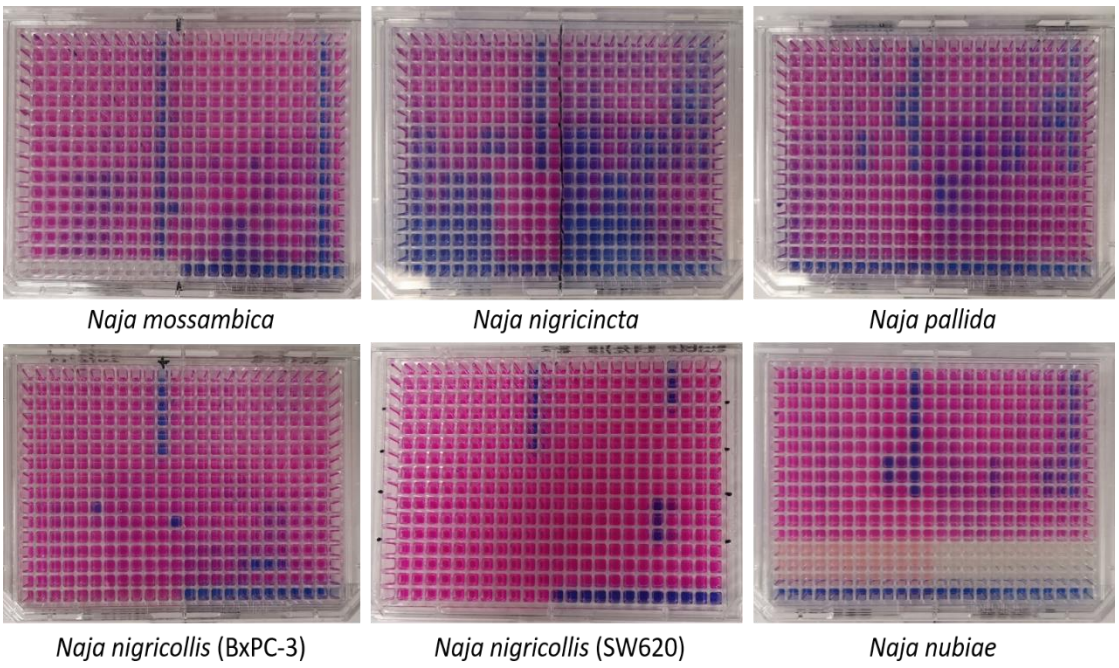



*Figure 4.17. Image of 384-well plate screens.*

*Each plate contains SW620 cells in columns 1-12 and BxPC-3 cells in columns 13-24 (except Naja nigricollis as there were too many fractions to fit all replicates on one plate so BxPC-3 and SW620 were screened on separate plates.) Pink colour (presence of mostly resorufin) indicates high cell viability whilst purple (mixture of both resazurin and resorufin) and blue colours (presence of mostly resazurin) indicate reduced or absent cell viability. Plate photos were taken after 120 minutes of resazurin exposure.*

Fluorescence intensity was measured for each sample across assay plates. Each sample was given a unique code to identify which species that fraction came from and which chronological ion exchange (i) and reverse phase (r) sample number it was, for example N.pal_i17r2 is ion exchange sample 17, reverse phase sample 2 from *Naja pallida*. Outlying data in the controls were identified by plotting boxplots for each negative, positive and resazurin only control for each screen. These boxplots can be found in Appendix II. Any data points lying outside of the "box and whisker" plot were excluded from the raw data sets during data processing.

The results from the 2D venom fraction screens were processed and the results are displayed graphically below from the following species (*Naja mossambica* in Figure 4.18, *Naja nigricincta* in Figure 4.19, *Naja nigricollis* in Figure 4.20 (fractions 1-50) and Figure 4.21 (fractions 51-95), *Naja nubiae* in Figure 4.22 and *Naja pallida* in Figure 4.23). The venom fraction samples were triaged according to their potency, selectivity and total protein yield. Only fractions with a percentage inhibition in the BxPC-3 cell line higher than 25% *and* higher than in the SW620 cell line were taken forwards for statistical analysis. Fractions causing over 25% inhibition are likely to be having a substantial biological effect and not just deviating from the baseline.

These selective fractions were analysed using the Kruskal-Wallis and Dunn's tests according to section 2.4. Results of these tests are shown in Appendix II. Statistical significance is displayed on the corresponding graph and is summarised in Table 4.3. *Naja mossambica* venom contained two fractions showing statistical significance in both cell lines - i26r1 and i26r2. *Naja nigricincta* venom contained three significant fractions in both cell lines – i18r2, i19r1 and i20r2. Although *Naja nigricollis* venom had the highest number of 2D fractions, none were selective for BxPC-3 cells over SW620 cells. *Naja nubiae* venom had just one fraction that was significant in both cell lines – i17r2. *Naja pallida* had five – i18r2, i18r3 and i20r1 were significant in both cell lines and i15r4 and i17r2 were significant in the BxPC-3 cell line only.

There were a total of 11 second-dimension fractions eliciting a statistically significant effect out of a total of 303 fractions (3.63% hit rate). Of these hits, only those with a total yield of at least 80 µg were taken forwards. This was due to the requirement of 60 µg protein to perform each dose-response curve with enough sample remaining to send off for mass spectrometry analysis (Chapter 5) and to take forward for qPCR analysis (Chapter 6). A total of 6 venom fractions across four cobra species met the selection criteria and were investigated further. The selected fractions are highlighted in Table 4.4. These were N.mos_i26r2, N.nct_i18r2, N.nub_i17r2, N.pal_i15r4, N.pal_i17r2 and N.pal_i18r2.

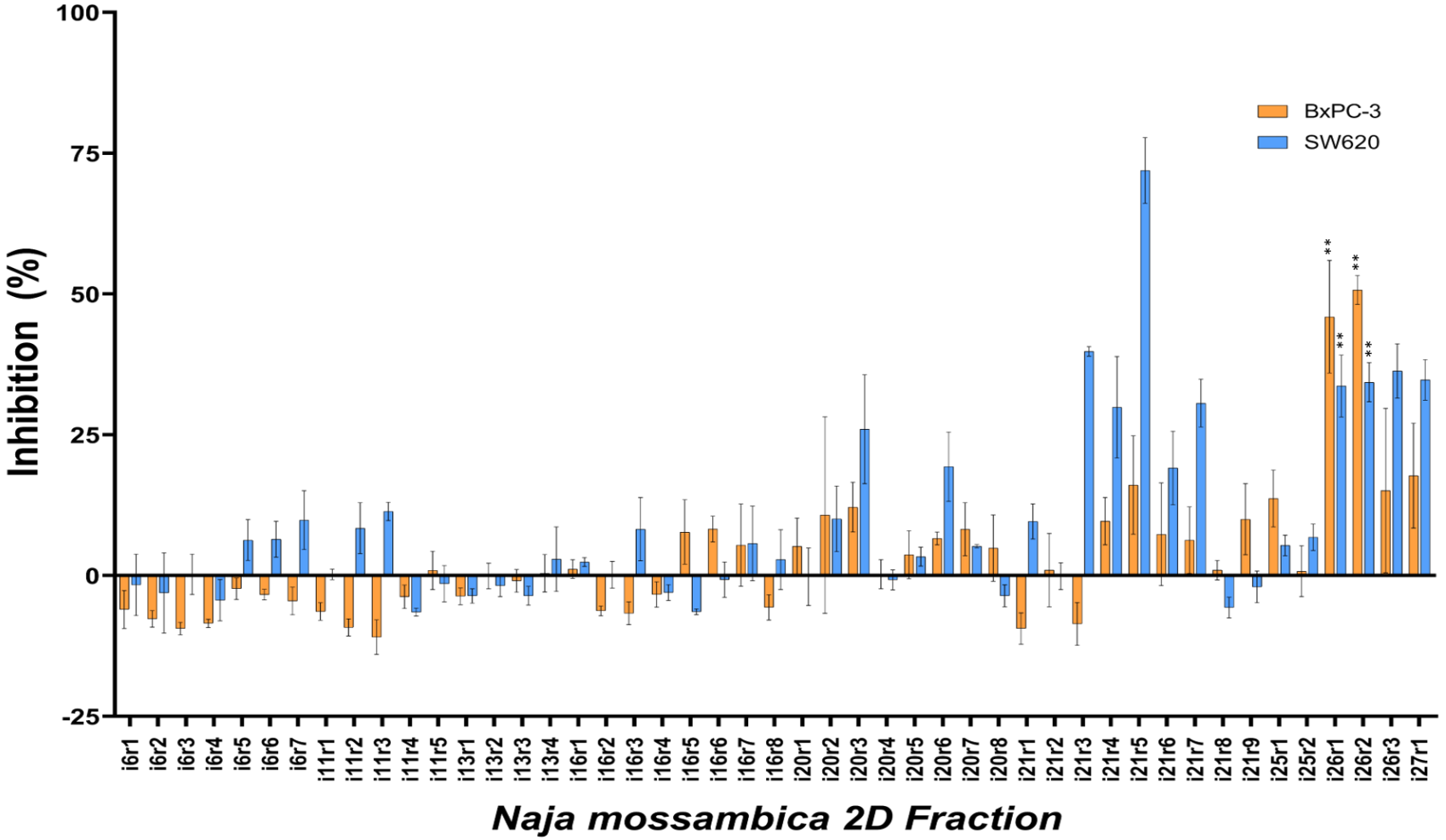



*Figure 4.18. Naja mossambica 2D fraction screening results.*

*Percentage inhibition is shown for BxPC-3 cells (orange) and SW620 cells (blue). Fractions that showed a statistically significant difference in percentage inhibition compared to the untreated control cells according to Kruskal-Wallis testing followed by Dunn's post-hoc test are indicated on the graph. ** = p <0.01. Error bars are based on SEM where n=3.*

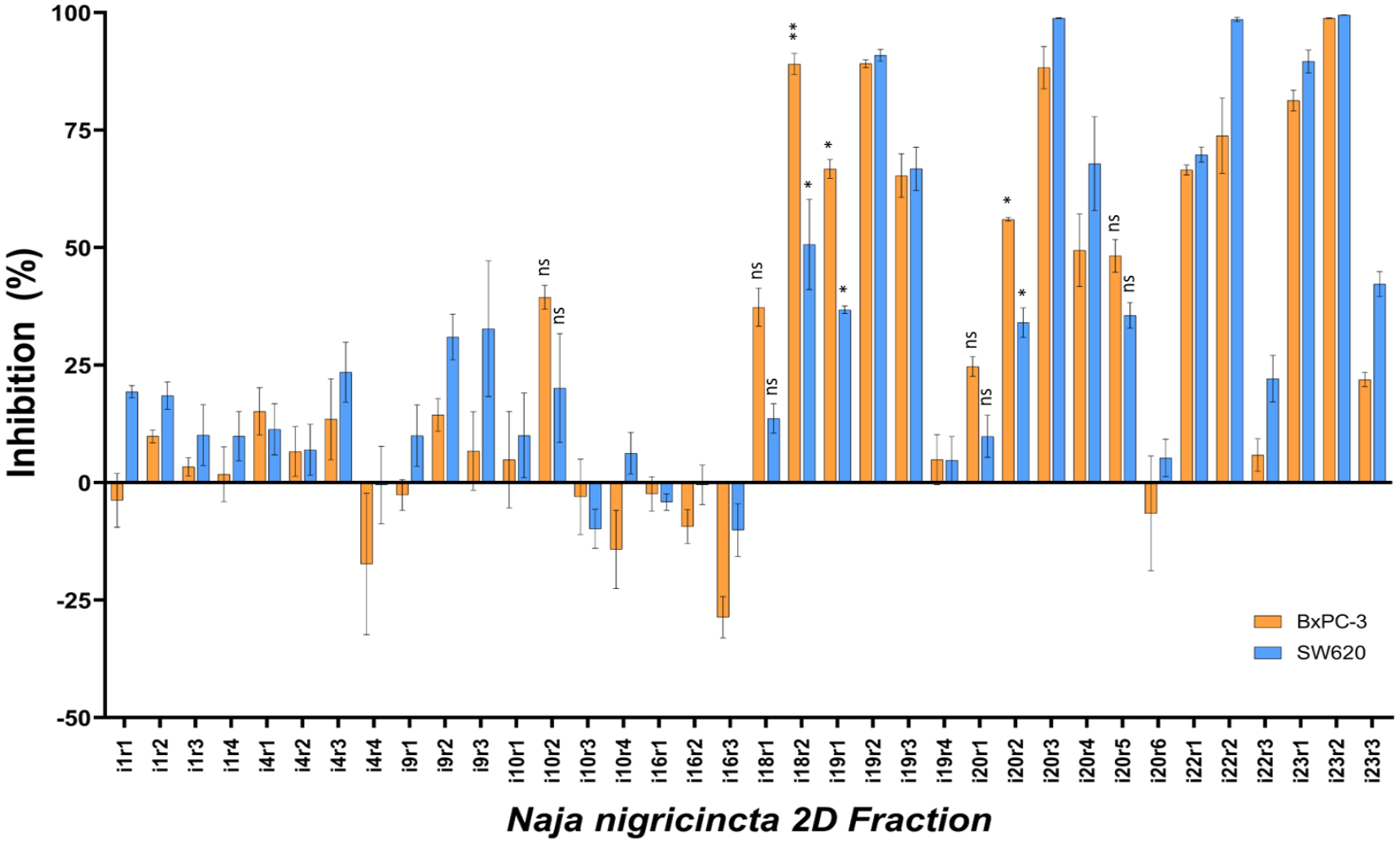



*Figure 4.19. Naja nigricincta 2D fraction screening results.*

*Percentage inhibition is shown for BxPC-3 cells (orange) and SW620 cells (blue). Fractions that showed a statistically significant difference in percentage inhibition compared to the untreated control cells according to Kruskal-Wallis testing followed by Dunn's post-hoc test are indicated on the graph. .ns = not significant (p>0.05), * = p<0.05, ** = p <0.01. Error bars are based on SEM where n=3.*

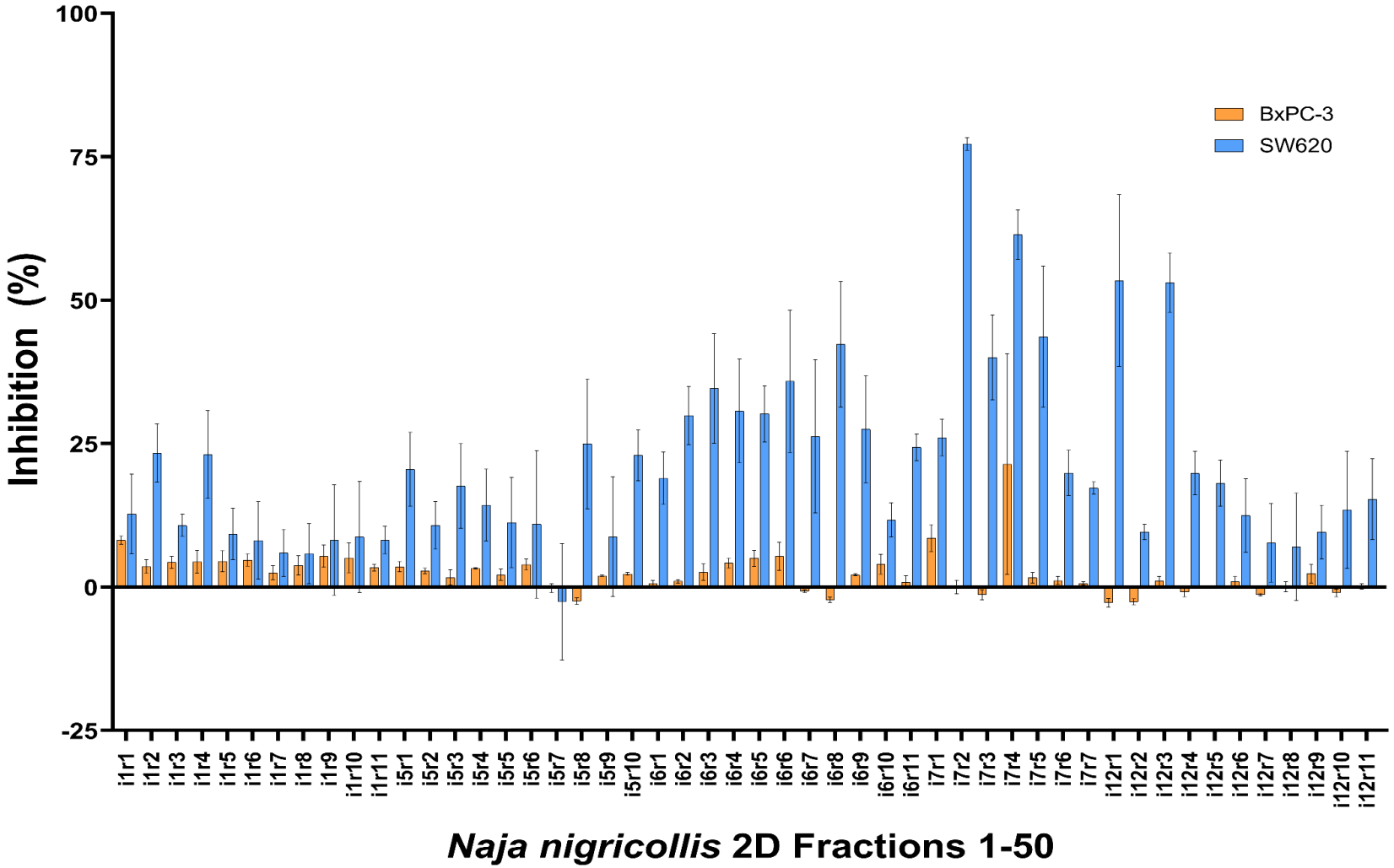



*Figure 4.20. Naja nigricollis 2D fractions 1-50 screening results.*

*Percentage inhibition is shown for BxPC-3 cells (orange) and SW620 cells (blue). Fractions that showed a statistically significant difference in percentage inhibition compared to the untreated control cells according to Kruskal-Wallis testing followed by Dunn's post-hoc test are indicated on the graph. Error bars are based on SEM where n=3.*

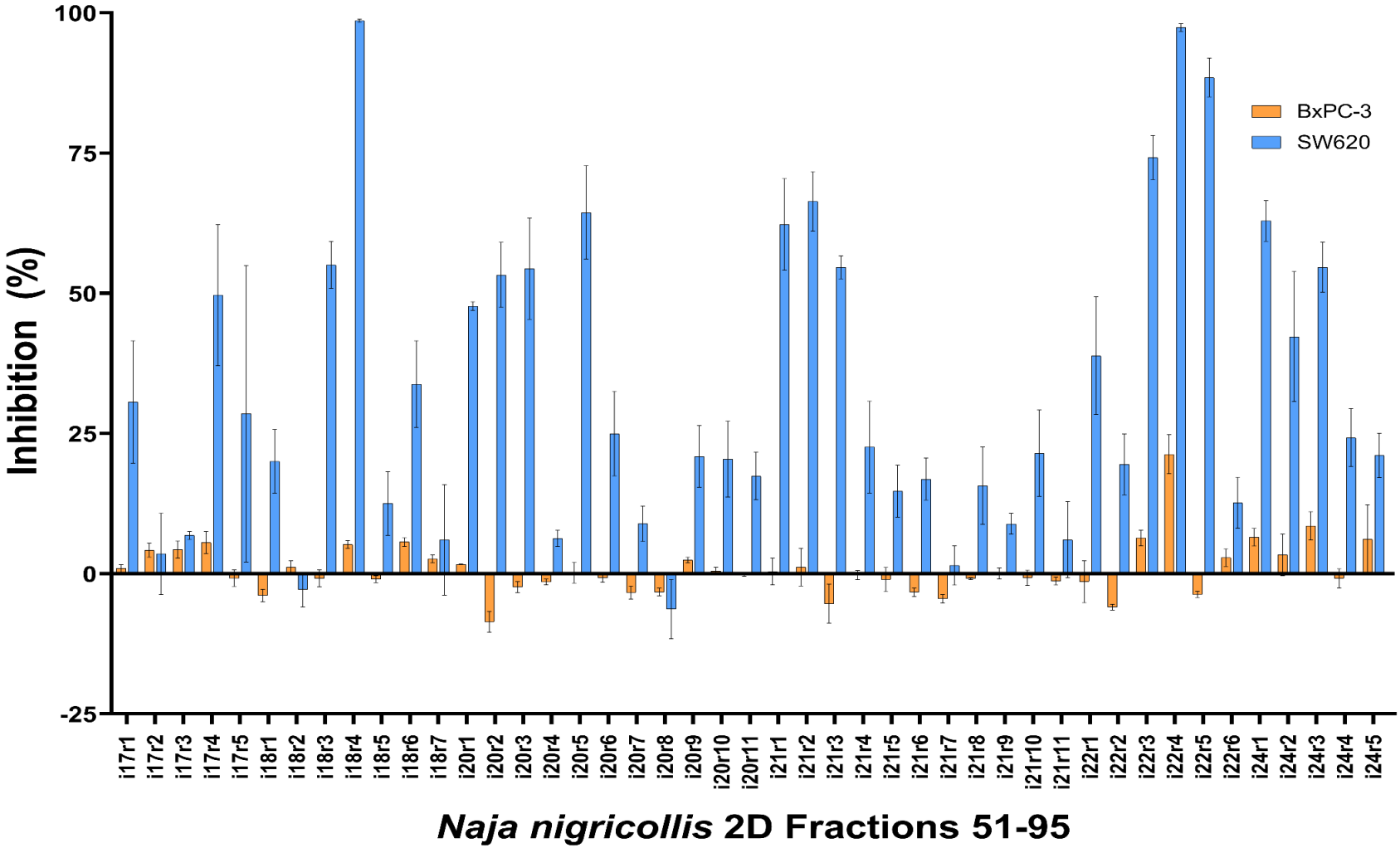



*Figure 4.21. Naja nigricollis 2D fractions 51-95 screening results.*

*Percentage inhibition is shown for BxPC-3 cells (orange) and SW620 cells (blue). Fractions that showed a statistically significant difference in percentage inhibition compared to the untreated control cells according to Kruskal-Wallis testing followed by Dunn's post-hoc test are indicated on the graph. Error bars are based on SEM where n=3.*

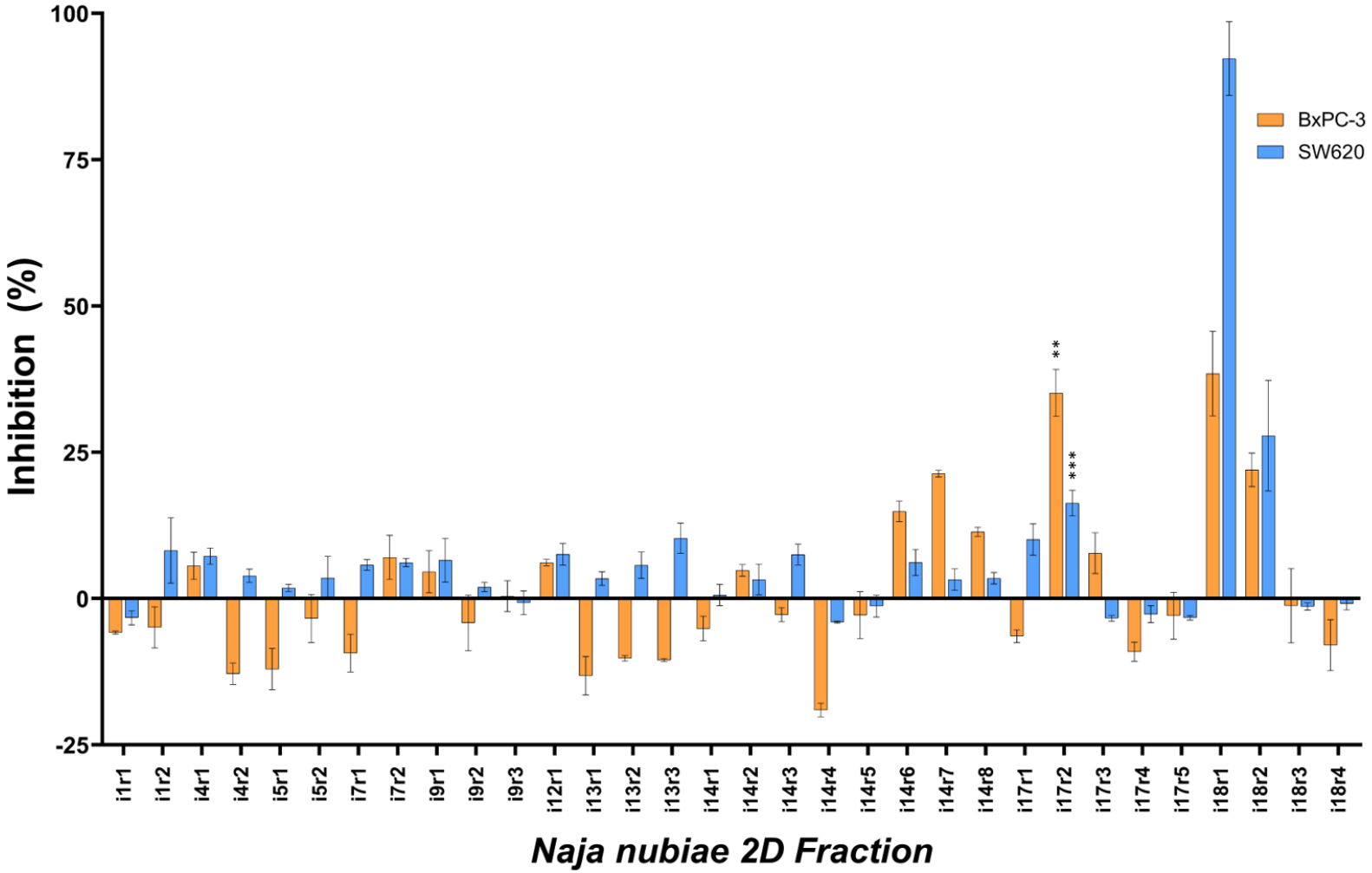



*Figure 4.22.Naja nubiae 2D fraction screening results.*

*Percentage inhibition is shown for BxPC-3 cells (orange) and SW620 cells (blue). Fractions that showed a statistically significant difference in percentage inhibition compared to the untreated control cells according to Kruskal-Wallis testing followed by Dunn's post-hoc test are indicated on the graph. ** = p <0.01. Error bars are based on SEM where n=3.*

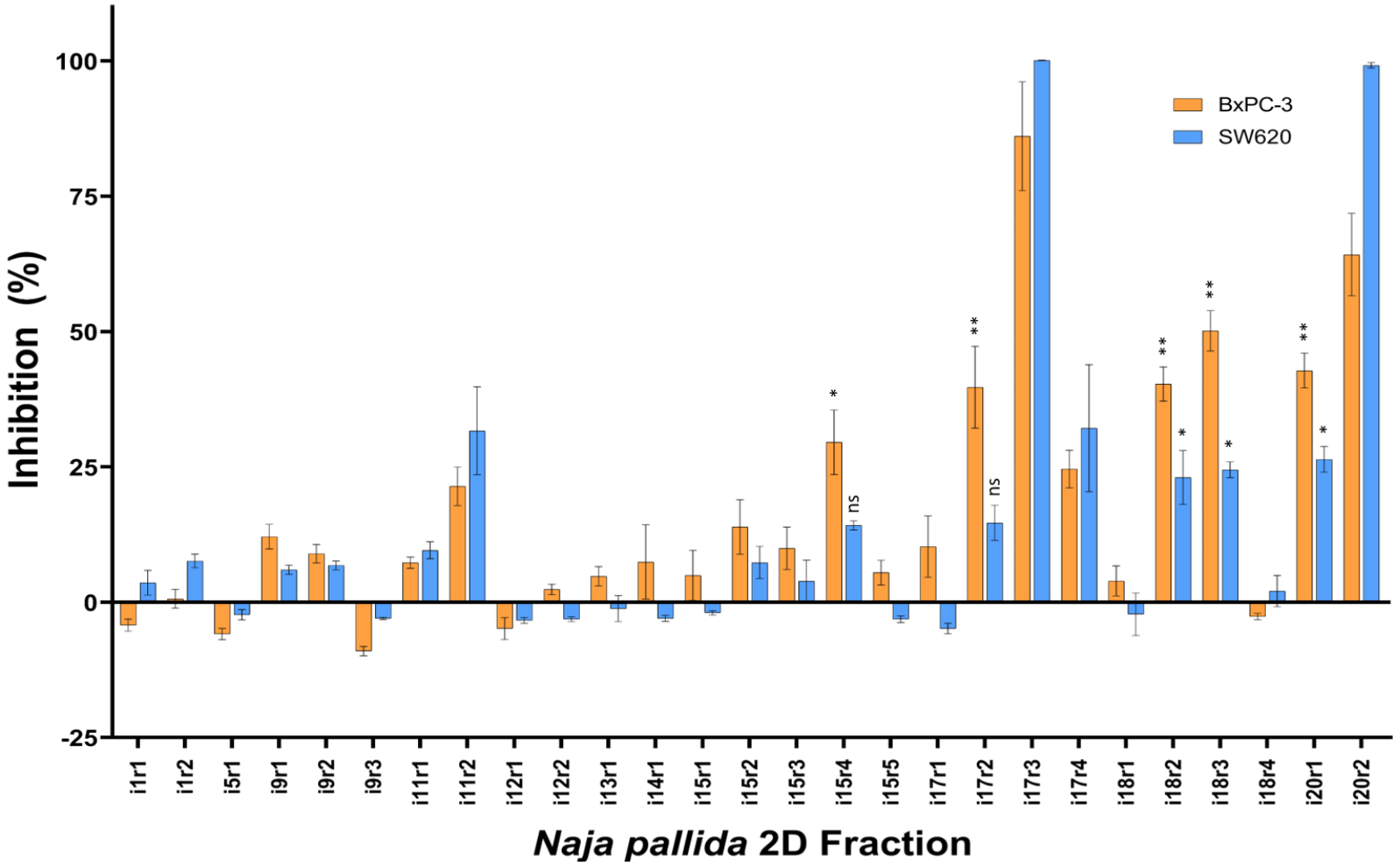



*Figure 4.23. Naja pallida 2D fraction screening results.*

*Percentage inhibition is shown for BxPC-3 cells (orange) and SW620 cells (blue). Fractions that showed a statistically significant difference in percentage inhibition compared to the untreated control cells according to Kruskal-Wallis testing followed by Dunn's post-hoc test are indicated on the graph. ns = not significant (p>0.05), * = p<0.05, ** = p <0.01. Error bars are based on SEM where n=3.*

*Table 4.3. Summary of the 2D cobra fractions tested for statistical significance.*

*Fractions were tested using Kruskal-Wallis testing followed by Dunn's multiple comparisons post-hoc test on selected fractions compared to the corresponding negative control data. * = p<0.05, ** = p<0.01, *** = p<0.001, ns = not significant (p>0.05) based on n=3.*

| Fraction | Species | IE Fraction | RP Fraction | BxPC-3 Significance | SW620 Significance |
|---|---|---|---|---|---|
| N.mos_ i26r1 | *Naja mossambica* | 26 | 1 | ns | * |
| N.mos_ i26r2 | *Naja mossambica* | 26 | 2 | * | * |
| N.nct_ i10r2 | *Naja nigricincta* | 10 | 2 | ns | ns |
| N.nct_ i18r1 | *Naja nigricincta* | 18 | 1 | ns | ns |
| N.nct_ i18r2 | *Naja nigricincta* | 18 | 2 | ** | * |
| N.nct_ i19r1 | *Naja nigricincta* | 19 | 1 | * | * |
| N.nct_ i20r1 | *Naja nigricincta* | 20 | 1 | ns | ns |
| N.nct_ i20r2 | *Naja nigricincta* | 20 | 2 | * | * |
| N.nct_ i20r5 | *Naja nigricincta* | 20 | 5 | ns | * |
| N.nub_ i17r2 | *Naja nubiae* | 17 | 2 | ** | *** |
| N.pal_ i15r4 | *Naja pallida* | 15 | 4 | * | ns |
| N.pal_ i17r2 | *Naja pallida* | 17 | 2 | ** | ns |
| N.pal_ i18r2 | *Naja pallida* | 18 | 2 | ** | * |
| N.pal_ i18r3 | *Naja pallida* | 18 | 3 | ** | * |
| N.pal_ i20r1 | *Naja pallida* | 20 | 1 | ** | * |

*Table 4.4. Summary of fractions that showed statistical significance against BxPC-3 cells.*

*The peak times for both dimensions of HPLC are displayed, along with percentage inhibition for both cell lines and the total yield of protein for each 2D fraction. Fractions with less than 80 µg total yield are shown in red text. The remaining fractions are highlighted according to the species and were taken forward for dose response curves.*

| Fraction | Species | IE Fraction | RP Fraction | BxPC-3 inhibition (%) | SW620 inhibition (%) | Total yield (µg) |
|---|---|---|---|---|---|---|
| N.mos_i26r2 | *Naja mossambica* | 26 | 2 | 50.71 | 34.32 | 101.03 |
| N.nct_18r2 | *Naja nigricincta* | 18 | 2 | 89.06 | 60.29 | 81.93 |
| N.nct_i19r1 | *Naja nigricincta* | 19 | 1 | 66.75 | 36.77 | 7.6 |
| N.nct_i20r2 | *Naja nigricincta* | 20 | 2 | 56.04 | 34.05 | 15.85 |
| N.nub_i17r2 | *Naja nubiae* | 17 | 2 | 35.14 | 16.31 | 86.03 |
| N.pal_i15r4 | *Naja pallida* | 15 | 4 | 29.54 | 14.19 | 244.7 |
| N.pal_i17r2 | *Naja pallida* | 17 | 2 | 39.71 | 14.66 | 91.63 |
| N.pal_i18r2 | *Naja pallida* | 18 | 2 | 40.31 | 23.03 | 86.4 |
| N.pal_i18r3 | *Naja pallida* | 18 | 3 | 50.14 | 24.47 | 17.73 |
| N.pal_i20r1 | *Naja pallida* | 20 | 1 | 42.79 | 26.4 | 35.95 |

### 4.3.7 HPLC Details for the Active 2D Venom Fractions

The HPLC traces for both dimensions of separation were compared to see the differences in retention times. The first-dimension ion exchange HPLC traces are overlaid in Figure 4.24. The second-dimension reverse phase HPLC traces are overlaid in Figure 4.25. The start times, end times and approximate peak time of each of the dimensions of the 'hit' fractions are displayed in Table 4.5(a).

All the first dimension IE fractions eluted between 44-65 minutes. This equates to 445-740 mM NaCl concentration in the buffers. In the second dimension, all fraction eluted between 11-14 minutes, or 29.9-38.2 % ACN. Details of the concentrations of the buffers and their significant contents at the time of fractions elution are displayed in Table 4.5 (b).

*Table 4.5. HPLC details for active cobra fractions.*

*(a) Comparison of the peak start time, end time and estimated peak time of the first and second dimensions for the active HPLC fractions. All times are in minutes from the start of the HPLC run. (b) The peak times of each dimension of HPLC converted into the percentage of relevant buffer B and converted into mM NaCl for 1D IE buffer and to % ACN in 2D RP HPLC.*

**a.**

| Code | Cobra species | 1st Dimension | | | | 2nd Dimension | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | IE fraction | Start (min) | End (min) | Peak (min) | RP fraction | Start (min) | End (min) | Peak (min) |
| N.mos_i26r2 | *Naja mossambica* | 26 | 64.50 | 68.50 | 68.91 | 2 | 13.51 | 15.50 | 13.98 |
| N.nct_i18r2 | *Naja nigricincta* | 18 | 47.98 | 51.98 | 50.97 | 2 | 11.60 | 13.59 | 11.76 |
| N.nub_i17r2 | *Naja nubiae* | 17 | 58.17 | 61.23 | 60.24 | 2 | 10.73 | 11.76 | 11.17 |
| N.pal_i15r4 | *Naja pallida* | 15 | 44.83 | 46.07 | 45.63 | 4 | 15.18 | 17.17 | 16.33 |
| N.pal_i17r2 | *Naja pallida* | 17 | 58.12 | 62.12 | 60.32 | 2 | 10.65 | 11.79 | 11.19 |
| N.pal_i18r2 | *Naja pallida* | 18 | 63.17 | 67.15 | 69.06 | 2 | 12.47 | 13.45 | 12.73 |

**b.**

| Code | Cobra species | 1st Dimension | | | | 2nd Dimension | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | IE fraction | Peak (min) | IE B (%) | NaCl (mM) | RP fraction | Peak (min) | RP B (%) | ACN (%) |
| N.mos_i26r2 | *Naja mossambica* | 26 | 68.91 | 73.64 | 736.38 | 2 | 13.98 | 42.96 | 34.37 |
| N.nct_i18r2 | *Naja nigricincta* | 18 | 50.97 | 51.21 | 512.13 | 2 | 11.76 | 38.52 | 30.82 |
| N.nub_i17r2 | *Naja nubiae* | 17 | 60.24 | 62.80 | 628.00 | 2 | 11.17 | 37.34 | 29.87 |
| N.pal_i15r4 | *Naja pallida* | 15 | 45.63 | 44.54 | 445.38 | 4 | 16.33 | 47.66 | 38.13 |
| N.pal_i17r2 | *Naja pallida* | 17 | 60.32 | 62.90 | 629.00 | 2 | 11.19 | 37.38 | 29.90 |
| N.pal_i18r2 | *Naja pallida* | 18 | 69.06 | 73.83 | 738.25 | 2 | 12.73 | 40.46 | 32.37 |

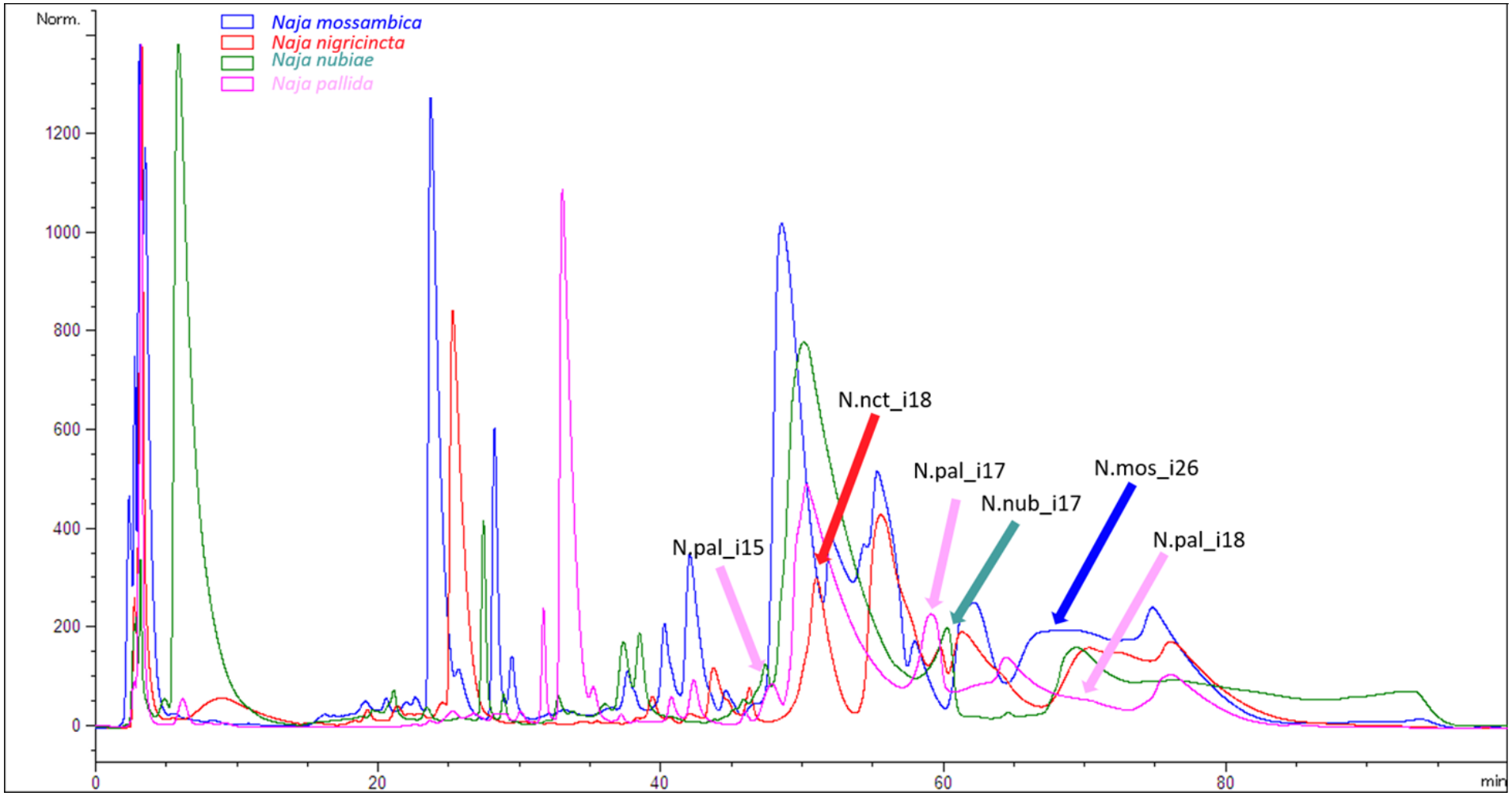



*Figure 4.24. Overlay of first-dimension preparative ion exchange traces from the four active cobra venoms.*

*The peak of each of the first-dimension fractions identified as a hit in second-dimension fraction screening is indicated with an arrow.*

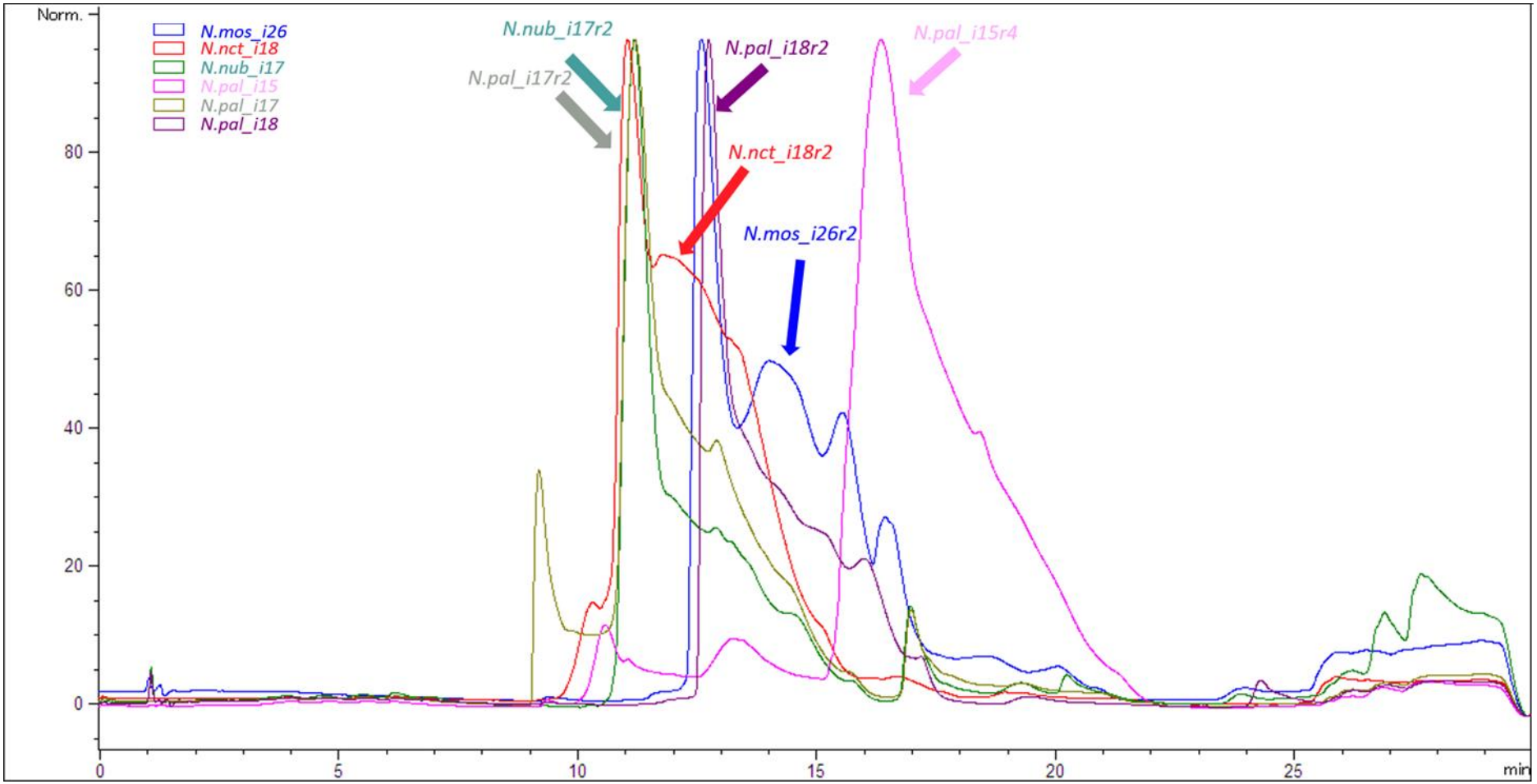



*Figure 4.25. Overlay of second-dimension preparative reverse phase traces from active cobra venoms.*

*The peak of each of the second-dimension fractions identified as a hit is indicated with an arrow.*

### 4.3.8 Dose Response Curves

Ten-point dose response curves were performed for each of the six 'hit' venom fractions. The starting dose was 160 µg/ml and were serially diluted by a factor of 2 each time. The results are displayed in Figure 4.26. The $IC_{50}$, $LogIC_{50}$ and Hillslope values for each curve are displayed in Table 4.6. The standard error of the mean for each sample at each concentration is shown on the graphs in Figure 4.26.

*Table 4.6. $IC_{50}$ values for the second-dimension venoms tested in dose response against BxPC-3 and SW620 cells.*

*$IC_{50}$ shows the concentration (µg/ml) at which 50% of the cell population were inhibited by the sample. Log $IC_{50}$ and Hillslope are also shown. Where it was not possible to calculate an accurate $LD_{50}$ value, - is displayed. All values based on n=3 technical repeats from n=1 biological repeat. Standard deviations for each point in the dose response curve are shown on the graphs in Figure 4.26.*

| | BxPC-3 | | | SW620 | | |
|---|---|---|---|---|---|---|
| | **$IC_{50}$ (µg/ml)** | **log$IC_{50}$** | **Hillslope** | **$IC_{50}$ (µg/ml)** | **log$IC_{50}$** | **Hillslope** |
| **N.mos_i26r2** | 29.23 | 1.47 | 1.048 | 55.78 | 1.75 | 4.965 |
| **N.nct_i18r2** | 23.64 | 1.37 | 5.92 | 17.85 | 1.25 | 4.74 |
| **N.nub_i17r2** | 71.62 | 1.86 | 1.25 | - | - | 1.46 |
| **N.pal_i15r4** | 35.45 | 1.55 | 1.74 | 28.33 | 1.45 | 2.94 |
| **N.pal_i17r2** | 46.10 | 1.79 | 1.23 | - | - | 1.44 |
| **N.pal_i18r2** | - | - | 1.32 | - | - | 2.5 |

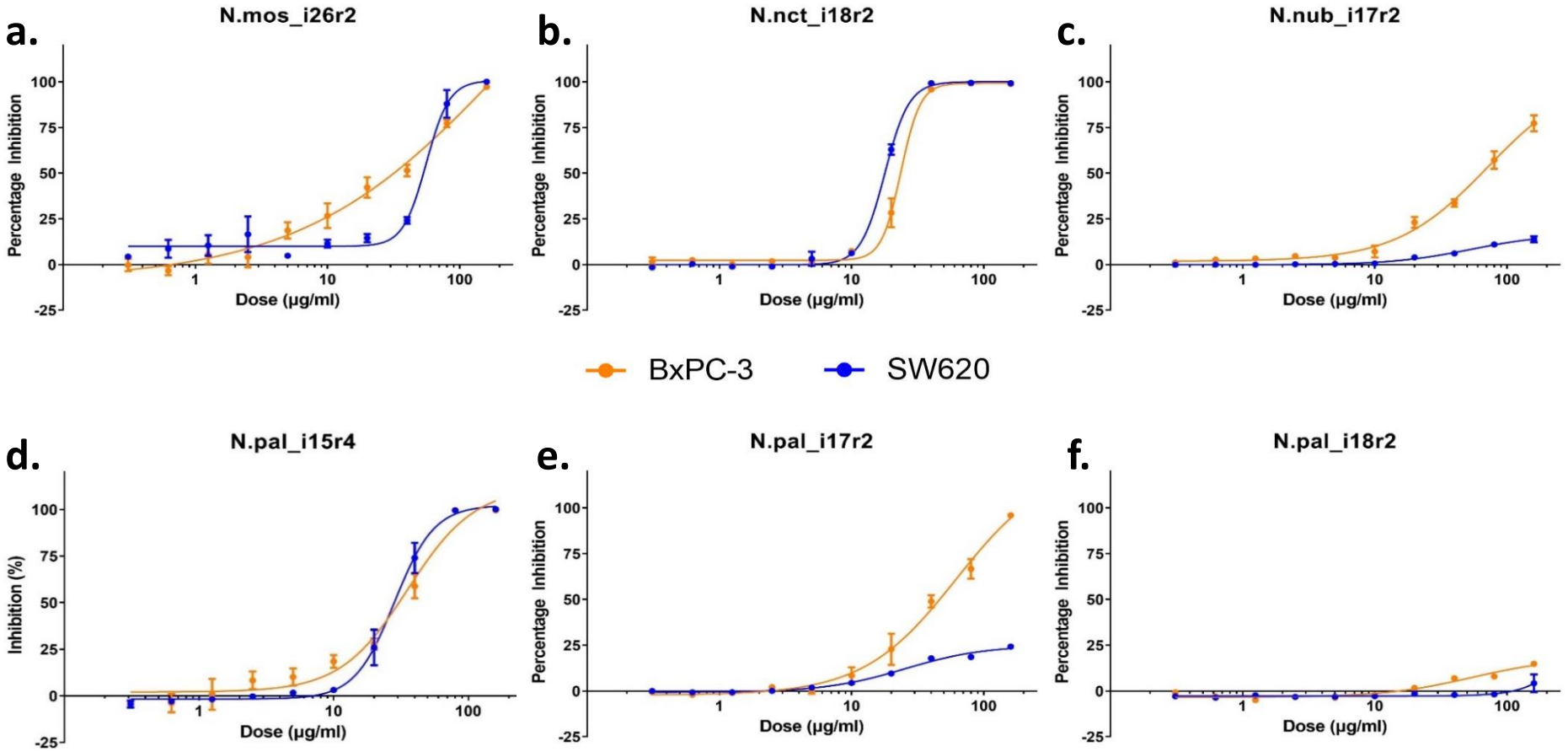



*Figure 4.26. Dose response curves for the selected cobra fraction samples.*

*BxPC-3 cell dose responses are shown in orange and SW620 cell dose responses are shown in blue. Error bars show the SEM based on n=3 technical repeats of n=1 biological repeat.*

The dose response curves for three of the cobra venom fractions (N.mos_i26r2, N.nct_i18r2 and N.pal_i15r4) reach the maximum efficacy possible for the resazurin assay of ~100% inhibition for both of the cell lines. The N.nct_i18r2 and N.pal_i15r4 fractions have dose response curves which show both cell lines reaching the same potency equally. The N.pal_i15r4 curves for each cell line are essentially superimposed on one another whilst the N.nct_i18r2 curves appear more distinct with the BxPC-3 curve shifted to the right. However, it is only one point in the centre of the logarithmic portion of the dose response curve which is different for this sample. The points flanking the logarithmic sigmoid portion of the curve are almost identical for both cell lines for this sample.

#### *4.3.8.1 Comparison of properties of N.nub_i17r2 and N.pal_i17r2 fraction samples*

The N.nub_i17r2 and N.pal_i17r2 samples show a large discrepancy between BxPC-3 and SW620 cells in terms of both potency and efficacy. In both cases, the cobra venom fractions displayed a difference in efficacy with the inhibition of BxPC-3 cells still increasing exponentially towards a high percentage inhibition whilst the SW620 cell response appears to be plateauing at approximately 25% of the maximal response. It is possible that the concentration of sample was not sufficient to reach the exponential phase of the dose response and this sample may have reached the same efficacy over time. The response of the SW620 cells also shows less potency with a much shallower dose response curve compared to the BxPC-3 response curve. Unfortunately, obtaining more sample to repeat the assay starting at a higher dose was not possible with the amount of protein for each sample obtained following 2D HPLC.

N.nub_i17r2 and N.pal_i17r2 were directly compared for HPLC retention time, percentage inhibition of both test cell lines and dose response pharmacology. N.nub_i17r2 and N.pal_i17r2 had first dimension retention times of 60.24 and 60.32 minutes respectively with second dimension retention times of 11.17 and 11.19 minutes, respectively.

In addition to the similar analytical characteristics of the two venom fractions, N.nub_i17r2 and N.pal_i17r2 caused similar percentage inhibition when resazurin was used to measure cytotoxicity. The percentage inhibition of each fraction to the cancer cells tested gave 35.14% and 39.71% inhibitions respectively against BxPC-3 cells and 16.31% and 14.66% inhibitions respectively against SW620 cells. The two venom fractions also gave similar pharmacological profiles in their dose response curves with the BxPC-3 treated cells heading towards 100% inhibition in both cases whilst the SW620 cells appeared to be plateauing at around just 25% inhibition.

## 4.4 Discussion

### 4.4.1 Data Normalisation Methods

Three different data analysis methods were compared to decide which would be the most appropriate following screening. Percent of control and normalised percent inhibition (NPI) use the untreated control wells to calculate an inhibition score relative to the control wells (Malo *et al.*, 2006). The Z-score uses the mean of all the tested samples as a baseline to compare each sample and does not rely on any controls for data normalisation (Brideau *et al.*, 2003).

In most cases, the three methods produce similar looking graphs with the same fractions appearing to be 'active' in each case. However, care should be taken if there are too few (or too many) active fractions in the screen as the baseline may become such that very small changes in cell viability result in very large changes in Z score. This could lead to false positive results where fractions that are the *most* active (but still essentially inactive), could be taken forwards (Brideau *et al.*, 2003).

Although there were no faults with the percent of control and normalised percentage inhibition (NPI) methods, it was decided that the Z score would be the best method to use for data analysis unless a clear reason not to use it occurs. This is due to the Z score being the only statistical scoring test of the three, incorporating the data variability into the output. This method is generally accepted for use in HTS campaigns as an easy method for data analysis and is widely used (Shun *et al.*, 2011). If for a particular data set the Z score was not appropriate, percentage inhibition was used.

Once the Z scores had been computed, two further tests were performed to determine which fractions showed activity. This was the Kruskal-Wallis test followed by Dunn's multiple cross comparisons test. Non-parametric tests were selected due to the absence of equal variance between the samples. The Kruskal-Wallis test was used to identify whether there were any statistically significant results in the data set and Dunn's post-hoc test was used to identify *which* fractions were demonstrating significance.

In addition, ROUT outlier tests were performed to measure activity in the screen. Active compounds in a screen are essentially outliers (Brideau *et al.*, 2003) and so this principle was utilised in order to verify the statistical based tests previously discussed. Unlike the Grubb's outlier test, which only has the power to identify one outlier, the ROUT outlier test may detect any number of potential outliers in a data set (Motulsky and Brown, 2006).

### 4.4.2 Finding Active Fractions from First Dimension RP and Second Dimension SEC HPLC

For each of the fractionated cobra venoms brought forward to first dimension RP fractionation from Chapter 3, at least one fraction was found to inhibit SW620 cells. Size exclusion chromatography (SEC) was performed on the active compounds from the first-dimension screening. The second-dimension fractions were then screened against SW620 cells, however, only one 'hit' was found. This could be due to the venom components giving synergistic effects when combined, the effects of which are reduced or even lost when the components are separated (Gao *et al.*, 2018). It may also simply be that the fractions were screened at too low a concentration to be effective (Macarron *et al.*, 2011). The buffers used for the second dimension HPLC contained various salts which ended up being highly concentrated when lyophilised. The one hit fraction from the screen was one of the fractions that was the most highly concentrated. The potentially toxic impact of having such high concentrations of salt was investigated. There was indeed a significant difference between the inhibition caused by the buffer at 20-fold concentration and the negative control sample. This brought into question the validity of the 'hits' identified using these buffers and meant that it was not possible to use a higher concentration of venom protein in the screen as this would have exacerbated the toxic effect. It was decided that this buffer was unsuitable for second dimension screening and a different approach was attempted.

### 4.4.3 Optimising the Organic SEC Buffers to Improve Chromatographic Separation

As the phosphate salt containing ion exchange buffers were unsuitable, organic solvent based reverse phase buffers were used for separation as lyophilisation removed the volatile solvents, leaving uncontaminated fraction samples.

Various factors play a role in separation of proteins and peptides by HPLC. Factors such as pH have an important effect with improved chromatographic separation occurring at lower pHs (Nugent *et al.*, 1988). Inclusion of ion-pairing acids such as trifluoroacetic acid (TFA), Heptafluorobutyric acid (HFBA) and phosphoric acid can have a profound effect on peak width and retention time of proteins during HPLC. The optimal concentrations for separation of proteins and peptides depends largely on the characteristics of each individual protein. Lower concentrations of ion-pairing acids improve recovery of hydrophobic proteins whilst higher concentrations improve more hydrophilic proteins (Nugent *et al.*, 1988).

It is possible for proteins to bind irreversibly to the stationary phase when TFA or another suitable ion-pairing acid is absent from the mobile phase (Bobaly *et al.*, 2015). The effect of acid addition should be closely monitored as there is an optimal concentration for each protein above or below which recovery of protein during HPLC is compromised (Bobaly *et al.*, 2015).

TFA was chosen as the ion-paring acid to be used for SEC separation. During the optimisation stage, the main difference between the buffers was the concentration of TFA used in each case. In the literature, concentrations of both 0.05% (Bonfim *et al.*, 2001, Ali *et al.*, 1999) and 0.1% (Nawarak *et al.*, 2003, Bonfim *et al.*, 2001, Juàrez *et al.*, 2006, Sanz *et al.*, 2006) can be found. TFA is regularly used in the separation of snake venoms. In the presented study, both concentrations were tested along with a more concentrated 0.2% TFA concentration. In the tested venom samples, the higher the TFA concentration, the more defined and separated the peaks became, so the highest tested concentration of TFA was selected. Lowering the flow rate made no difference to the quality of separation, so the higher flow rate was selected. Cell screening was run in the usual manner, however the second dimension SEC fractions failed to provide any active fractions.

It was unclear why the second dimension SEC HPLC did not produce any actives. It may be due to the synergistic activity some venom toxins exhibit (Kini, 2002, Borkow *et al.*, 1993) and by separating them from one another, they may lose most or all of their potency.

It was decided that the phases would be reversed to see if the SEC HPLC worked successfully as the first dimension. From the first-dimension separation phase, two fractions were selected as active and taken forwards to second dimension separation. Although they did not exceed the set Z score threshold, this threshold was set fairly arbitrarily and it is common to take forward a proportion of the highest scoring compounds during a screen (Brideau *et al.*, 2003). Both of the selected fractions were further fractionated using RP HPLC, however no hits were found in either case. This could be due to the previously mentioned synergism, or because the stability of the venom proteins was being compromised in some way. This was investigated further using nanoDSF technology and is discussed below.

### 4.4.4 Investigation of Venom Fraction Protein Stability using nanoDSF Technology

Lyophilised venoms protected from air exposure can be stable for at least 80 years with very little reduction in activity compared to freshly collected venom samples (Jesupret *et al.*, 2014). However, choice of buffer in which to resuspend dried venoms can have a large impact on their activity.

Protein unfolding characteristics were investigated for *Naja nigricollis* venom using the nanoTemper Prometheus. The venom was made up with various diluents to observe changes in melt temperatures when exposed to different diluent conditions. All diluents (except purified water) made a significant difference to the melt temperature, causing protein unfolding at much lower temperatures in some cases.

It is possible that exposure to different diluents, particularly the HPLC buffers containing TFA, cause conformational changes to native protein structures, leading to a reduction or even loss of activity *in vitro* (Uversky and Finkelstein, 2019). It has been noted that TFA concentrations exceeding 0.01% can cause significant conformational changes in protein structure (Bobály *et al.*, 2014). Exposure to high levels of TFA may cause irreversible changes to the protein molecules, rendering them less potent or even disrupting their activity entirely. When lower concentrations of TFA were used in the original first dimension HPLC, more molecules displaying bioactivity were retained. At this low concentration of TFA, however, adequate SEC chromatographic separation was not achieved.

It was decided that a combination of ion exchange (IE) chromatography followed by RP HPLC to de-salt and further purify the fractions would be used to separate the whole venoms.

### 4.4.5 Two-Dimensional Cobra Venom Screening

Two-dimensional screening revealed at least one active fraction from each of the four venoms tested. To be deemed an active fraction, the sample was required to cause at least 25% inhibition. This was to reduce the chance of false positives being taken forwards (Sink *et al.*, 2010). In addition, to be taken forwards to the dose response stage, the fractions needed to show selectivity for the BxPC-3 cell line over the SW620 cell line. This was to find fractions that were targeting something specifically in PC cells over CRC cells, or that had a more potent effect in PC cells. PC is notoriously difficult to treat and novel treatment options are urgently needed (Hidalgo *et al.*, 2015, Ilic and Ilic, 2016) and therefore this study is focussed primarily on potential treatment options for PC. The focus of the current study aside, any venom fractions which showed selectivity in CRC cells over PC cells could be investigated further as potential CRC treatment options in the future.

To further this study, the active fractions should be tested in non-cancerous human cell lines to assess their cytotoxicity to non-cancerous cells compared with cancerous cells. To become a lead molecule for potential development as a drug, the fractions should ideally cause less cytotoxicity in non-cancerous human cells compared with cancerous cells. In addition, it would be preferable if the selected fraction target cancerous cells preferentially over non-cancerous cells such that it would accumulate in greater quantities in the cancerous cells when introduced to a patient's body. This could be achieved, for example, by targeting different metabolic requirements of tumour cells versus healthy cells (Ngoi *et al.*, 2020, Vander Heiden, 2011, Ko *et al.*, 2010), or different energy requirements of mitochondria in tumour cells through oxidative phosphorylation (Esparza-Molto and Cuezva, 2020).

The fractions were tested for statistical significance using Kruskal-Wallis testing followed by Dunn's post-hoc test due to the non-parametric nature of the data. Fractions that were identified as having activity were selected and dose response curves were performed on the six selected 'hits' to determine their approximate potency and to compare the pharmacological response of the two different cell lines. Due to restrictions on the quantity of each sample available, it was not possible to go above 160 µg/ml as the highest dose for the dose-response assay. To continue the study, more sample should be obtained to increase the maximum concentration tested for those samples to give all samples the opportunity to reach their maximal response.

The dose-response curves for N.mos_i26r2, N.nct_i18r2 and N.pal_i15r4 were quite similar and showed a classic sigmoidal dose-response relationship showing maximum potency with a similar efficacy in each case. These fractions did not show selectivity for either cell line over the other.

The dose response curves for N.nct_i18r2 and N.pal_i15r4 show a very steep slope mostly with just one or two points falling on the curve between minimal and maximal potency. This suggests a very narrow range between the venom fractions having no measurable effect and having maximal effect. This could be due to activation of an apoptotic pathway or other genetic cytotoxic mechanism that once the fraction concentration exceeds a particular threshold, the response increases dramatically and reaches the maximal response within one or two $\log_2$ increases in concentration. The effect of selected fractions on cancer related gene expression pathways will be investigated further in Chapter 6.

The cells treated with N.mos_i26r2 show a slightly shallower dose response gradient in BxPC-3 cells compared with N.nct_i18r2 and N.pal_i15r4 treated samples. This suggests a pathway of cytotoxicity that is being activated by lower doses but requires a higher concentration of the fraction to reach its maximum effect. It was not possible to accurately determine an $IC_{50}$ value for this venom fraction in this cell line using Graphpad Prism 9.0.0 (GraphPad Software, San Diego, USA) as the graph did not plateau at the top of the graph. Treatment of cells with N.pal_i18r2 gave no measurable effect in either cell line up to the 160 µg/ml dose. To better assess the potency and pharmacological profiles, treatment using both N.mos_i26r2 and N.pal_i18r2 samples could be repeated at a higher dose if this becomes a viable option in the future. This would require harvesting more crude venom and several rounds of the 2D HPLC protocol discussed in section 2.5.6 (first dimension) and 2.5.7 (second dimension).

### 4.4.6 N.nub_17r2 and N.pal_i17r2 Comparison

N.nub_i17r2 and N.pal_i17r2 have different pharmacological profiles compared to the three fractions previously mentioned. They have very similar retention times at both first dimension

(IE) and second dimension (RP) HPLC. This suggests very similar characteristics of both proteins. As each amino acid has unique chromatographic properties, even small changes to amino acid composition of proteins can lead to large changes in retention time in one or more form of HPLC (Marek *et al.*, 2018). This suggests the amino acid sequences of these fractions are either identical or have a small number of mutations leading to proteins with very similar ionic and hydrophilic characteristics.

Both fractions caused a similar degree of cytotoxicity against the two cancer cell lines tested. Both fractions appeared to be having a full inhibitory effect against BxPC-3 cells whilst they both seem to only have a partial inhibitory effect on SW620 cells with the dose response curve appearing to be heading towards its maximal inhibition at a much lower percentage inhibition compared to BxPC-3 cells (Ezechias and Cajthaml, 2018).

Taking together the retention times of both dimensions, the percentage inhibition during the 2D screen and the shape of the dose response curves, it appears that N.nub_i17r2 and N.pal_i17r2 may be the same biochemical entity. This will be further investigated by mass spectrometry in Chapter 5.

### 4.4.7 Chapter Conclusion

Overall, 11 fractions were identified that have a significant cytotoxic effect out of the 303 tested. Six out of 11 were selected based on their potency and availability of sample. Of these, five fractions proved to have interesting dose-response relationships. Limited resources necessitated choosing four out of five venom fractions to take forwards for Mass Spectroscopy (MS) analysis in Chapter 5. The purified toxins from N.nct_i18r2 and N.pal_i15r4 were selected to be taken forwards as they have very similar pharmacological profiles despite having quite different cytotoxicity values and HPLC retention times. N.nub_i17r2 and N.pal_i17r2 on the other hand had extremely similar cytotoxicity values and HPLC retention times and so this pair were also selected for MS analysis.

# Chapter 5 - Identifying Active 2D Venom Fractions and Determining Residues Important for Their Activity

## 5.1 Introduction

As the fractions identified in Chapter 4 originated from a natural source (venom), rather than a synthetically or recombinantly produced protein, the identities of these venom components were unknown. To progress the active fractions as potential lead molecules, the sequence of the proteins needed to be ascertained. Knowing the sequences of the active venom components allowed them to be compared using structure activity relationship (SAR) analysis. This method is useful in drug discovery as it allows the structures of many molecules (including proteins and peptides) to be compared and to assess the impact of each atom (or AA) on the activity of the molecule or protein. This helps to identify key areas of the molecule/protein that lead to the desired effect against the target of interest (or disease state) and which changes cause improved activity or potency of the molecule/protein.

### 5.1.1 Mass Spectrometry

Mass spectrometry (MS) is a highly sensitive technique for biophysical characterisation, able to elucidate structures of molecules including large and more complex biomolecules such as peptides, proteins, polysaccharides, etc. through fragmentation of their ions (De Hoffmann, 2005). MS offers many advantages for the study of protein structure including small sample quantity requirement, almost no limit to protein size and the ability to assess proteins in their native (or near-native) states (Zhang *et al.*, 2014). In addition, MS is a direct measure and does not require any kind of labelling, making it less susceptible to potential interference and artefacts such as auto-fluorescence, fluorescence quenching and light scattering (Mclaren *et al.*, 2021).

MS involves ionisation of samples followed by separation and detection of the ionised components based on their mass to charge ratio (m/z). There are three main stages of MS: the ionisation stage, mass analysis stage and spectrum analysis stage. There are a variety of analyte ionisation techniques with the most frequently used methods for protein MS being electrospray ionisation (ESI) and matrix-assisted laser desorption/ionisation (MALDI) due to their ability to ionise larger, highly polar molecules (Rozanova *et al.*, 2021). The three main mass analysers used for protein MS are quadrupole, time of flight (ToF) and Orbitrap (Rozanova *et al.*, 2021). Pairing together two or more mass analysers is known as tandem MS (or MS/MS) and allows further separation of analytes.

Effect-directed analysis (EDA) is a technique used to identify environmental toxin components, particularly those causing disease or other adverse effects. EDA consists of extraction, fractionation, bioassay testing which is often cell-based, and finally identification using a suitable analytical technique (Hong *et al.*, 2016, Jonker *et al.*, 2015). This methodology is often applied to study soil, crude oil, contaminated wastewater and other such environmental concerns, utilising ToF-MS and Orbitrap MS techniques (Hong *et al.*, 2016), but may also be applied to more complex mixtures such as venom and other animal toxins. Indeed, a multitude of venoms have been analysed this way, usually utilising a form of chromatography (most commonly liquid), followed by MS analysis. Details of a few examples of venoms which have been separated in this manner are shown in Table 5.1.

MS is increasingly being utilised for high throughput applications and is proving to be a valuable technique for elucidating protein structure, dynamics and ligand-protein and protein-protein interactions (Gheyi and Molina-Martin, 2018). Acoustic mist ionisation MS (AMI-MS) offers increased throughput by utilising modified acoustic liquid handling technology to dispense ultra-low volumes of analytes directly from 384- or 1536-well microtiter plates into the mass spectrometer using sound waves, analysing each well in less than one second (Sinclair *et al.*, 2019a, Sinclair *et al.*, 2019b, Belov *et al.*, 2020).

*Table 5.1.Examples of venoms which have been analysed using various forms of MS.*

*RP HPLC = reverse phase high performance liquid chromatography, SEC = size exclusion chromatography, SDS-PAGE = sodium dodecyl sulphate- polyacrylamide gel, IE = ion exchange, HILIC = hydrophilic interaction chromatography, ESI = electro-spray ionisation, MALDI = matrix-assisted laser desorption/ionisation, ToF = time of flight, MS/MS = tandem MS, Q(q)ToF=quadrupole (quadrupole) time of flight.*

| Species | Name(s) | Separation type | Ionisation method | Mass analyser technique | Author |
|---|---|---|---|---|---|
| Invertebrates | | | | | |
| Ant | *82 species* | RP HPLC | MALDI | ToF | (Touchard *et al.*, 2015) |
| Red-tailed Bumblebee | *Bombus lapidarius* | - | ESI, MALDI | QqToF MS/MS | (Favreau *et al.*, 2006) |
| Social Paper Wasp | *Polistes dominulus* | - | MALDI | ToF | (Turillazzi *et al.*, 2006) |
| East European Hornet | *Vespa orientalis* | SEC, HPLC | ESI, MALDI | ToF | (Klochkov *et al.*, 2008) |
| Funnel-web Spider | *Illawarra wisharti, Hadronyche cerberea, Hadronyche infensa Toowoomba* | RP HPLC | MALDI | ToF | (Palagi *et al.*, 2013) |
| Amazonian scorpion | *Tityus cambridgei* | RP HPLC | ESI, MALDI | ToF | (Batista *et al.*, 2004) |
| Snakes | | | | | |
| Egyptian cobra | *Naja haja legionis* | SEC, RP HPLC, SDS-PAGE, tryptic digest | nano-ESI | QToF MS/MS | (Malih *et al.*, 2014) |
| Eastern Indian Cobra | *Naja naja* | IE HPLC, SDS-PAGE (+ tryptic digest), RP HPLC | nano-ESI | Orbitrap MS/MS | (Dutta *et al.*, 2017) |
| African spitting cobras (Mozambique spitting and Zebra cobra) | *Naja mossambica and Naja nigricincta* | SDS-PAGE (+ tryptic digest), RP-HPLC | nano-ESI | Quadrupole, Orbitrap | (Katali *et al.*, 2020) |
| African cobras (Egyptian, Mali and Black-necked spitting cobra) | *Naja haje, Naja katiensis and Naja nigricollis* | SDS-PAGE, HILIC | nano-ESI | Orbitrap, Quadrupole | (Adamude *et al.*, 2021) |

### 5.1.2 Bioinformatics

Bioinformatics is a highly interdisciplinary field incorporating mathematics, physics and computer science to address biological questions. It involves applying computation and analysis to interpret large sets of biological data (Bayat, 2002). In this chapter, some Bioinformatic tools were applied to help ascertain the sequences of the venom fraction proteins from the MS data.

#### 5.1.2.1 Protein Basic Local Alignment Search Tool (BLASTp)

BLASTp is a bioinformatic tool which is part of the National Centre for Biotechnology Information (NCBI) suite that compares the sequence of two or more proteins (Altschul *et al.*, 1990) and is located at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins. Submitting a protein sequence query to BLASTp produces a results page containing related sequences showing the overall score of that protein in relation to the searched protein. This score is calculated by aligning the two sequences, assigning a value to each pair of AAs and adding these values over the length of the protein (Wheeler and Bhagwat, 2007). The BLASTp summary also lists other scores which show how well the aligned sequence matches the search query. These include the percent of the query which is covered in the search (the query cover), a score for the random background noise (the Expect value, or E-value), and the percent identity (Per. Ident) showing the degree of identical AA in the same position between the two sequences.

#### 5.1.2.2 Multiple Sequence Alignments

Another bioinformatic technique used in this chapter is Multiple Sequence Alignment (MSA). This involves aligning two or more sequences such that the homology between similar sequences may be studied. One such tool is called Clustal Omega and is provided by the European Bioinformatics Institute as part of the European Molecular Biology Laboratory (EMBL-EBI) (https://www.ebi.ac.uk/Tools/msa/clustalo/).

These aligned sequences can then be interrogated further to see where mutations have occurred in the AA sequence and contribute towards determining the evolutionary relationship between different organisms.

#### 5.1.2.3 Databases and Web Tools

The Mascot server is a web-based client provided by Matrix Science at https://www.matrixscience.com/. Mascot is used for identifying proteins from Peptide Mass Fingerprint and MS/MS data by performing data queries on MS mass lists. It is compatible with data output from a wide range of MS instruments. Mascot searches a variety of commonly used sequence databases including SwissProt (https://www.uniprot.org/uniprot/?query=reviewed:yes), non-redundant proteins from the National Centre for Biotechnology Information (NCBInr)

(https://www.ncbi.nlm.nih.gov/refseq/about/nonredundantproteins/) and the European Molecular Biology Laboratory Expressed Sequence Tags (EMBL EST) divisions (https://www.ebi.ac.uk/ena/browser/home).

Swiss-Prot is a protein sequence databank containing annotated, non-redundant protein sequences cross-referenced to a number of complementary databases such as the EMBL, NCBI and Protein Data Bank (PDB) (Bairoch and Apweiler, 1996).

The PDB is a database containing three-dimensional structures of proteins, nucleic acids and other large biomolecules obtained by X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy or cryo-electron microscopy (cryoEM). There are several member organisations internationally that contribute to the Worldwide Protein Data Bank (wwPDB). One such organisation is the Research Collaboratory for Structural Bioinformatics (RCSB) which is located here: https://www.rcsb.org/.

#### 5.1.2.4 Structure Activity Relationships

A structure-activity relationship (SAR) describes the relationship between the structure of a chemical or biological entity and the biological activity observed on the target (or targets) of interest. Using this information, chemical or biological entities may be structurally modified to improve one or more property and lead to a better drug-like molecule. These improved properties may include increased potency, improved target selectivity, decreased toxicity or improved bioavailability, amongst others (Guha, 2013).

WebProAnalyst is a web tool available for assessing SAR between similar proteins and is located at http://wwwmgs.bionet.nsc.ru/mgs/programs/panalyst/ . WebProAnalyst uses aligned protein sequences and their activities (e.g. $LD_{50}$) to determine correlations between different AAs. The output may include a sequence-activity correlation (or determination) coefficient (SACC/SADC). The SADC is a calculation of the proportion of the variation in observed activity due to AA substitution at each location in the protein sequence, and the SACC is the square root of the SADC (Ivanisenko *et al.*, 2005).

### 5.1.3 Chapter 5 Aims

The aims of this chapter were to:

- identify protein sequences of the selected venom fractions (N.pal_i15r4, N.pal_i17r2, N.nct_i18r2 and N.nub_i17r2) using MS intact mass analysis and peptide mapping.
- build consensus sequences based on MS and similar proteins existing in the Swissprot database through BLASTp searches and multiple sequence alignments.
- perform SAR to determine the most important residues affecting potency.

## 5.2 Results

### 5.2.1 Mass Spectrometry

The venom samples that were sent for mass spectrometry (MS) analysis are detailed in the following pages. For each sample, a summary of the MS data is provided with multiple sequence alignments (MSA) performed for each of the identified fragments. In addition, a 'template' sequence as proposed by Konshina *et al*. (2017) to be conserved in approximately 80% of cobra cardiotoxins is included in each MSA, plus the top matches for that protein according to the Mascot database.

Details of the actual masses measured by MS along with the details of the closest matched protein from the SwissProt database according to Mascot are displayed in Table 5.2. The MS spectra are shown in Figure 5.1. The dose required to cause death in 50% of the test population ($LD_{50}$) represents the acute toxicity of a compound. These values were obtained from the Swissprot database.

*Table 5.2. Details of the actual mass measured by MS analysis plus the closest match in the Swissprot database according to Mascot.*

*The species (taxonomy), protein name, Swissprot database ID, mass and measured $LD_{50}$ values (where available in Swissprot) are displayed.*

| Sample | N.pal_i15r4 | N.pal_i17r2 | N.nct_i18r2 | N.nub_i17r2 |
|---|---|---|---|---|
| **Main peak (Da)** | 6,818 | 6,735 | 6,708 | 6,735 |
| **Minor peak(s) (Da)** | 6,952 | 6,904 | 7,281<br>7,522 | |
| **Closest Match for main peak** | | | | |
| **Taxonomy** | *Naja pallida* | *Naja nivea* | *Naja mossambica* | *Naja nivea* |
| **Protein** | Cytotoxin 1 | Cytotoxin 1 | Cytotoxin 4 | Cytotoxin 1 |
| **Swissprot ID** | 3SA1_NAJPA | 3SA1_NAJNI | 3SA4_NAJMO | 3SA1_NAJNI |
| **Mass (Da)** | 6,827 | 6,697 | 6,715 | 6,697 |
| **$LD_{50}$** | unknown | 2.9 mg/kg | 1.97 mg/kg | 2.9 mg/kg |

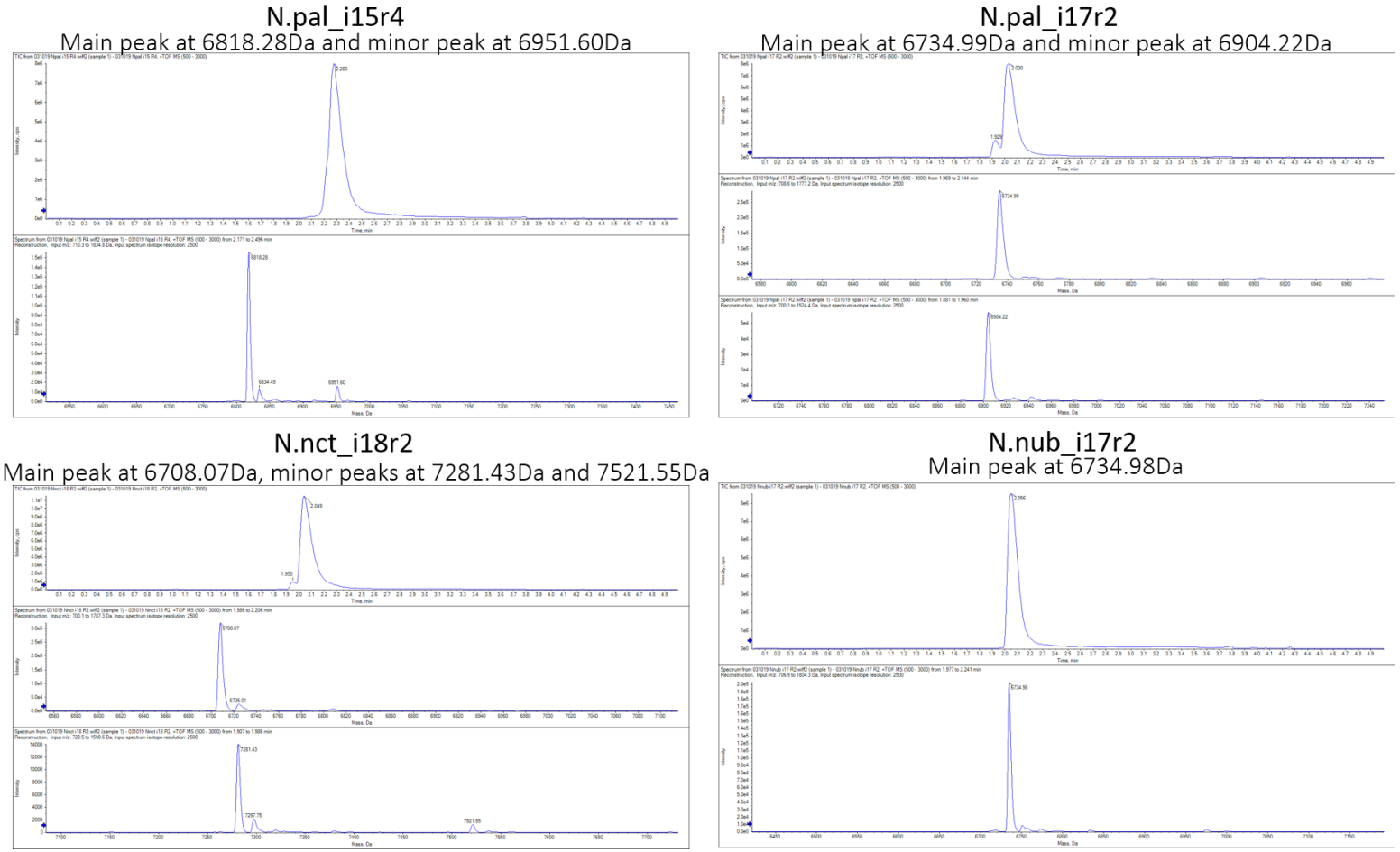



*Figure 5.1. MS spectra obtained from the four cobra venom fractions analysed.*

*(a) N.pal_i15r4 has a main peak at 6818.28 Da and a minor peak at 6951.60 Da. (b) N.pal_i17r2 has a main peak at 6734.99 Da and a minor peak at 6904.22 Da. (c) N.nct_i18r2 has a main peak at 6708.07 Da and minor peaks at 7281.43 and 7521.55 Da. (d) N.nub_i17r2 has a main peak at 6734.98 Da. No minor peaks were detected for this sample.*

## 5.2.2 Bioinformatic Analysis

The following sections break down the individual venom fractions that were selected in Chapter 4. For each venom fraction, the fragments identified during MS peptide mapping were compared to the sequences of the top results from the Swissprot database identified by Mascot. The fragments were also aligned with a 'template' sequence based on 80% sequence homology of cardiotoxin proteins (Konshina *et al.*, 2017), as shown in Figure 5.2.

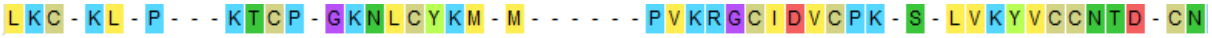



*Figure 5.2. Template cardiotoxin sequence*

*This sequence represents conservation of approximately 80% when all cobra cytotoxins from the Uniprot database are aligned (Konshina et al., 2017)*

### *5.2.2.1 N.pal_i15r4*

The N.pal_i15r4 venom fraction sample had one major peak and one minor peak.

**Major Peak**

The N.pal_i15r4 major peak had two cardiotoxin entry matches according to MS peptide mapping matched with the Mascot database (see Appendix IV for details of peptide matches). These entries – 3SA1_NAJPA and 3SA1_NAJNA – were aligned with the template cardiotoxin structure (Figure 5.2) and the peptide fragments identified by MS (details of peptide mapping found in Appendix IV). This multiple sequence alignment (MSA) is shown in Figure 5.3.

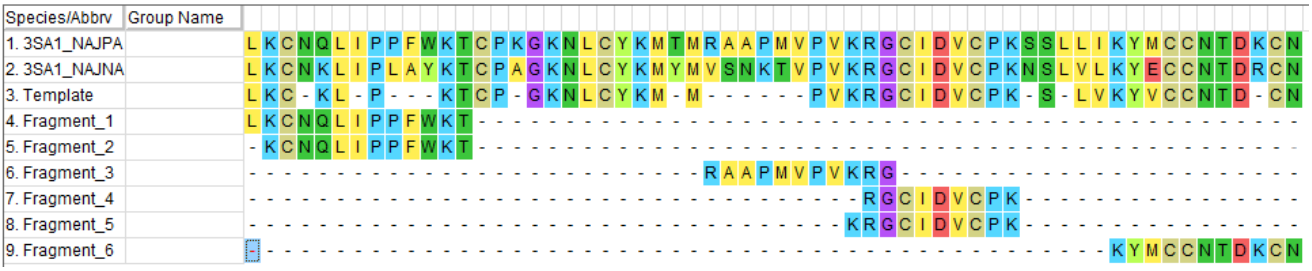



*Figure 5.3. Multiple sequence alignment of the N.pal_i15r4 fragments identified by mass spectrometry.*

*The fragment sequences were aligned with the two top scoring hits from the Mascot data search along with the template sequence (Konshina et al., 2017).*

This alignment gave a consensus sequence of:

**LKCNQLIPPFWKTCP-GKNLCYKM-MRAAPMVPVKRGCIDVCPK-SL--KYMCCNTDKCN.**

A BLASTp search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was performed on the consensus sequence. Details of the top five matches originating from *Afronaja* subgenera (according to classification by Wallach et al. (2009)) for this search are shown in Table 5.3.

*Table 5.3. Details of the five closest BLASTp matches to N.pal_15r4.*

*BLASTp search was performed against consensus sequence LKCNQLIPPFWKTCP-GKNLCYKM-MRAAPMVPVKRGCIDVCPK-SL--KYMCCNTDKCN. Only proteins from species within the Afronaja subgenera were included in this top 5 list.*

| | Protein name | Species | Locus | Uniprot ID | Total score | Query cover | E Value | % Identity |
|---|---|---|---|---|---|---|---|---|
| 1 | Cytotoxin 1 | *Naja pallida* | 3SA1_NAJPA | P01468.1 | 98.2 | 100% | 2e-28 | 90% |
| 2 | Cytotoxin 1 | *Naja mossambica* | 3SA1_NAJMO | P01467.1 | 96.3 | 100% | 1e-27 | 88.33% |
| 3 | Cytotoxin 2 | *Naja mossambica* | 3SA2_NAJMO | P01469.1 | 92.4 | 100% | 5e-26 | 85% |
| 4 | Cytotoxin 3 | *Naja mossambica* | 3SA3_NAJMO | P01470.1 | 90.1 | 100% | 3e-25 | 83.33% |
| 5 | Naniproin | *Naja nigricollis* | 3SAN_NAJNG | P0DSN1.1 | 88.2 | 100% | 2e-24 | 81.67% |

These top 5 matches were aligned using MEGA7 and are shown in Figure 5.4, along with the template cardiotoxin structure (Figure 5.2) and the peptide fragments.

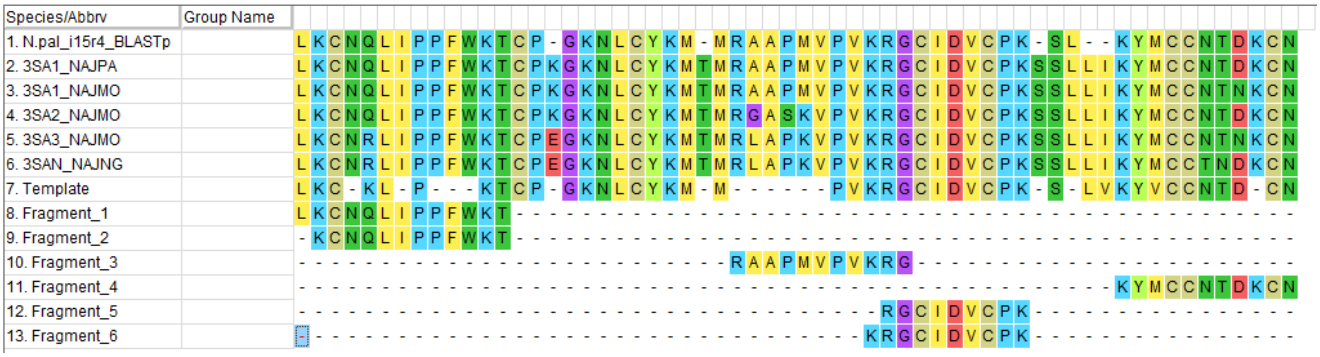



*Figure 5.4. Multiple sequence alignment against N.pal_i15r4.*

*N.pal_i15r4 consensus sequence used for BLASTp search, the top five results from the BLASTp search, the template sequence and the six fragments identified by MS.*

AA residues that were identical for all five of the top *Afronaja* sequences were assumed to be correct for the unknown N.pal_i15r4 peptide, giving the consensus sequence:

**LKCNQLIPPFWKTCP-GKNLCYKMTMRAAPMVPVKRGCIDVCPKSSLL-KYMCCNTDKCN**

Estimations of the molecular weight of the protein sequence when different AA residues were inserted into the unknown gaps are shown in Figure 5.5. The measured molecular weight for the unknown N.pal_i15r4 protein was 6818.28 Da.

LKCNQLIPPFWKTCP-GKNLCYKMTMRAAPMVPVKRGCIDVCPKSSLL-KYMCCNTDKCN

----------1----------2----------

| 1 | 2 | Complete Sequence | Mass (Da) |
|---|---|---|---|
| K | I | LKCNQLIPPFWKTCP**K**GKNLCYKMTMRAAPMVPVKRGCIDVCPKSSLL**I**KYMCCNTDKCN | 6827.41 |
| K | V | LKCNQLIPPFWKTCP**K**GKNLCYKMTMRAAPMVPVKRGCIDVCPKSSLL**V**KYMCCNTDKCN | 6813.38 |
| E | I | LKCNQLIPPFWKTCP**E**GKNLCYKMTMRAAPMVPVKRGCIDVCPKSSLL**I**KYMCCNTDKCN | 6828.35 |
| E | V | LKCNQLIPPFWKTCP**E**GKNLCYKMTMRAAPMVPVKRGCIDVCPKSSLL**V**KYMCCNTDKCN | 6814.32 |

*Figure 5.5. Estimation of molecular weight of theoretical CT molecules based upon the consensus sequence.*

*Consensus sequence is LKCNQLIPPFWKTCP-GKNLCYKMTMRAAPMVPVKRGCIDVCPKSSLL-KYMCCNTDKCN. Unknown AA residues chosen from the different AAs present in the top five matched Afronaja species CT sequences. Calculations performed at https://web.expasy.org/compute_pi/.*

Unfortunately, this MW does not perfectly match any of the options given in Figure 5.5. A matrix of other possible amino acid substitutions can be found in Appendix V. From this matrix, theoretically adding the molecular weight for a proline and a valine in the unknown regions (most likely proline in position 1 and valine in position 2), could add to the desired 6818.28 Da molecular weight. However, it was not possible to determine if this gives the correct full sequence without further experimental investigation and using an incorrect sequence may bias the SAR analysis. Therefore, these regions were left blank and the consensus sequence was used for N.pal_i15r4 from this point onwards:

**LKCNQLIPPFWKTCP-GKNLCYKMTMRAAPMVPVKRGCIDVCPKSSLL-KYMCCNTDKCN**

### Minor Peak

The N.pal_i15r4 venom fraction sample also had a minor but significant peak in the MS spectrum. This minor peak had top matches in the Mascot database to Basic Phospholipase $A_2$ from *Naja melanoleuca, Naja mossambica, Naja nigricollis* and two *Oxyuranus* species. These matches only had a score of 44 (compared to a score of 242 for the main peak) indicating that the match and/or sequence coverage is not particularly high. The intact mass from the peak on the MS trace (6951.6 Da) is much smaller than the anticipated ~14,000 Da of these Phospholipases in the Uniprot database and this could explain the lower score given.

### *5.2.2.2 N.pal_i17r2*

**Major Peak**

N.pal_i17r2 had four cardiotoxin entry matches according to MS peptide mapping matched with the Mascot database (see Appendix IV for details of peptide mapping). These entries – 3SA1_NAJNI, 3SA2_NAJHA, 3SA6_NAJHA and 3SA9_NAJHA – were aligned with the template cardiotoxin structure (Figure 5.2) and the fragments identified by MS (details of peptide mapping found in Appendix IV). This MSA is shown in Figure 5.6.

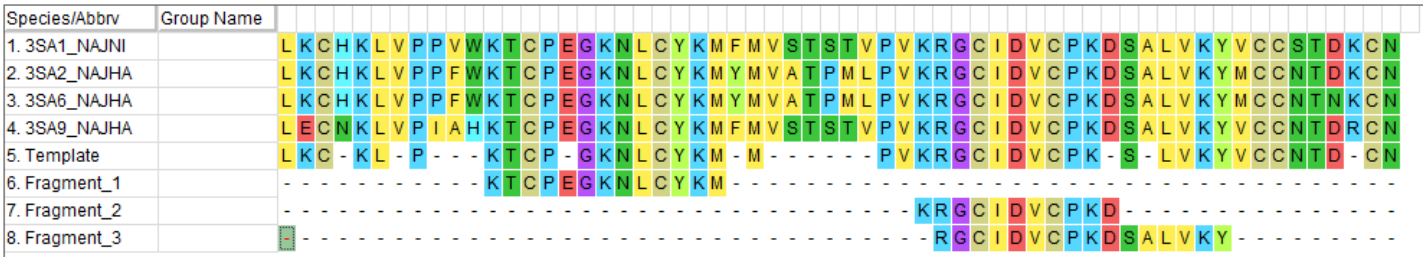



*Figure 5.6. Multiple sequence alignment of the N.pal_i17r2 fragments identified by mass spectrometry*

*The fragment sequences were aligned with the four top scoring hits from the Mascot data search along with the template sequence (Konshina et al., 2017).*

This alignment gave a consensus sequence of:

**LKC-KLVP---KTCPEGKNLCYKM-MV-T---PVKRGCIDVCPKDSALVKY-CC-T--CN.**

A BLASTp search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was performed on the consensus sequence. Details of the top five matches originating from *Afronaja* subgenera (according to classification by Wallach et al. (2009)) for this search are shown in Table 5.4.

*Table 5.4. Details of the five closest BLASTp matches for N.pal_i17r2.*

*BLASTp search was performed against consensus sequence LKC-KLVP---KTCPEGKNLCYKM-MV-T---PVKRGCIDVCPKDSALVKY-CC-T--CN. Only proteins from species within the Afronaja subgenera were included in this top 5 list.*

| | Protein name | Species | Locus | Uniprot ID | Total score | Query cover | E Value | Per. Ident |
|---|---|---|---|---|---|---|---|---|
| 1 | Cytotoxin 4 | *Naja mossambica* | 3SA4_NAJMO | P01452.1 | 63.9 | 93% | 6e-15 | 71.7% |
| 2 | Cytotoxin 5 | *Naja mossambica* | 3SA5_NAJMO | P25517.3 | 59.3 | 93% | 4e-13 | 67.92% |
| 3 | Naniproin | *Naja nigricollis* | 3SAN_NAJNG | P0DSN1.1 | 58.9 | 100% | 6e-13 | 64.29% |
| 4 | Cytotoxin 3 | *Naja mossambica* | 3SA3_NAJMO | P01470.1 | 58.5 | 93% | 9e-13 | 66.04% |
| 5 | Cytotoxin 1 | *Naja mossambica* | 3SA1_NAJMO | P01467.1 | 56.6 | 93% | 5e-12 | 64.15% |

Figure 5.7 shows these five proteins aligned against the consensus sequence, the template sequence (Figure 5.2) and the fragments identified in MS peptide mapping. AA residues that were identical for all five of the top *Afronaja* sequences were assumed to be correct for the unknown N.pal_i17r2 sequence.

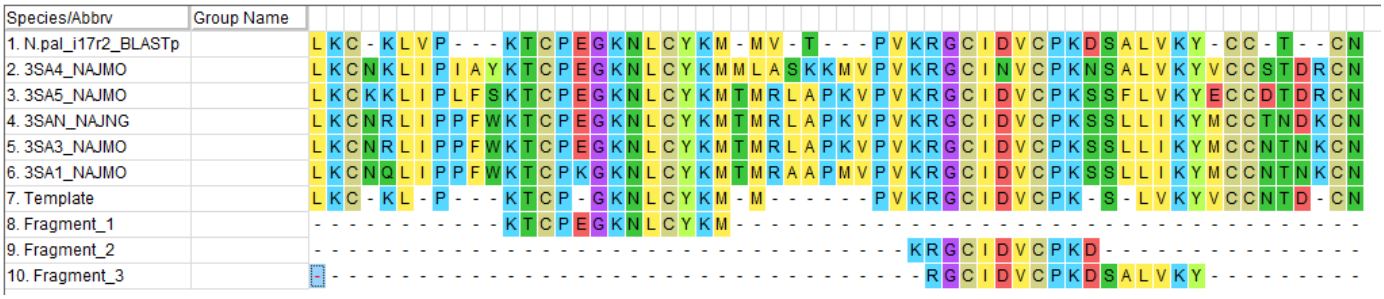



*Figure 5.7. Multiple sequence alignment for N.pal_i17r2*

*Alignment of the consensus sequence used for a BLASTp search, the top five results from cobras in the Afronaja subgenera from the BLASTp search, the template sequence and the three fragments identified by MS.*

Unfortunately, the alignment shown in Figure 5.7 did not provide any further agreement for the missing AA residues and therefore N.pal_i17r2 had an overall consensus sequence for SAR of:

**LKC-KLVP---KTCPEGKNLCYKM-MV-T---PVKRGCIDVCPKDSALVKY-CC-T--CN**.

### Minor Peak

N.pal_i17r2 had a significant minor peak identified with a mass of 6904.22 Da, however the peptide mapping could not find any database matches for this peptide.

### *5.2.2.3 N.nct_i18r2*

**Major Peak**

N.nct_i18r2 had three cardiotoxin entry matches according to MS peptide mapping matched with the Mascot database (see Appendix IV for details of peptide mapping). These entries – 3SA1_NAJPA, 3SA2_NAJMO and 3SA4_NAJMO – were aligned with the template cardiotoxin (Figure 5.2) and the fragments identified by MS (details of peptide mapping found in Appendix IV). This MSA is shown in Figure 5.8.

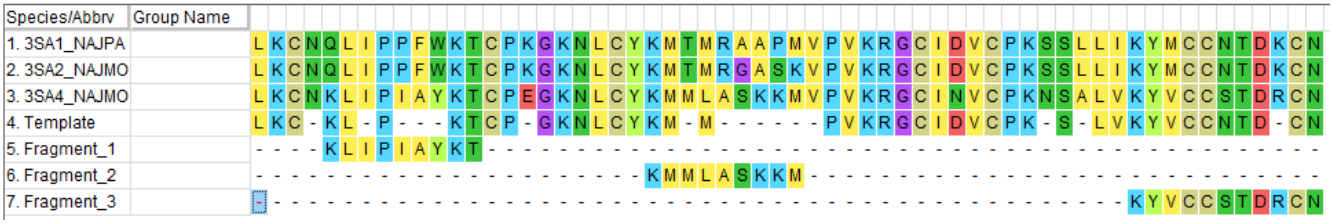



*Figure 5.8. Multiple sequence alignment of the three N.nct_i18r2 fragments identified by mass spectrometry.*

*The fragment sequences were aligned with the three top scoring hits from the Mascot data search along with the template sequence (Konshina et al., 2017).*

This alignment gave a consensus sequence of:

**LKCNKLIPIAYKTCP-GKNLCYKMMLASKKMVPVKRGCI-VCPK-S-L-KYVCCSTDRCN**

A BLASTp search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was performed on the consensus sequence. Details of the top five matches originating from *Afronaja* subgenera (according to classification by Wallach et al. (2009)) for this search are shown in Table 5.5.

*Table 5.5. Top 5 results from a BLASTP search against N.nct_i18r2 consensus sequence.*

*The consensus sequence for N.nct_i18r2 was LKCNKLIPIAYKTCP-GKNLCYKMMLASKKMVPVKRGCI-VCPK-S-L-KYVCCSTDRCN. Any proteins originating from cobras other than Afronaja were omitted.*

| | Protein name | Species | Locus | Uniprot ID | Total score | Query cover | E Value | Per. Ident |
|---|---|---|---|---|---|---|---|---|
| 1 | Cytotoxin 4 | *Naja mossambica* | 3SA4_NAJMO | P01452.1 | 92 | 100% | 4e-23 | 88.33% |
| 2 | Cytotoxin 1 | *Naja pallida* | 3SA1_NAJPA | P01468.1 | 73.2 | 100% | 1e-15 | 68.33% |
| 3 | Cytotoxin 1 | *Naja mossambica* | 3SA1_NAJMO | P01467.1 | 71.2 | 100% | 7e-15 | 66.67% |
| 4 | Cytotoxin 2 | *Naja mossambica* | 3SA2_NAJMO | P01469.1 | 70.5 | 100% | 1e-14 | 66.67% |
| 5 | Cytotoxin 5 | *Naja mossambica* | 3SA5_NAJMO | P25517.3 | 67.8 | 100% | 3e-16 | 68.33% |

These top 5 matches were aligned using MEGA7 and are shown in Figure 5.9, along with the template cardiotoxin structure (Figure 5.2) and the MS peptide fragments.

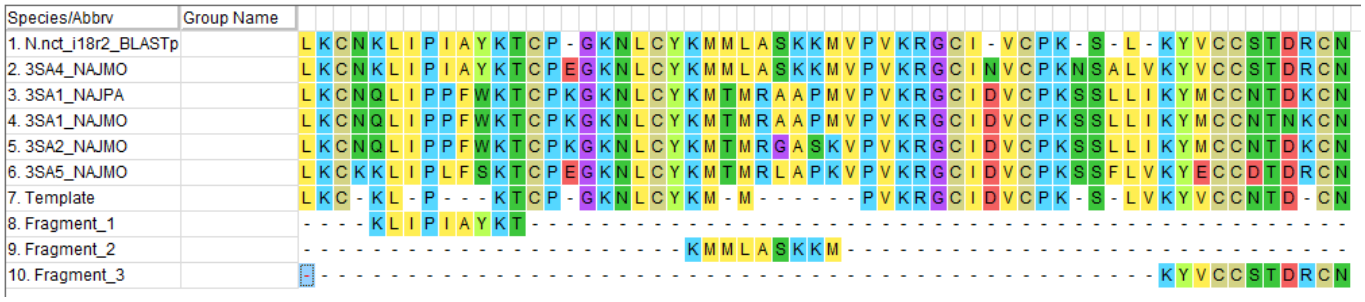



*Figure 5.9. Multiple sequence alignment of the five highest scoring proteins against N.nct_i18r2.*

*Alignment of venom proteins from cobras from the Afronaja subgenera from the N.nct_i18r2 consensus sequence BLASTp query along with the template sequence and the MS fragment sequences.*

AA residues that were identical for all five of the top *Afronaja* sequences were assumed to be correct for the unknown N.nct_i18r2 peptide, giving the sequence:

**LKCNKLIPIAYKTCP-GKNLCYKMMLASKKMVPVKRGCI-VCPK-S-L-KYVCCSTDRCN**

An attempt was made to work out the unknown amino acids (AAs) in the protein sequence by replacing each unknown AA with a possible AA from that position in at least one of the top five matched protein sequences, then calculating the predicted molecular weight of the constructed protein to see if any match the molecular weight measured by MS. This AA replacement is shown in Figure 5.10.

LKCNKLIPIAYKTCP-GKNLCYKMMLASKKMVPVKRGCI-VCPK-S-L-KYVCCSTDRCN

----------1----------2----3-4-5----------

| | 1 | 2 | 3 | 4 | 5 | | Total | MW |
|---|---|---|---|---|---|---|---|---|
| 1 | E | N | N | A | V | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCINVCPKNSALVKYVCCSTDRCN | 6715.26 |
| 2 | E | N | N | A | I | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCINVCPKNSALIKYVCCSTDRCN | 6729.29 |
| 3 | E | N | N | L | V | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCINVCPKNSLLVKYVCCSTDRCN | 6757.34 |
| 4 | E | N | N | L | I | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCINVCPKNSLLIKYVCCSTDRCN | 6771.37 |
| 5 | E | N | S | A | V | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCINVCPKSSALVKYVCCSTDRCN | 6688.23 |
| 6 | E | N | S | A | I | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCINVCPKSSALIKYVCCSTDRCN | 6702.26 |
| 7 | E | N | S | L | V | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCINVCPKSSLLVKYVCCSTDRCN | 6730.32 |
| 8 | E | N | S | L | I | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCINVCPKSSLLIKYVCCSTDRCN | 6744.34 |
| 9 | E | D | N | A | V | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCIDVCPKNSALVKYVCCSTDRCN | 6716.25 |
| 10 | E | D | N | A | I | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCIDVCPKNSALIKYVCCSTDRCN | 6730.27 |
| 11 | E | D | N | L | V | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCIDVCPKNSLLVKYVCCSTDRCN | 6758.33 |
| 12 | E | D | N | L | I | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCIDVCPKNSLLIKYVCCSTDRCN | 6772.35 |
| 13 | E | D | S | A | V | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCIDVCPKSSALVKYVCCSTDRCN | 6689.22 |
| 14 | E | D | S | A | I | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCIDVCPKSSALIKYVCCSTDRCN | 6703.25 |
| 15 | E | D | S | L | V | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCIDVCPKSSLLVKYVCCSTDRCN | 6731.30 |
| 16 | E | D | S | L | I | | LKCNKLIPIAYKTCPEGKNLCYKMMLASKKMVPVKRGCIDVCPKSSLLIKYVCCSTDRCN | 6745.33 |
| 17 | K | N | N | A | V | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCINVCPKNSALVKYVCCSTDRCN | 6714.32 |
| 18 | K | N | N | A | I | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCINVCPKNSALIKYVCCSTDRCN | 6728.35 |
| 19 | K | N | N | L | V | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCINVCPKNSLLVKYVCCSTDRCN | 6756.40 |
| 20 | K | N | N | L | I | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCINVCPKNSLLIKYVCCSTDRCN | 6770.43 |
| 21 | K | N | S | A | V | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCINVCPKSSALVKYVCCSTDRCN | 6687.29 |
| 22 | K | N | S | A | I | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCINVCPKSSALIKYVCCSTDRCN | 6701.32 |
| 23 | K | N | S | L | V | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCINVCPKSSLLVKYVCCSTDRCN | 6729.37 |
| 24 | K | N | S | L | I | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCINVCPKSSLLIKYVCCSTDRCN | 6743.40 |
| 25 | K | D | N | A | V | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCIDVCPKNSALVKYVCCSTDRCN | 6715.30 |
| 26 | K | D | N | A | I | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCIDVCPKNSALIKYVCCSTDRCN | 6729.33 |
| 27 | K | D | N | L | V | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCIDVCPKNSLLVKYVCCSTDRCN | 6757.38 |
| 28 | K | D | N | L | I | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCIDVCPKNSLLIKYVCCSTDRCN | 6771.41 |
| 29 | K | D | S | A | V | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCIDVCPKSSALVKYVCCSTDRCN | 6688.28 |
| 30 | K | D | S | A | I | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCIDVCPKSSALIKYVCCSTDRCN | 6702.31 |
| 31 | K | D | S | L | V | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCIDVCPKSSLLVKYVCCSTDRCN | 6730.36 |
| 32 | K | D | S | L | I | | LKCNKLIPIAYKTCPKGKNLCYKMMLASKKMVPVKRGCIDVCPKSSLLIKYVCCSTDRCN | 6744.39 |

*Figure 5.10. Attempts at working out the amino acid sequence of N.nct_i18r2 based on the MS weight.*

*MW prediction calculated by https://web.expasy.org/compute_pi/. Sequences within 10 Da of the measured MW (6708 Da) are highlighted in green, however, none of the sequences perfectly matched the measured molecular weight.*

None of the predicted MWs perfectly matched the MW measured by MS. N.nct_i18r2 is likely to be extremely similar to 3SA4_NAJMO with point mutations at one or two different sites, leading to a slightly different AA sequence and therefore overall MW. Since it was not possible to determine exactly where these differences occur without further experimental data, a consensus sequence of was used for N.nct_i18r2 from this point onwards:

**LKCNKLIPIAYKTCP-GKNLCYKMMLASKKMVPVKRGCI-VCPK-S-L-KYVCCSTDRCN**

**Minor Peak**

N.nct_i18r2 had two significant minor peaks detected by intact mass analysis. One of these minor peaks (measured mass 7,281.43 Da) matched sequences for two Basic $PLA_2$s in the database, each with a score of 79 (compared to a score of 155 for the main peak matching to a cytotoxin). The matched $PLA_2$s were PA2B3_NAJMO (mass 14,091 Da) and PA2B4_NAJNG (mass 14,057 Da).

### 5.2.2.4 N.nub_i17r2

5.2.2.5 N.nub_i17r2 had only one major peak and no minor peaks. This sample had five cardiotoxin entry matches according to MS peptide mapping matched with the Mascot database (see Appendix IV for details of peptide mapping). These entries – 3SA1_NAJNI, 3SA2_NAJHA, 3SA6_NAJHA, 3SA9_NAJHA and 3SA0_NAJSP – were aligned with the template cardiotoxin structure (Figure 5.2) and the fragments identified by MS (details of peptide mapping found in Appendix IV). This MSA is shown in Figure 5.11.

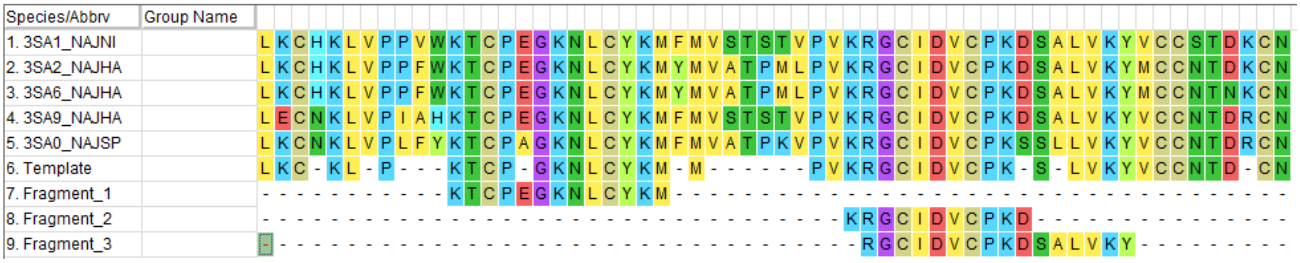



*Figure 5.11. Multiple sequence alignment of N.nub_i17r2 fragments identified by mass spectrometry.*

*The fragment sequences were aligned with the five top scoring hits from the Mascot data search along with the template sequence (Konshina et al., 2017).*

The alignment gave a consensus sequence of:

**L-C-KLVP---KTCPEGKNLCYKM-MV-T---PVKRGCIDVCPKDSALVKY-CC-T--CN**

A BLASTp search was performed for this sequence (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Details of the top five matches originating from *Afronaja* subgenera (according to classification by Wallach et al. (2009)) for this search are shown in Table 5.6.

*Table 5.6. Top 5 results from a BLASTP search against N.nub_i17r2 consensus sequence.*

*The consensus sequence for N.nub_i17r2 was L-C-KLVP---KTCPEGKNLCYKM-MV-T---PVKRGCIDVCPKDSALVKY-CC-T--CN. Any proteins originating from cobras other than Afronaja were omitted.*

| | Protein name | Species | Locus | Uniprot ID | Total score | Query cover | E Value | Per. Ident |
|---|---|---|---|---|---|---|---|---|
| 1 | Cytotoxin 4 | *Naja mossambica* | 3SA4_NAJMO | P01452.1 | 63.5 | 89% | 6e-12 | 71.43% |
| 2 | Cytotoxin 5 | *Naja mossambica* | 3SA5_NAJMO | P25517.3 | 58.5 | 93% | 4e-10 | 68.75% |
| 3 | Naniproin | *Naja nigricollis* | 3SAN_NAJNG | P0DSN1.1 | 58.2 | 95% | 6e-10 | 63.46% |
| 4 | Cytotoxin 3 | *Naja mossambica* | 3SA3_NAJMO | P01470.1 | 58.2 | 89% | 8e-10 | 65.31% |
| 5 | Cytotoxin 1 | *Naja mossambica* | 3SA1_NAJMO | P01467.1 | 56.2 | 89% | 4e-9 | 63.27% |

These top 5 matches were aligned using MEGA7 and are shown in Figure 5.12, along with the template cardiotoxin structure (Figure 5.2) and the peptide fragments.

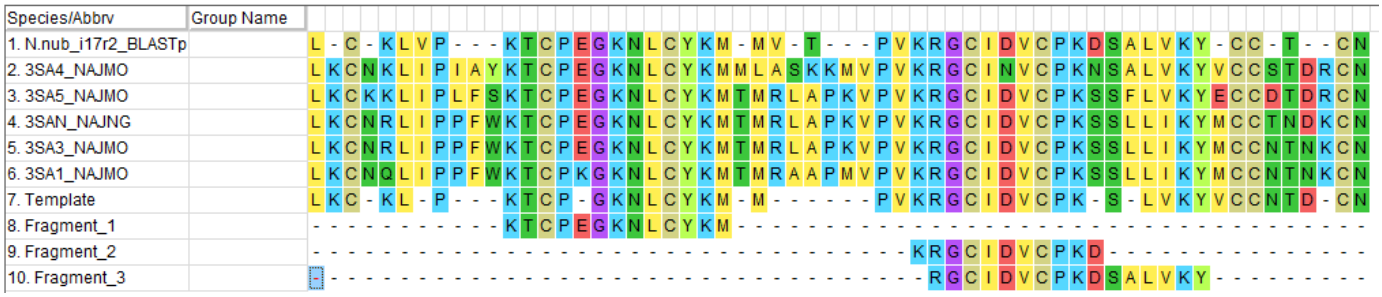



*Figure 5.12. Multiple sequence alignment of the five highest scoring proteins against N.nub_i17r2.*

*The five highest scoring proteins from cobras from the Afronaja subgenera in a multiple sequence alignment with the N.nub_i17r2 sequence put into BLASTP and the template sequence.*

It was not possible to confidently predict the final AA sequence of the N.nub_i17r2 CT. A final consensus sequence was used for further analysis:

**LKC-KLVP---KTCPEGKNLCYKM-MV-T---PVKRGCIDVCPKDSALVKY-CC-T--CN**

### 5.2.3 Database Cardiotoxin Multiple Sequence Alignments

There are many cardiotoxins previously studied and available to view in various databases. SwissProt is a database containing a sub-selection of proteins from the UniProt database which have been manually annotated to contain information extracted from literature sources and contain computational analysis evaluated by curators. SwissProt may be found at the following link: https://www.uniprot.org/uniprot/?query=reviewed:yes.

The SwissProt database contains 25 cobra cardiotoxin entries annotated with $LD_{50}$ values. These entries are detailed in Table 5.7. 24 of these cardiotoxins were aligned by descending potency (Figure 5.13) and grouped by sub-genera (Figure 5.14). AA residues that are conserved in every sequence are denoted by a * at the top of each column. S- and P-types of CTs are indicated with black and red arrows, respectively. SwissProt entry 3SOFB was excluded from this MSA as it has a distinctly different structure compared to the other CTs including mutations at usually conserved positions 6, 22, 24, 34, 39, 41 and 51 and insertions between positions 3-4 and 28-29.

### 5.2.4 Structure Activity Relationships

Structure Activity Relationships (SAR) show the relationship between the structure (AA sequence in the case of proteins) and the biological activity of the compound or protein. The structures of the cobra cytotoxins which have $LD_{50}$ values associated with them in the SwissProt database were compared to determine AAs contributing significantly to the activity of each cytotoxin. Figure 5.15 shows the SAR for all the CTs with associated $LD_{50}$ values **except** 3SOFB which is of low potency (25.2 mg/kg) CT from *Naja annulifera*. The same SAR **including** 3SOFB can be found in Appendix V.

*Table 5.7. Details of the proteins included in the SAR analysis including UniProt code, UniProt accession number, species, group name, $LD_{50}$ value and literature reference that described the protein.* 3SOFB_NAJHA is highlighted in grey as it was not included in the MSA analysis below.

| Code | Accession number | Species | Group name | LD50 (mg/kg) | Reference |
|---|---|---|---|---|---|
| **3SA1_NAJMO** | P01467 | *Naja mossambica* | Afronaja | 0.83 | (Chien *et al.*, 1994) |
| **3SA2_NAJMO** | P01469 | *Naja mossambica* | Afronaja | 1.11 | (Chien *et al.*, 1994) |
| **3SA7A_NAJKA** | P01445 | *Naja kaouthia* | Naja | 1.2 | (Joubert and Taljaard, 1980) |
| **3SA1_NAJME** | P01448 | *Naja melanoleuca* | Boulangerina | 1.37 | (Carlsson and Joubert, 1974) |
| **3SA2_NAJNI** | P01463 | *Naja nivea* | Uraeus | 1.5 | (Botes and Viljoen, 1976) |
| **3SA3_NAJNI** | P01458 | *Naja nivea* | Uraeus | 1.6 | (Botes and Viljoen, 1976) |
| **3SA3_NAJMO** | P01470 | *Naja mossambica* | Afronaja | 1.82 | (Chien *et al.*, 1994) |
| **3SA5_NAJHA** | P01464 | *Naja annulifera* | Uraeus | 1.82 | (Joubert, 1977) |
| **3SA4_NAJMO** | P01452 | *Naja mossambica* | Afronaja | 1.97 | (Chien *et al.*, 1994) |
| **3SA2_NAJHA** | P01462 | *Naja annulifera* | Uraeus | 1.98 | (Joubert, 1977) |
| **3SA2_NAJAT** | P01442 | *Naja atra* | Naja | 2.1 | (Kaneda *et al.*, 1977) |
| **3SA4_NAJAT** | P01443 | *Naja atra* | Naja | 2.1 | (Kaneda *et al.*, 1977) |
| **3SA1_NAJKA** | P0CH80 | *Naja kaouthia* | Naja | 2.5 | (Debnath *et al.*, 2010) |
| **3SA5_NAJHH** | P01457 | *Naja haje haje* | Uraeus | 2.6 | (Joubert and Taljaard, 1978) |
| **3SA8_NAJHA** | P01460 | *Naja annulifera* | Uraeus | 2.6 | (Joubert, 1976b) |
| **3SA7_NAJHA** | P01466 | *Naja annulifera* | Uraeus | 2.61 | (Joubert, 1977) |
| **3SA5_NAJMO** | P25517 | *Naja mossambica* | Afronaja | 2.85 | (Bougis *et al.*, 1983) |
| **3SA1_NAJNI** | P01456 | *Naja nivea* | Uraeus | 2.9 | (Botes and Viljoen, 1976) |
| **3SA3_NAJHA** | P01459 | *Naja annulifera* | Uraeus | 2.98 | (Joubert, 1976a) |
| **3SA1_NAJHA** | P01455 | *Naja annulifera* | Uraeus | 3 | (Weise *et al.*, 1973) |
| **3SA4_NAJHA** | P01461 | *Naja annulifera* | Uraeus | 4 | (Joubert, 1976a) |
| **3SAA_NAJHA** | P01453 | *Naja annulifera* | Uraeus | 4.14 | (Joubert, 1976b) |
| **3SA9_NAJHA** | P01454 | *Naja annulifera* | Uraeus | 4.9 | (Joubert, 1976b) |
| **3SA6_NAJHA** | P01465 | *Naja annulifera* | Uraeus | 5.09 | (Joubert, 1977) |
| **3SOFB_NAJHA** | P62390 | *Naja annulifera* | Uraeus | 25.2 | (Joubert, 1976a) |

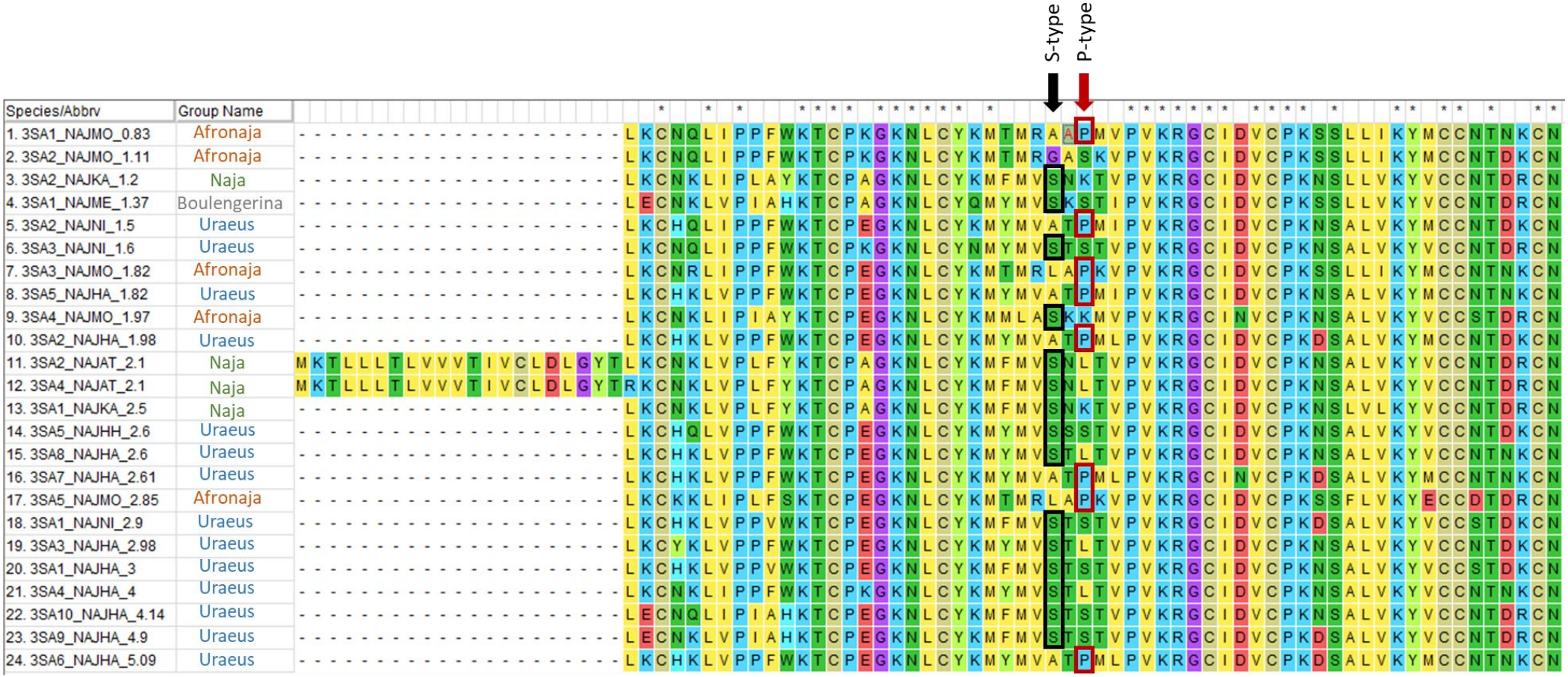



*Figure 5.13. Multiple sequence alignment of 24 cardiotoxins from the Swissprot database that have $LD_{50}$ data available, arranged by potency.*

*The cardiotoxins are listed from the most toxic (lowest $LD_{50}$ value) at the top, descending to the least toxic (highest $LD_{50}$ value) at the bottom. Columns that are strictly conserved have * at the top of the column. Multiple sequence alignment performed using EMBL-EBI Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and imported into MEGA7 where they are displayed here. The $LD_{50}$ value is shown as a value in mg/kg at the end of each Swissprot code in the first column (e.g. 0.83 mg/kg for 3SA1_NAJMO_0.83). S-type CTs (those with a serine AA in position 28) are indicated in black; P-type CTs (those with a proline AA in position 30) are indicated in red.*

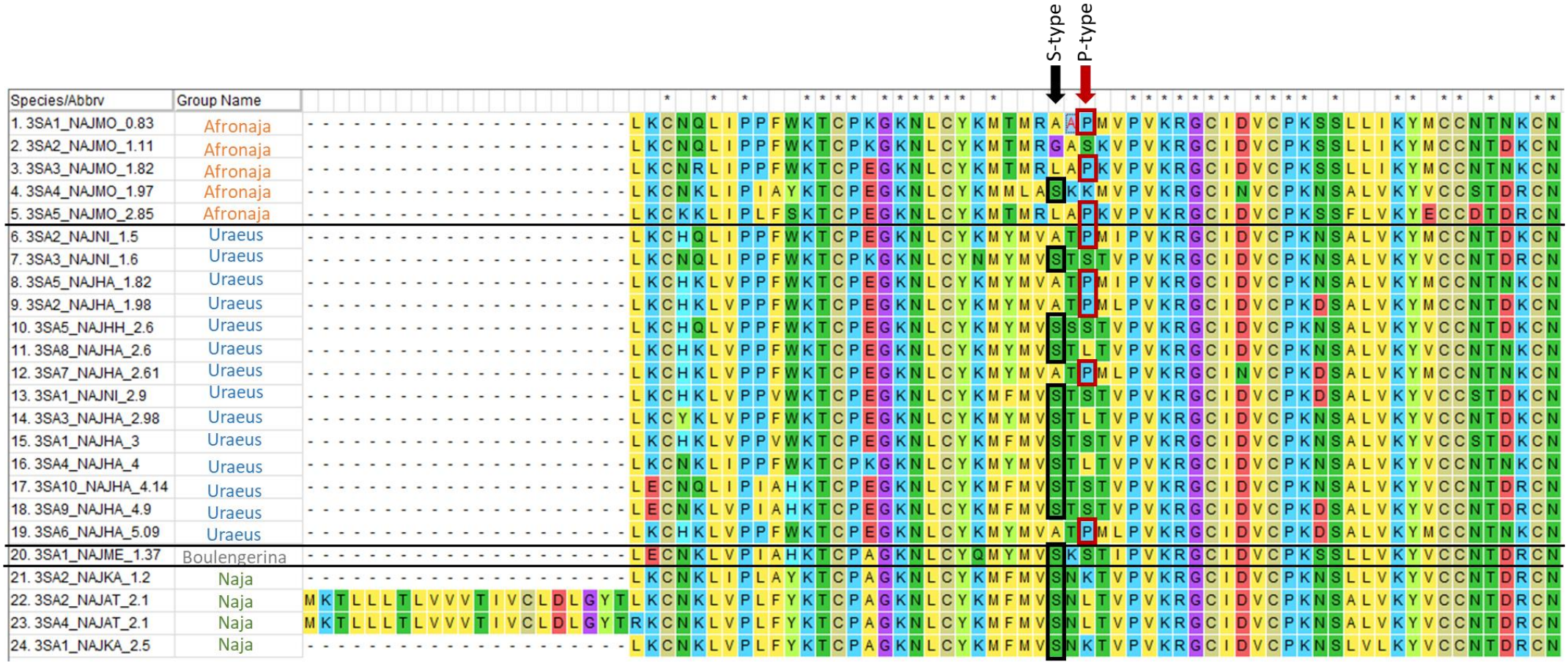



*Figure 5.14.Multiple sequence alignment of 24 cardiotoxins from the Swissprot database that have $LD_{50}$ data available, ordered by subgenera.*

*The cardiotoxins are grouped according to their subgenera (as described by Wallach* et al *(2009)). Columns that are strictly conserved have * at the top of the column. Multiple sequence alignment performed using EMBL-EBI Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and imported into MEGA7. S-type CTs (those with a serine AA in position 28) are indicated in black; P-type CTs (those with a proline AA in position 30) are indicated in red.*

| Code | Cytotoxin | Species | $LD_{50}$ (mg/kg) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3SA1 | 1 | *Naja mossambica* | 0.83 | L | K | C | N | Q | L | I | P | P | F | W | K | T | C | P | K | G | K | N | L | C | Y | K | M | T | M | R | A | A | P | M | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | I | K | Y | M | C | C | N | T | N | K | C | N |
| 3SA2 | 2 | *Naja mossambica* | 1.11 | L | K | C | N | Q | L | I | P | P | F | W | K | T | C | P | K | G | K | N | L | C | Y | K | M | T | M | R | G | A | S | K | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | I | K | Y | M | C | C | N | T | D | K | C | N |
| 3SA2 | 2 | *Naja kaouthia* | 1.2 | L | K | C | N | K | L | I | P | L | A | Y | K | T | C | P | A | G | K | N | L | C | Y | K | M | F | M | V | S | N | K | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | L | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA1 | 1 | *Maja melanoleuca* | 1.37 | L | E | C | N | K | L | V | P | I | A | H | K | T | C | P | A | G | K | N | L | C | Y | Q | M | Y | M | V | S | K | S | T | I | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA2 | 2 | *Naja nivea* | 1.5 | L | K | C | H | Q | L | I | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | T | P | M | I | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | M | C | C | N | T | D | K | C | N |
| 3SA3 | 3 | *Naja nivea* | 1.6 | L | K | C | N | Q | L | I | P | P | F | W | K | T | C | P | K | G | K | N | L | C | Y | N | M | Y | M | V | S | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA3 | 3 | *Naja mossambica* | 1.82 | L | K | C | N | R | L | I | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | T | M | R | L | A | P | K | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | I | K | Y | M | C | C | N | T | N | K | C | N |
| 3SA5 | 5 | *Naja annulifera* | 1.82 | L | K | C | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | T | P | M | I | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | M | C | C | N | T | N | K | C | N |
| 3SA4 | 4 | *Naja mossambica* | 1.97 | L | K | C | N | K | L | I | P | I | A | Y | K | T | C | P | E | G | K | N | L | C | Y | K | M | M | L | A | S | K | K | M | V | P | V | K | R | G | C | I | N | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | S | T | D | R | C | N |
| 3SA2 | 2 | *Naja annulifera* | 1.98 | L | K | C | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | T | P | M | L | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | M | C | C | N | T | D | K | C | N |
| 3SA2 | 2 | *Naja atra* | 2.1 | L | K | C | N | K | L | V | P | L | F | Y | K | T | C | P | A | G | K | N | L | C | Y | K | M | F | M | V | S | N | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA4 | 4 | *Naja atra* | 2.1 | R | K | C | N | K | L | V | P | L | F | Y | K | T | C | P | A | G | K | N | L | C | Y | K | M | F | M | V | S | N | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA1 | 1 | *Naja kaouthia* | 2.5 | L | K | C | N | K | L | V | P | L | F | Y | K | T | C | P | A | G | K | N | L | C | Y | K | M | F | M | V | S | N | K | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | L | V | L | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA5 | 5 | *Naja haje* | 2.6 | L | K | C | H | Q | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | S | S | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | K | C | N |
| 3SA8 | 8 | *Naja annulifera* | 2.6 | L | K | C | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | S | T | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | N | K | C | N |
| 3SA7 | 7 | *Naja annulifera* | 2.61 | L | K | C | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | T | P | M | L | P | V | K | R | G | C | I | N | V | C | P | K | D | S | A | L | V | K | Y | M | C | C | N | T | N | K | C | N |
| 3SA5 | 5 | *Naja mossambica* | 2.85 | L | K | C | K | K | L | I | P | L | F | S | K | T | C | P | E | G | K | N | L | C | Y | K | M | T | M | R | L | A | P | K | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | F | L | V | K | Y | E | C | C | D | T | D | R | C | N |
| 3SA1 | 1 | *Naja nivea* | 2.9 | L | K | C | H | K | L | V | P | P | V | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | F | M | V | S | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | V | C | C | S | T | D | K | C | N |
| 3SA3 | 3 | *Naja annulifera* | 2.98 | L | K | C | Y | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | S | T | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | K | C | N |
| 3SA1 | 1 | *Naja annulifera* | 3 | L | K | C | H | K | L | V | P | P | V | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | F | M | V | S | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | S | T | D | K | C | N |
| 3SA4 | 4 | *Naja annulifera* | 4 | L | K | C | N | K | L | I | P | P | F | W | K | T | C | P | K | G | K | N | L | C | Y | K | M | Y | M | V | S | T | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | N | K | C | N |
| 3SA10 | 10 | *Naja annulifera* | 4.14 | L | E | C | N | Q | L | I | P | I | A | H | K | T | C | P | E | G | K | N | L | C | Y | K | M | F | M | V | S | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA9 | 9 | *Naja annulifera* | 4.9 | L | E | C | N | K | L | V | P | I | A | H | K | T | C | P | E | G | K | N | L | C | Y | K | M | F | M | V | S | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA6 | 6 | *Naja annulifera* | 5.09 | L | K | C | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | T | P | M | L | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | M | C | C | N | T | N | K | C | N |

| Key: | high | 0.8-1 | medium | 0.6-0.8 | low | 0.4-0.6 | very low | 0.2-0.4 | none | 0-0.2 |
|---|---|---|---|---|---|---|---|---|---|---|

*Figure 5.15. Amino acids within the cardiotoxin sequences that are important for the resultant toxicity according to SAR analysis.*

*The low toxicity ($LD_{50}$ value =25.2 (mg/kg)) 3SOFB from Naja annulifera was excluded from SAR analysis. The key shows how important the amino acid in that position in the sequence is to the activity of the protein. Rankings are based on the SADC values and fall between 0-1.*

Many of the cardiotoxins detailed in the Swissprot database also have entries in the protein data bank (PDB) – a database containing the structural assemblies of many biological molecules. The RCSB PDB database (Berman *et al.*, 2000, Burley *et al.*, 2019) may be found at https://www.rcsb.org/. The worldwide PDB (Berman *et al.*, 2003) may be found at http://www.wwpdb.org/.

Cytotoxin 4 from *Naja mossambica* venom was the closest match available according to the Mascot database for three out of four of the analysed venom components (N.pal_i17r2, N.nct_i18r2 and N.nub_i17r2). This cardiotoxin is in the Swissprot database under entry 3SA4_NAJMO and has a PDB ID of 1CDT. This PDB entry (found at https://www.rcsb.org/structure/1CDT) details the refined crystal structure of the homodimer from X-ray crystallography experimental data at 2.5 Å. Modified 3D models of the 1CDT molecule showing the regions identified as contributing significantly to activity according to the CT SAR (Figure 5.15) are shown in Figure 5.16.

The four unknown CT molecules identified by MS had consensus sequences detailed in sections 5.2.2.1 (N.pal_i15r4), 5.2.2.2 (N.pal_i17r2), 5.2.2.3 (N.nct_i18r2) and 5.2.2.4 (N.nub_i17r2). These sequences are aligned in Figure 5.17(a) and had SAR performed on them Figure 5.17(b). Approximate $LD_{50}$ values were calculated for each sequence and are shown in 4.3.8. Unfortunately, if any of the sequences contained gaps with unknown AA residues, that numbered column could not be included in the SAR analysis. Despite this, four AA residues (7, 26, 27 and 29) were identified as having a 'high' probability of having a significant impact on activity and one AA residue (5) was identified as having 'medium' probability of significance.

Only one residue - the AA in position 29 - was identified as being significant in both the literature and the presented study.

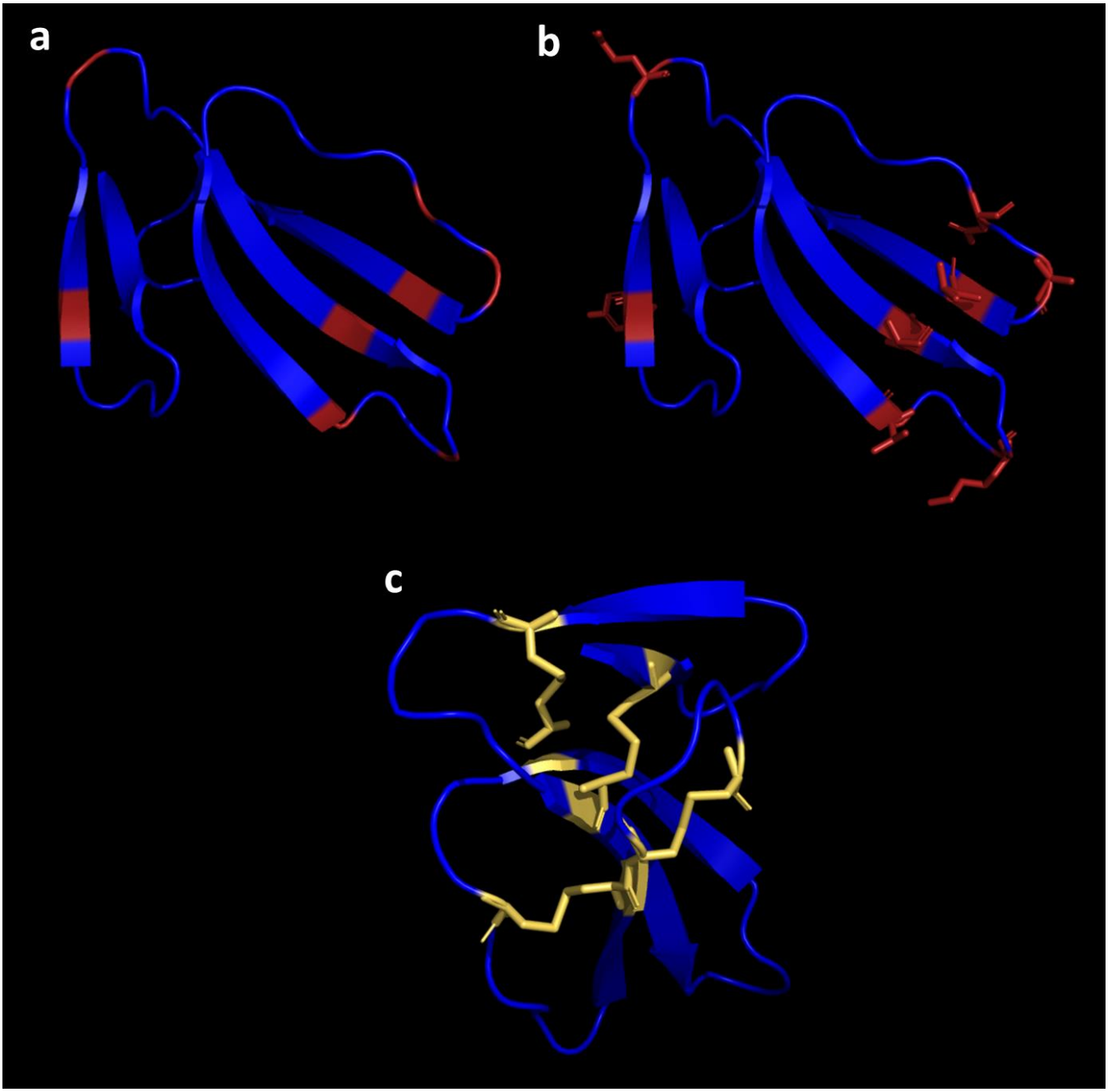



*Figure 5.16. Three dimensional models of cytotoxin 4 from Naja mossambica venom.*

*Models created using The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. (a) Structure of PDB ID 1CDT (Rees et al., 1990) (blue) with amino acid residues determined to be of significance (SADC >0.6) according to literature SAR in section 5.2.4 marked in red. (b) as a. but with significant residues shown as 'sticks' to show how these residues lie in 3D space. (c) Cysteine residues shown as 'sticks' in yellow. These cysteine residues form the four disulphide bonds conserved in all cardiotoxin molecules and provide a large amount of structural stability.*

**a.**

| Species/Abbrv | Group Name | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | * | * | * | | | * | | * | | | | * | * | * | * | | * | * | * | * | * | * | * | * | | | | | | | | | | | * | * | * | * | * | | | * | | * | | | * | * | | | * | | | * | * |
| 1. N.pal_i15r4 | | L | K | C | N | Q | L | I | P | P | F | W | K | T | C | P | - | G | K | N | L | C | Y | K | M | T | M | R | A | A | P | M | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | - | K | Y | M | C | C | N | T | D | K | C | N |
| 2. N.pal_i17r2 | | L | K | C | - | K | L | V | P | - | - | - | K | T | C | P | E | G | K | N | L | C | Y | K | M | - | M | V | - | T | - | - | - | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | - | C | C | - | T | - | - | C | N |
| 3. N.nct_i18r2 | | L | K | C | N | K | L | I | P | I | A | Y | K | T | C | P | - | G | K | N | L | C | Y | K | M | M | L | A | S | K | K | M | V | P | V | K | R | G | C | I | - | V | C | P | K | - | S | - | L | - | K | Y | V | C | C | S | T | D | R | C | N |
| 4. N.nub_i17r2 | | L | K | C | - | K | L | V | P | - | - | - | K | T | C | P | E | G | K | N | L | C | Y | K | M | - | M | V | - | T | - | - | - | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | - | C | C | - | T | - | - | C | N |

**b.**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| >N.pal_i15r4 | L | K | C | N | Q | L | I | P | P | F | W | K | T | C | P | - | G | K | N | L | C | Y | K | M | T | M | R | A | A | P | M | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | - | - | K | Y | M | C | C | N | T | D | K | C | N |
| >N.pal_i17r2 | L | K | C | - | K | L | V | P | - | - | - | K | T | C | P | E | G | K | N | L | C | Y | K | M | - | M | V | - | T | - | - | - | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | - | C | C | - | T | - | - | C | N |
| >N.nct_i18r2 | L | K | C | N | K | L | I | P | I | A | Y | K | T | C | P | - | G | K | N | L | C | Y | K | M | M | L | A | S | K | K | M | V | P | V | K | R | G | C | I | - | V | C | P | K | - | S | - | L | - | K | Y | V | C | C | S | T | D | R | C | N |
| >N.nub_i17r2 | L | K | C | - | K | L | V | P | - | - | - | K | T | C | P | E | G | K | N | L | C | Y | K | M | - | M | V | - | T | - | - | - | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | - | C | C | - | T | - | - | C | N |

| Key: | high | 0.8-1 | medium | 0.6-0.8 | low | 0.4-0.6 | very low | 0.2-0.4 | none | 0-0.2 |
|---|---|---|---|---|---|---|---|---|---|---|

**c.**

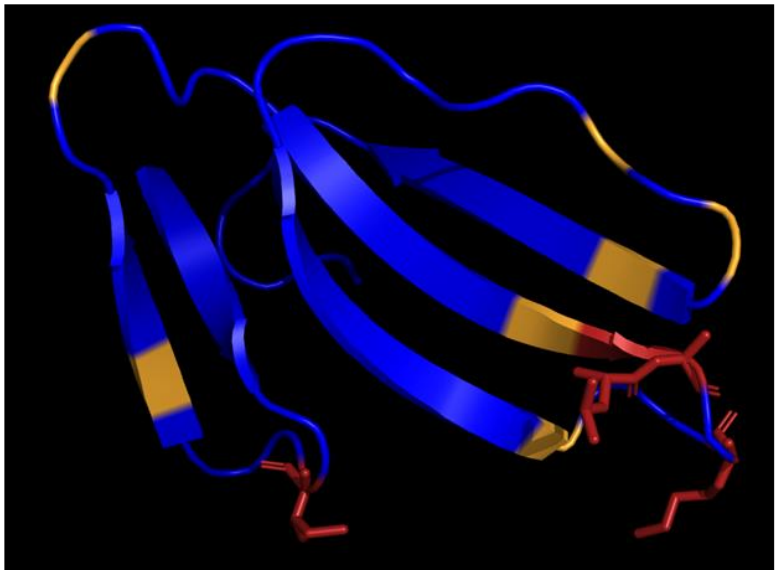


*Figure 5.17. Comparison of the four cardiotoxin sequences identified in this study.*

*(a) MSA of the four cardiotoxin sequences identified by MS analysis using MEGA7. (b) SAR highlighting the most significant AA residues according to SADC value. SAR performed using WebProAnalyst. (c) 3D model of cardiotoxin 4 from Naja mossambica (PDB ID: 1CDT) created using PyMOL. The general regions indicated as potentially significant from the SAR analysis of the database cardiotoxins (as shown in Figure 5.15) are displayed in orange. The residues identified as significant according to the SAR in part b of this figure are displayed in red.*

## 5.3 Discussion

### 5.3.1 Cytotoxins

The cobra venom components discovered to be active in Chapter 4 were identified by mass spectroscopy (MS) analysis using intact mass analysis and peptide mapping techniques. The fragments from all four of the submitted samples - N.pal_i15r4, N.pal_i17r2, N.nct_i18r2 and N.nub_i17r2 - were matched using Mascot against existing peptide sequences from the Swissprot database and were identified as being cardiotoxins (CTs). The identified fragments were aligned with the top Mascot matches and a highly conserved CT scaffold sequence (as described by (Konshina *et al.*, 2017)). The sequences were then searched using BLASTp searches. From what is known of the evolution of cobra (*Naja*) species, African spitting cobra (*Afronaja*) venoms are quite evolutionarily distinct from other subgenera of *Naja* species (Wallach *et al.*, 2009, Panagides *et al.*, 2017). Considering these distinctions, coupled with the large gaps in protein sequence data from MS peptide mapping, it was deemed appropriate when performing BLAST searches to omit any protein sequences from any species other than those in the *Afronaja* subgenera. Unfortunately, none of the CT samples could be confidently identified using the MS data alone as there was not 100% sequence coverage from any of the identified peptide fragments. Sequences as complete as could be confidently assumed were compared along with their $LD_{50}$ values to predict the SAR of the CTs.

CTs have a remarkably stable structure which is highly conserved throughout different species of cobra (Kessler *et al.*, 2017, Reeks *et al.*, 2015). Unlike 3-finger neurotoxins which have positive selection pressures to evolve in tandem with prey receptors, CTs have a less specific mechanism, and their evolution is constrained by negative selection (Sunagar et al., 2013). Local conformational changes in CT sequences can cause large changes in functional activity and affect their structure and dynamic behaviour (Konshina et al., 2017).

CTs can be categorised into two sub-groups - S- and P-type - depending on the presence of Ser-28 or Pro-30 residue in their AA sequence. S-type CTs appear to exhibit a higher depolarising effect against muscle cells than P-type, whilst P-type demonstrate a larger haemolytic activity (Chien *et al.*, 1994). S- or P-type identity does not appear to have a direct impact on $LD_{50}$ values in whole mice, as illustrated in the MSA in section 5.2.3. However, the different types do appear to have different spatial structures upon membrane binding. In P-type CTs, loop II of the three-finger structure remains very similar in aqueous solution compared to in the membrane, whilst S-type CT undergoes a confirmational change upon membrane binding compared to when in aqueous solution (Konshina *et al.*, 2012). The tips of all three 'fingers' of both P- and S-type CTs penetrate into the lipid bilayer of the cell membrane, however, the membrane binding site of S-

type CTs have a reduced hydrophobicity compared to that of P-type, resulting in a stronger interaction between P-type CTs and cell membranes (Dubovskii *et al.*, 2005). This is due to an additional phospholipid binding site located on loop II in addition to the binding region in loop I present in all type of CT (Chien *et al.*, 1994). This could explain why mutations in the loop I and II binding sites identified by SAR on the presented data caused the most significant change in activity of the tested CTs.

The venom proteins found within *Naja mossambica* and *Naja nigricincta* have been found to be over 50% identical, with the majority of proteins having a molecular mass of less than 17 kDa (Katali *et al.*, 2020). The recently classified *Naja ashei* also falls into the African spitting cobra category (Wüster and Broadley, 2007). This species was found to be closely related to *Naja mossambica* and *Naja nigricollis* and contains a very similar catalogue of toxins with an abundance of 3FTxs and $PLA_2$ molecules found in its venom (Hus *et al.*, 2018).

### 5.3.2 Structure Activity Relationship

Multiple sequence alignments (MSA) and structure activity relationship (SAR) analysis were performed on all cobra cytotoxins with entries in the SwissProt database (https://www.uniprot.org/uniprot/?query=reviewed:yes). The $LD_{50}$ values for CTs described in the UniProt database are measured from intravenous, subcutaneous or intraperitoneal injection in mice. This data can be difficult to compare directly as there is a tendency to require a higher dosage when injecting intraperitoneally and especially subcutaneously compared to intravenous administration. In addition, the data originating from the literature could not be directly compared to the data presented in this study as the test models were very different and so not directly comparable. Whilst the literature used $LD_{50}$ values in mg/kg in injected mouse studies, this study presents $IC_{50}$ values (the dose required to inhibit 50% of the population) in mg/ml on human cancer cells in 2D culture. In addition to this difference, the presented study used a specific cell toxicity assay in specific cell lines to measure the $IC_{50}$ value, whilst injection into a mouse model could lead to other whole organism effects from the CTs, such as cardiotoxicity, neurotoxicity or ion channel and organ specific effects.

Despite having an $LD_{50}$ value in the database, SwissProt entry 3SOFB_NAJHH - Cytotoxin 11 from *Naja haje* - was excluded from analysis as it has a distinctly different structure to the other CTs including several insertions and mutations in regions that are conserved in all of the other CT entries. Despite being potentially interesting due to the 10-fold lower potency, this entry would have biased the algorithm used as part of the SAR analysis and made it harder to determine changes in activity caused by more subtle structural changes.

Looking at the SAR analysis from sequences in the literature, the AA regions that had the most significant impact on activity when substituted tended to be towards the loops or 'finger' parts of the three-finger structure. Given that these loops are the regions that interact with cell membranes (Konshina *et al.*, 2012), it is not too surprising that variations in these regions cause changes in potency in whole organism studies. In addition, the other regions of the protein are highly conserved and contribute to the structure and stability of the protein (Kessler *et al.*, 2017).

With regards to the SAR performed on the data presented in this study, the $IC_{50}$ values used for the SAR were generated in Chapter 4. These may not be completely accurate as not all of the curves reached the 'top' of the sigmoidal dose response curve. If it becomes possible in the future to obtain new supplies of fractionated venom, the dose response should be repeated but with a higher top dose such that the full pharmacological response may be measured and a more accurate $IC_{50}$ recorded. This, combined with a different method of protein identification to give a full and accurate protein sequence, would give improved confidence to the SAR.

Unfortunately since none of the cobra venom fractions could be completely identified using the MS techniques applied, there were gaps in the MSA used for the SAR. Due to this, any column with at least one missing AA could not be assigned a SADC value or any significance determined for this part of the sequence. Of the AAs that were identified as contributing significantly to the activity, three out of four were located at the tip on loop II in the three-finger toxin structure and the remaining AA is at the tip of loop I. Only one AA was identified as contributing significantly to activity in both the literature and data presented in this study - an AA in the loop II region. In reality, there are probably more overlaps, but due to the aforementioned reasons of gaps in sequences and disparity of measurement, the two sets of data are not directly comparable.

It should be borne in mine that the Uniprot database is not fully populated and is ever expanding with the addition of new protein sequences. It is possible and indeed likely that some of the proteins described in this chapter are novel and this is the reason they show as incomplete matches to the database, rather than the data being of poor quality. Additional mutations are possible in addition to those currently represented within the current databases. This also explains why it was not possible to accurately estimate the AA sequences for the cardiotoxin sequences which only had two or three 'gaps' of unknown AAs. These gaps could theoretically be filled by any amino acid, and their identification rely heavily on the total mass. Any sequence 'gaps' filled with AAs from the template CT sequence (Figure 5.2) are susceptible to skewing the mass of the predicted sequences if incorrect, which makes prediction of the remaining 'gaps' based on mass futile.

In Chapter 4, it was suggested that N.nub_i17r2 and N.pal_i17r2 could be the same protein due to their similar chromatographic retention times in both IE and RP HPLC and their similar pharmacological profiles. These two samples also only had a mass difference of 0.01 Da according to intact mass analysis. This further supports the idea that these two samples are the same protein or identical homologues.

To build upon the SAR investigation, the significance of each AA in a CT protein sequence could be systematically investigated using a technique such as positional scanning. Positional scanning is used to identify AA residues in a protein sequence that are important for the shape, stability and function of a protein. This involves substituting each AA in a sequence with another amino acid, for example alanine, allowing the importance of a side chain functional group to be assessed (Morrison and Weiss, 2001).

Positional scanning has been utilised in several venom studies, including a study which identified the residues important for blocking $Na_v$ ion channels (Moyer *et al.*, 2018). They used Alanine scanning mutagenesis to discover important residues for blocking the $Na_v1.7$ channel, and Glutamine scanning mutagenesis identified mutations leading to 500-fold selectivity of $Na_v1.7$ over $Na_v1.4$ channel blocking.

Neff *et al* (2020) performed a similar positional scanning mutagenesis experiment with Huwentoxin-IV (HwTx-IV), a cysteine knot peptide from the venom of the theraphosid *Haplopelma schmidti*, to investigate the SAR between the toxin and its activity at $Na_v1.2$ and $Na_v1.7$ ion channels. Substitution of each position with alanine caused a decrease in activity at both channels for every sequence iteration but did highlight the most significant residues being in loop 4 and at the C terminus. The authors then explored substitutions with other amino acids and found varying amounts of decreased potency depending on the position. Although many substitutions increased selectivity for the unwanted $Na_v1.2$, they did find that substitutions in Arg-26 on loop 4 favoured selectivity for $Na_v1.7$ by more than 3-fold.

This technique of positional scanning could be applied to a promising looking CT to determine the residues in the sequence contributing most to the activity, and to potentially improve the potency, selectivity, or other aspect of the peptide. This would, however, require a full and accurate starting AA sequence and recombinantly produced protein from an expression system where the DNA can be sequentially mutated to produce the many AA substituted variations of the protein.

Recombinantly producing CTs in the first place can be a challenge by itself, since these peptides have a complex three-dimensional structure containing several disulphide bridges, requiring a

recombinant cell line able to perform post-translational modifications. This in turn presents a problem as CTs are by their very nature cytotoxic and so their build up within a cell could cause cell injury, leading to apoptosis or damage to the cell membranes of the recombinant cells once the CT concentration rises above a certain concentration. This difficulty could be overcome if it were possible to modify a venom gland expression system, such as the one described by Yamanouye *et al.* (2006). More recently, venom gland organoids have been developed, allowing the secretion of functionally active venom in culture (Post *et al.*, 2020). Although this is a highly welcome advance for the sustainable production of venom for both antivenom studies and studies utilising venom for other medical applications, this method does not yet allow modification of the venom gland DNA such that positional scanning could be performed.

Recombinant venom proteins (with or without modification) could be expressed in bacteria and then post-translational modifications performed later. Successful production of recombinant analogue of cytotoxin I from *Naja oxiana* venom expressed in *Escherichia coli* (*E.coli)* in inclusion bodies has been demonstrated (Shulepko *et al.*, 2017). This recombinant protein was refolded in an L-arginine, urea and glutathione containing folding buffer and showed similar spatial structure in NMR and similar cytotoxic activity against C6 rat glioma cells compared to the wild type venom protein. The recombinantly produced protein was also tested for homogeneity and purity using analytical HPLC, SDS-PAGE gels and MS. Alternatively, high throughput expression of oxidised, correctly folded, disulphide containing recombinant venom peptides also produced in *E.coli* has been demonstrated (Turchetto *et al.*, 2017). The ability to express and correctly fold venom proteins in this manner could give a strong starting point for producing AA substituted protein variants for positional scanning.

As previously mentioned, peptide mapping was unable to provide full sequences for any of the submitted venom peptides. This technique could be repeated with a larger quantity of sample to increase the chances of obtaining fragments covering the full peptide sequence. Alternatively, a different methodology could be applied to determine the unknown sequence. Other method which could potentially be applied include protein X-ray crystallography, nuclear magnetic resonance (NMR), cryogenic electron microscopy (cryoEM) or de novo peptide sequencing.

Protein crystallography is an extremely robust method of determining an atomic level of structural analysis of proteins as well as protein-ligand interactions. The method involves supersaturating a protein solution until nucleation occurs and eventually a crystal forms (Holcomb *et al.*, 2017). The crystal is then mounted in the path of an X-ray beam and a diffraction pattern from the crystal collected (Smyth and Martin, 2000). This diffraction pattern must then be deconvoluted so the protein structure may be ascertained. X-ray crystallography requires

extensive optimisation of crystallisation conditions which can be an extremely time-consuming activity. There is also a requirement for access to a synchrotron or free-electron laser and cryo-storage of crystals until this access is available.

NMR spectroscopy and cryoEM provide information about macromolecular structure and molecular interactions such as distance between interacting molecules, which can include proteins and peptides. Although neither technology produces data quite as high resolution as X-ray crystallography, they can be used as an alternative, for example when crystals fail to form or if access to the appropriate equipment is not possible.

De novo peptide sequencing does not rely on database searching to align amino acid fragments identified by MS, but rather directly identifying amino acids in the protein chain. However, even the leading de novo peptide algorithms are only up to 75% accurate in their amino acid identification (Frank *et al.*, 2007). In addition to this, some amino acids have the same mass (e.g. leucine and isoleucine) and so cannot be reliably distinguished. In this case, database searching could be used in conjunction if the protein of interest is already documented in the used database but has limited application for novel peptides and proteins or those with modifications and mutations.

### 5.3.3 Phospholipase-$A_2$ Contained in Samples

Two of the tested samples (N.pal_i15r4 and N.nct_i18r2) contained minor peaks which matched as $PLA_2$ enzymes in the Uniprot database. It could be that these venom fraction samples (N.pal_i15r4 and N.nct_i18r2) contained a low level of Phospholipase $A_2$ contamination, or that the cytotoxin and $PLA_2$ subunits have formed a synergistic complex (Doley and Kini, 2009, Pucca *et al.*, 2020).

$PLA_2$ contributes between 4-31% of protein content of cobra venoms. There is an increased abundance of $PLA_2$ enzymes in all spitting lineages (African spitting, Asian spitting and *H.haemachatus*) compared to non-spitting cobras. This is thought to be due to a duplication of the $PLA_2$ gene in the ancestor of African Spitting cobras concurrently with the venom spitting trait (Kazandjian *et al.*, 2021). There is significantly less $PLA_2$ content in the venom of cobras from the *Uraeus* (savanna dwelling non-spitting African) sub-genus compared to the *Naja* (Asian) and *Afronaja* (African spitting) subgenera and particularly *Boulengerina* (African forest-dwelling) species (Tan *et al.*, 2019).

It has long been documented that $PLA_2$s are able to coelute with cardiotoxins during ion exchange chromatography and that this combination can act synergistically to increase the observed toxicity (Hodges *et al.*, 1987). IV injection of $PLA_2$ in mice followed 15 minutes later by

cardiotoxin injection significantly decreases the dose required for lethality to as little as 18% of the $LD_{50}$ (Bougis *et al.*, 1987). One theory suggests that $PLA_2$s act pre-synaptically and 3FTxs exert their effects post-synaptically, leading to increased cytotoxicity (Doley and Kini, 2009). The effects of CTXs have been shown to be potentiated by $PLA_2$ and stimulation of sensory neurons by a mixture of both CTXs and $PLA_2$ is enhanced compared to stimulation by CTXs alone (Kazandjian *et al.*, 2021). These observations may explain why the two venoms with potential $PLA_2$ contamination show the lowest $IC_{50}$ values in Chapter 4.

Cobra CTs have been repeatedly shown to target and interfere with mitochondrial membrane structure, function and integrity (Zhang *et al.*, 2019a, Gasanov *et al.*, 2015, Dubovskii *et al.*, 2014). CTs have less polar AA residues compared to neurotoxins and this may allow CTs to bind to membranes by interacting with the phosphatidylserine head groups of lipids (Konshina *et al.*, 2011, Dubovskii and Utkin, 2015) and that P-type CTs interact more strongly with membranes compared to S-type, although their mechanism of activity is similar for both (Dubovskii *et al.*, 2005). Local conformational changes can cause large changes in functional activity and affect membrane binding as well as the structure and dynamic behaviour of CTs (Konshina *et al.*, 2017). Lysosomes appear to be another primary target for CTs from various *Naja* species and their accumulation there correlates with their cytotoxic effects (Feofanov *et al.*, 2005). Cytotoxins also appear to be responsible for the ocular effects commonly documented in spitting cobra species (Ismail *et al.*, 1993).

### 5.3.4 Chapter Conclusion

MS was used to identify the fractions of interest - N.pal_i15r4, N.pal_i17r2, N.nct_i18r2 and N.nub_i17r2 - from Chapter 4. All these fractions were identified as having closest database matches to a variety of cobra cytotoxins with N.pal_i15r4 and N.nct_i18r2 also having a minor peak matching to cobra $PLA_2$ sequences identified in the sample. In the following chapter, the effect of these cytotoxins on a variety of cancer related genes will be investigated.

# Chapter 6 - Investigating the Effect of Cobra Cytotoxins on Gene Expression in Cancer Cells

## 6.1 Introduction

The purpose of this chapter is to explore the effect of the cardiotoxins identified in chapter 5 on relevant cancer genes. The first section demonstrates the optimal parameters for assessing gene expression by qPCR with reference to the minimum information for publication of quantitative real-time PCR experiments (MIQE) guidelines. The second section of this chapter demonstrates which of the selected genes are being influenced by the tested cytotoxic venom components.

### 6.1.1 Real Time Quantitative Polymerase Chain Reaction (qPCR)

In order to gain an understanding of what effect the cytotoxins identified in chapter 5 were having within the cell, real-time quantitative polymerase chain reaction (qPCR) experiments were performed.

qPCR is superior to standard PCR reactions as it is faster and has much higher sensitivity than regular PCR. qPCR also allows detection of small quantities of DNA, even in the presence of dominant populations (Postollec *et al.*, 2011). qPCR may be used to quantify DNA samples or combined with reverse transcription (RT) to provide accurate estimates of quantity of transcribed mRNA. Fluorescently labelled nucleotides are incorporated into the amplification product (amplicon) and fluorescence is measured after every cycle, allowing an amplification curve to be plotted. A fluorescence threshold is chosen at a level where gene products may be distinguished from background fluorescence. From this, the quantification cycle (Cq) number - the quantity of cycles required for a sample to reach the fluorescence threshold - may be calculated for each sample. The Cq is relative to the template DNA in the reaction and from this either the absolute or relative quantity of DNA in the original sample may be calculated.

To produce good quality and meaningful qPCR results, it is imperative that primer pairs for the chosen genes are designed carefully and quality controlled thoroughly prior to commencing the main qPCR experiments. In 2009, the MIQE guidelines were published (Bustin *et al.*, 2009). These guidelines set out the minimum information required when designing and evaluating qPCR experiments to ensure high quality, repeatable, publication quality data are generated. These guidelines have been observed whilst designing and undertaking this study.

In this study, qPCR was used to monitor the effect of cell treatment with cytotoxins on mRNA transcription levels of key cancer genes compared to untreated control cells of the same type.

### 6.1.2 Choice of Genes Included in the Study

Nineteen cancer related genes were selected to address the research question and assess some of the effects cytotoxins may be having on cancer cells.

It must be noted at this point that genes were selected based on the belief that there were two PC cell lines in the presented study. qPCR optimisation experiments and most qPCR experiments were undertaken prior to cell authentication confirming that the intended Mia PaCa2 PC cell line was in fact the SW620 colorectal cell line. Unfortunately, this meant that some of the selected genes were not as relevant to the tested cell lines as intended when they were chosen. It is hoped that some value can still be taken from this chapter as the expression of some genes were still determined to be significantly altered following addition of the cytotoxin samples.

Various forms of cancer may be caused by a combination of the activation of oncogenes and/or inhibition of tumour suppressor genes.

There are several tumour suppressor genes and oncogenes implicated in the development of PC. In terms of tumour suppressor genes, 70% of PDAC tumours contain mutations in Tumour Protein p53 (TP53), 55% in SMAD family member 4 (SMAD4) and 90% in cyclin-dependent kinase inhibitor 2A (CDKN2A) (Liu *et al.*, 2016, Tuveson and Neoptolemos, 2012). Simultaneously, over 90% of metastasising PDAC tumours contain mutations in Kirsten rat sarcoma (KRAS), which lead it to act as a powerful oncogene impacting the majority of early pancreatic lesions as well as later PDAC development (Liu *et al.*, 2016, Tuveson and Neoptolemos, 2012). Extensive literature exists for a multitude of other important genes that are involved in the development, progression and prognosis of PC. A total of 18 (19 including NOTCH1) 'test' genes were selected to give a variety of different mechanisms that have relevance in cancer, particularly pancreatic cancer. These genes are discussed in more detail below.

There is an overlap between some of the key genes involved in colorectal cancer (CRC) and those involved in PDAC, including CDKN2A, Tp53 and KRAS (Yurgelun *et al.*, 2017).

In addition, four reference (also known as 'housekeeping') genes were selected to act as control genes for quality control and normalisation purposes. Reference genes are genes which are stably transcribed, and their transcription levels are not normally affected by the addition of compounds that may be being assessed in the study. 18S ribosomal RNA (18S), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), β-Actin (Actin) and α-tubulin (Tubulin) were selected as reference genes for this study as they are well documented reference genes which are normally reported as being stably transcribed (Hu *et al.*, 2018, Pereira-Fantini *et al.*, 2016, Tang *et al.*, 2017).

#### 6.1.2.1 KRAS

The Kirsten rat sarcoma (KRAS) oncogene is one of the most commonly mutated forms of RAS and is heavily implicated in the development of cancers with particularly high mortality rates including PC, CRC and lung cancer (Waters and Der, 2018). KRAS is connected to many signalling pathways and its overexpression results in an increase in cell survival and proliferation. KRAS encodes a small Guanosine triphosphatase (GTPase) and mutation in the KRAS gene causes permanent binding of this GTPase to GTP, activating multiple downstream signalling pathways including phosphoinositide 3 kinase Akt mammalian target of rapamycin (PI3K-AKTmTOR) (Bournet *et al.*, 2016). Activated KRAS also binds to and activates the RAF kinase family which go on to phosphorylate MEK1 and MEK2 kinases, which in turn activate ERK1 and ERK2 kinases (Cicenas *et al.*, 2017). This signalling cascade eventually ends in cell proliferation through activation of transcription factors ELK1 and c-Jun which are activated by phosphorylated ERK kinases (Cicenas *et al.*, 2017). KRAS mutations may also cause aberrant activation of the RAS/mitogen-activated protein kinase (MAPK) pathway and lead to metastasis, changes to the tumour microenvironment, increased cell survival and proliferation and increased invasive potential (Aubert *et al.*, 2020). There is also a link between KRAS and transforming growth-factor beta (TGF-β) with TGF-β being activated to provoke tumour invasiveness and metastasis by mutated forms of KRAS with this pathway also having relevance in CRC (Poulin and Haigis, 2017).

PC causes high glucose uptake and metabolism with elevated glycolytic rates (the Warburg effect) (Chen *et al.*, 2016). This may in part be due to the KRAS oncogene which regulates enzymes which accelerate the use of glutamine (Falasca *et al.*, 2016). Mitochondrial activity also poses an interesting additional target to limit the progression of PC due to the apparent dependence of some tumours on mitochondrial energy production (Falasca *et al.*, 2016).

KRAS is implicated in more than 20% of cancers including colorectal and lung cancers with over 90% of metastasising PDAC cases containing mutations in KRAS oncogenes (Tuveson and Neoptolemos, 2012, Cicenas *et al.*, 2017). 95% of PCs contain activating KRAS mutations, 99% of which are attributed to a point mutation on codon G12 (50% specifically G12D) of the KRAS oncogene (Cicenas *et al.*, 2017).

With regard to this study, the BxPC-3 cell line was determined to be wild-type for KRAS mutations (Deer *et al.*, 2010) whilst SW620 cells have an activating KRAS mutation (Li *et al.*, 2014).

#### 6.1.2.2 Tp53

Tumour protein 53 (Tp53) is a tumour suppressor gene encoding the p53 protein and acts as a transcription factor with a significant role in regulating a multitude of genes involved in DNA repair, apoptosis and senescence, amongst other functions (Duffy *et al.*, 2020). Wild type Tp53 acts as an effective tumour suppressor gene having a central role in promoting apoptosis following DNA damage (Hafner *et al.*, 2019).

Mutations present in some forms of PC cause loss of this tumour suppression function and can even acquire tumour-promoting functions (Stojanovic *et al.*, 2017). Mutation in Tp53 is common in PC with approximately 70% PDAC tumours containing a Tp53 mutation (Tuveson and Neoptolemos, 2012). 50-75% of these Tp53 mutations result from an initial mutation in the KRAS gene, usually causing expression of a $p53^{R175H}$ mutant form (as opposed to loss of protein expression) which has been shown to promote metastasis in PDAC (Morton *et al.*, 2010).

Tp53 point mutations or deletion of a single copy of Tp53 in combination with KRAS activation is enough to induce PDAC and to stimulate progression of pre-invasive lesions to invasive PDAC (Mello *et al.*, 2017). This highlights the significance of Tp53's contribution to tumour suppression. Tp53 also has a role in the development of CRC and induction of Tp53 expression by drug molecules can encourage apoptosis in CRC cells (Gong *et al.*, 2017).

SW620 cells are Tp53 inactivated (Liu *et al.*, 2018) and BxPC-3 cells have a p53Y220C mutation which leads to expression of dysfunctional p53 protein which has compromised thermal stability (Malhotra *et al.*, 2021).

#### 6.1.2.3 SMAD4 and TGF-β

Transforming growth-factor β (TGF-β) has a significant role in the regulation of cell proliferation and apoptosis and can play a role in tumorigenesis by stimulating angiogenesis and metastasis (Zhang *et al.*, 2019b). TGF-β contributes to signalling by activating SMAD transcription factors which act as tumour suppressors, however the inactivation of SMAD4 in PDAC means that this tumour suppression function is compromised (Ahmed *et al.*, 2017). The loss of this function of TGF-β is one of the characteristics of many cancers (Zhang *et al.*, 2019b) and SMAD4 gene is commonly mutated or deleted in PDAC (Yasutome *et al.*, 2005). Legendre *et al.* (2014) report that it is in fact translocation of the SMAD4 gene, leading to the production of a truncated version, rather than a deletion that may account for PC cells' lack of response to TGF-β.

Inactivation of the SMAD4 gene is a common mutation in PC and is seen in more than 50% of cases of PDAC (Ahmed *et al.*, 2017) and loss of heterozygosity can be seen in approximately 60% of PCs (Zhao *et al.*, 2018). Prognosis and life expectancy of PC patients whose tumour expressed SMAD4 protein is significantly better than patients whose tumours are SMAD4 protein deficient

(Liu, 2001). It should also be noted, however, that SMAD4 abnormalities appear to occur more often in PC cell lines than in primary adenocarcinomas (Bartsch *et al.*, 1999).

SMAD4 gene loss alone does not encourage tumour growth, however when combined with other mutations such as KRAS, p53 or p16 (in PC), or the inactivation of APC (in CRC), tumour growth is promoted (Zhao *et al.*, 2018, Izeradjene *et al.*, 2007). The TGF-β/SMAD4 signalling pathway is closely linked with a variety of other cell signalling pathways including PI3K/AKT, MAPK and the WNT pathway and can be regulated through these other pathways, leading to further complexities (Zhao *et al.*, 2018). Reduction of Pten in tandem with SMAD4 inactivation increases cell proliferation and activates the Notch1 signalling pathway (Xu *et al.*, 2010). 100% of pancreatic tumours appear to have some alterations in TGF-β signalling, including those related to SMAD3 and SMAD4 (Jones *et al.*, 2008).

In terms of the cell lines used in this study, BxPC-3 cells contain homozygous deletions in the SMAD4 gene leading to lack of production of SMAD4 protein (Deer *et al.*, 2010). The SW620 cell line is also deficient in the SMAD4 gene and has been used in other studies as a SMAD4-null control cell line, although it does still have some intrinsic expression (Yan *et al.*, 2018b).

#### *6.1.2.4 CDKN2A/p16INK4a*

Part of the mammalian cell cycle process is regulated by cyclin proteins which depend on cyclin-dependent kinases (CDKs) (Nabel, 2002). CDKs may be inactivated by CDK inhibitor proteins (CDIs) which can act as tumour suppressors to prevent cell cycle progression (Zhao *et al.*, 2016). The CDK inhibitor 2A (CDKN2A) gene encodes for the $P16^{INK4a}$ CDI protein responsible for inhibiting cell cycle phase transition G1/S (Cicenas *et al.*, 2017). When genes encoding CDKs are lost due to factors such as deletion, inactivating mutations or epigenetic silencing, this tumour supressing function is lost (Zhao *et al.*, 2016). Expression of $P16^{INK4a}$ may be lost due to CDKN2A gene hypermethylation of the promotor region or loss of heterozygosity (Bian *et al.*, 2002). Between 14-25% of aberrant expression of CDKN2A has been attributed to promotor hypermethylation (El-Naggar *et al.*, 1997, Jiao *et al.*, 2007). CDKN2A is also associated with CRC gene silencing and promotor hypermethylation being key mechanisms of its downregulation (Shima *et al.*, 2011).

Patients who go on to develop PCs often have mutations in the CDKN2A gene. These mutations lead to an increased incidence and somatic mutations may be found in up to 95% of pancreatic tumours (Mcwilliams *et al.*, 2011, Attri *et al.*, 2005). There may be homozygous deletions or inactivation of the gene, leading to loss of activity (Caldas *et al.*, 1994) and deficiency of CDKN2A is one of the most frequent genetic changes leading to the development of PC (Deer *et al.*, 2010).

BxPC-3 appears to be wild-type for the CDKN2A gene, but does not produce detectable $P16^{INK4a}$ protein product (Deer *et al.*, 2010, Bruno *et al.*, 2017). These cells contain a homozygous deletion on exon 2-3 in the CDKN2A gene. SW620 cells contain full methylation on the CDKN2A promotor, repressing gene transcription completely.

#### *6.1.2.5 MMP-9*

The matrix metalloproteases (MMPs) are zinc-dependant endopeptidases that degrade components of the extracellular matrix (ECM) and have a role in the invasion of tumours, playing a key role in the metastatic cascade (Murray *et al.*, 2004). MMP-9 can act as a biomarker for cancer and can contribute to invasion, metastasis and angiogenesis (Huang, 2018). MMP9 is specifically implicated in PC showing significantly higher serum levels than that of healthy controls (Tian *et al.*, 2008).

MMPs are also commonly implicated in cell invasiveness in CRC with MMP-1, -2, -7, -9 and -13 giving the worst patient outcomes and MMP-12 having a protective effect (Said *et al.*, 2014). Overexpression of MMP-9 correlates with metastasis and an associated reduced survival time in CRC (Yang *et al.*, 2014), although it should be noted that the SW620 CRC cells used in this study do not express measurable levels of MMP9 (Murray *et al.*, 2004).

#### *6.1.2.6 ERBb Gene Family (EGFR, HER2, HER3, HER4)*

A study was conducted by Lee *et al.* (2007) looking at EGFR mutations in PDAC. The actual number of EGFR mutations discovered were extremely low (one patient out of 66, 1.5%). This combined with other data suggest a very low incidence of EGFR mutation mediated PDAC. The amount of mutations found in EGFR were much lower than some other cancers including lung cancer (59%), cholangiocarcinoma (14%) and CRC (12%). The proportion of increased EGFR copy number (41%) was also lower than other solid tumours and presence of additional copies of EGFR did not seem to impact PC patient survival significantly.

BxPC-3 cells show relatively high expression of EGFR receptors with approximately 2.03 x $10^5$ EGFR receptors/cell whereas their expression of HER2 is low at 1.1 x $10^4$ receptors per cell (Larbouret *et al.*, 2012, Mozzi *et al.*, 2015). SW620 on the other hand have an almost undetectable level of EGFR receptors. They are essentially EGFR negative (Yang *et al.*, 2004). They also have approximately 50% the expression level of HER2 compared to BxPC-3 cells (Rusnak *et al.*, 2007). Rusnak *et al*. (2007) compared the EGFR protein expression as a percentage expression of the highly overexpressing cell line HN5 and HER2 expression levels to the BT474 HER2 overexpressing cell line. BxPC-3 cells expressed 15.8% and SW620 expressed 0.34% the number of receptors compared to HN5 cells. HER2 protein expression as a percentage

expression of HN5 HER2: BxPC-3 expressed 2.89% of the HER2 receptors of BT474 cells and SW620 expressed 1.57% (Rusnak *et al.*, 2007).

#### *6.1.2.7 STAT3*

Members of the signal transducer and activator of transcription (STAT) family act as transcription factors following phosphorylation by Janus family kinases (JAKs) (Cowan *et al.*, 2015). STAT3 in particular is implicated in PC and phosphorylated STAT3 is found in increased amounts in PDAC ducts compared to normal pancreatic ducts (Cowan *et al.*, 2015). STAT3 plays an important role in metastasis in colorectal cancer (Zhang *et al.*, 2018)

#### *6.1.2.8 WNT/NOTCH*

The WNT signalling pathway has an important function in gene regulation having roles in cell differentiation, cell proliferation, migration and the progression of various cancers including PC and CRC (Cheng *et al.*, 2019).

WNT/β-catenin signalling has been identified as one of the key 12 signalling pathways associated with PC. All pancreatic tumours appear to have at least some alterations in the WNT/NOTCH signalling pathway (Jones *et al.*, 2008), however WNT/β-catenin signalling is not required for growth and survival in the BxPC-3 cell line and these cells show very low expression of WNT activity when unstimulated (Olsen *et al.*, 2014).

Almost all CRC show increased activation of the WNT pathway and is a key driver in the initiation and progression of the disease (Schatoff *et al.*, 2017).

#### *6.1.2.9 Protein kinase B (AKT1)*

Protein kinase B (Akt1) is a serine/threonine-protein kinase and is a member of the phosphoinositide 3 kinase Akt mammalian target of rapamycin (PI3K-AKTmTOR) signalling pathway which is frequently responsible for cell survival, growth and angiogenesis in cancers including PC (Xu *et al.*, 2018b, Kishor Roy *et al.*, 2017). Hyperactivation of Akt1 is thought to be a major mechanism through which PC forms and presents as an interesting potential pharmacological target (Si *et al.*, 2017, Albury *et al.*, 2015). The PI3K/AKT signalling pathway is also important in colorectal cancer (CRC) and its activation contributes significantly to CRC progression (Wei *et al.*, 2017).

#### *6.1.2.10 Bcl-2/BAX/BAD*

B-cell Lymphoma 2 (Bcl-2), Bcl-2-associated X protein (BAX) and Bcl-2 antagonist of cell death (BAD) are members of the Bcl-2 protein family having involvement in mitochondrial apoptosis regulation (Yang *et al.*, 2020a). Bcl-2 itself is an anti-apoptotic protein and as such its overexpression or upregulation can stimulate tumour growth by preventing apoptotic activities on tumour cells. BAX possesses a proapoptotic function by forming pores in the outer

mitochondrial membrane whilst BAD acts as a sensitiser which encourages apoptosis by supressed anti-apoptotic protein function and can activate the 'death machinery' in mitochondria (Yan *et al.*, 2018a, Yang *et al.*, 2020a).

#### *6.1.2.11 Caspases*

Apoptosis has long been known to have an important, programmed role in the normal regulation of animal cells (Kerr *et al.*, 1972) with evasion of apoptosis contributing to avoidance of cell death, one of the "hallmarks of cancer" (Hanahan and Weinberg, 2011). Caspases are enzymes involved in initiating and performing apoptosis by cleaving proteins important for the structure and regulation of the cell (Boice and Bouchier-Hayes, 2020). Altered caspase function can have consequences in a variety of diseases including cancer with dysregulated caspase activity known to contribute to metastatic invasion, excessive proliferation and evasion of apoptosis, amongst other tumorigenic processes (Xu *et al.*, 2018a). Caspases 3 and 7 are members of the executioner caspase group (Boice and Bouchier-Hayes, 2020) and are required for amplification of upstream apoptotic pathways. The alteration of these caspases, therefore, has significant consequences for apoptotic cell death in cancer cells (Mccomb *et al.*, 2019).

### 6.1.3 Chapter 6 Aims

The aims of this chapter were to:

- design appropriate primer pairs for genes of interest and validate reference gene primers
- fulfil the qPCR criteria outlined in the minimum information for publication of quantitative real-time PCR experiments (MIQE) guidelines
- assess the effect of the cytotoxins identified in Chapter 5 on relevant cancer genes

## 6.2 Results

### 6.2.1 qPCR Optimisation

#### *6.2.1.1 Investigation of primer design to assess suitability for qPCR*

Following their design, primers required testing to ensure they produced the correct amplicon, did not amplify unintended products, provided a reasonable Cq value with low variation and did not create primer dimers. Figure 6.1 shows results of agarose gel analysis of primer pair suitability. The results are summarised in Table 6.1. The HER4 gene was neither matching the anticipated amplicon size, nor amplifying a single band and so was excluded from the study. The CDNK2A primer pair could not be verified as it did not produce a visible agarose gel DNA band in either cell line (Figure 6.1g) and was therefore excluded from the study. Amplicon sequences and details for genes of interest can be found in section 2.10.1.3.

*Table 6.1. Summary of primer suitability for downstream qPCR.*

*Based on whether the cDNA band shown on the agarose gels (Figure 6.1) matched the expected size based on the expected amplicon and whether there is amplification in addition to the intended amplicon (i.e. are there multiple bands present?). Green colouration indicates the primer pair is suitable and is amplifying the target amplicon. Orange indicates a potential problem with the primer pairs due to incorrect sized amplicon or the presence of additional bands. n/a indicates that there were no bands present to analyse on the agarose gel. Genes highlighted in grey were excluded from analysis.*

| Gene | Expected size (bp) | Matches expected size | | Single band | |
|---|---|---|---|---|---|
| | | BxPC-3 | SW620 | BxPC-3 | SW620 |
| GAPDH | 87 | ✔ | ✔ | ✔ | ✔ |
| 18S | 83 | ✔ | ✔ | ✔ | ✔ |
| Actin | 140 | ✔ | ✔ | ✔ | ✔ |
| Tubulin | 144 | ✔ | ✔ | ✔ | ✔ |
| EGFR | 106 | ✔ | ✘ | ✔ | ✔ |
| HER2 | 74 | ✔ | ✔ | ✔ | ✘ |
| HER3 | 134 | ✔ | ✔ | ✘ | ✘ |
| HER4 | 64 | ✘ | ✘ | ✘ | ✘ |
| AKT1 | 84 | ✔ | ✔ | ✔ | ✔ |
| KRAS | 118 | ✔ | ✔ | ✔ | ✔ |
| SMAD4 | 108 | ✘ | ✔ | ✔ | ✔ |
| TGFB | 125 | ✔ | ✔ | ✘ | ✘ |
| TP53 | 145 | ✔ | ✔ | ✘ | ✘ |
| MAPK14 | 170 | ✔ | ✔ | ✔ | ✔ |
| Caspase 3 | 191 | ✔ | ✔ | ✔ | ✔ |
| Caspase 7 | 137 | ✔ | ✔ | ✔ | ✔ |
| MMP9 | 79 | ✔ | n/a | ✔ | n/a |
| STAT3 | 73 | ✔ | ✔ | ✔ | ✔ |
| BAX | 148 | ✔ | ✔ | ✔ | ✔ |
| BAD | 128 | ✔ | ✔ | ✔ | ✔ |
| BCL-2 | 113 | ✔ | ✔ | ✔ | ✔ |
| WNT1 | 85 | ✔ | n/a | ✔ | n/a |
| NOTCH1 | 193 | ✔ | ✔ | ✔ | ✔ |
| CDKN2A | 139 | n/a | n/a | n/a | n/a |

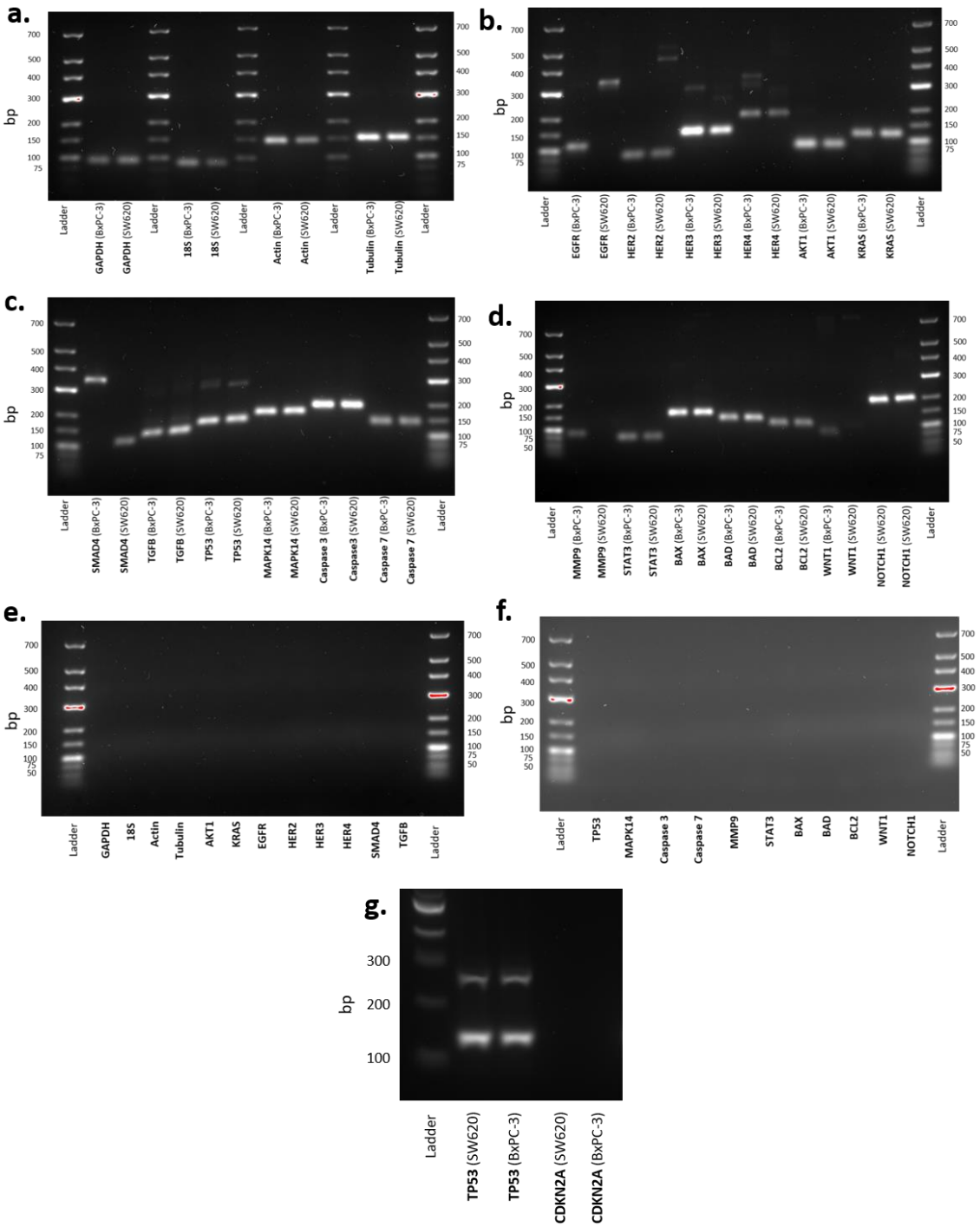



*Figure 6.1. 2% Agarose gels for each gene in the gene study*

*Gels confirm amplicon size (bp) is as expected from the primer design stage. a) shows the reference genes (GAPDH, 18S, β-Actin and α-Tubulin) from BxPC-3 6h cDNA and SW620 24h cDNA. b)-d) show the amplicons from the test genes from BxPC-3 6h cDNA and SW620 6h cDNA. e) and f) show clear gels from the NTC samples from SW620 plates. The BxPC-3 samples had equivalent clear NTC gels (not shown). g) gel to show the lack of bands following attempted amplification of the CDKN2A gene in both cell lines (TP53 bands included to show the gel ran successfully).*

#### *6.2.1.2 Investigating the effect of miniaturising from 24- to 96-well on the RNA purification samples on RNA yield, RNA quality and amplification efficiency*

To determine the feasibility of this scaling down to 96-well format, cDNA was reverse transcribed from RNA harvested from SW620 cells plated into 24- and 96-well plates. The concentration and quality (represented by the A260/280 and A260/230 ratios) of the RNA and cDNA are shown in Table 6.2. The spectral patterns for the harvested RNA and cDNA are shown in Figure 6.2. The A260/230 ratio was low for the RNA purified from the 96-well plate cells. All other ratios are within the expected range. (A260/280 ratio of ~2.0 for RNA and ~1.8 for DNA and A260/230 ratio of between 2-2.2 for pure nucleic acids.) The 96-well plate RNA concentration was measured at 37.1 ng/µl rather than the anticipated ~50 ng/µl which would be predicted from the 10X seeding density of the 24-well plate. This potentially inaccurate reading, combined with the low A260/230 ratio may have led to more RNA being reverse transcribed than intended, explaining the higher concentration of cDNA following reverse transcription. Spectral patterns for the buffers supplied with the QIAGEN RNeasy Mini Kit are shown in Figure 6.2.

*Table 6.2. Summary of concentration and quality of RNA and cDNA from harvested SW620 cells from 24 and 96-well plates.*

*Values in red lies outside the normal range (A260/280 ratio of ~2.0 for RNA and ~1.8 for DNA and A260/230 ratios of 2-2.2 for pure nucleic acids).*

| Sample | Type of nucleic acid | Concentration (ng/µl) | A260/280 | A260/230 |
|---|---|---|---|---|
| 24-well | RNA | 498.0 | 2.10 | 2.08 |
| 96-well | RNA | 37.1 | 2.01 | 0.44 |
| 24-well | cDNA | 205.48 | 1.81 | 1.91 |
| 96-well | cDNA | 619.58 | 1.85 | 1.98 |

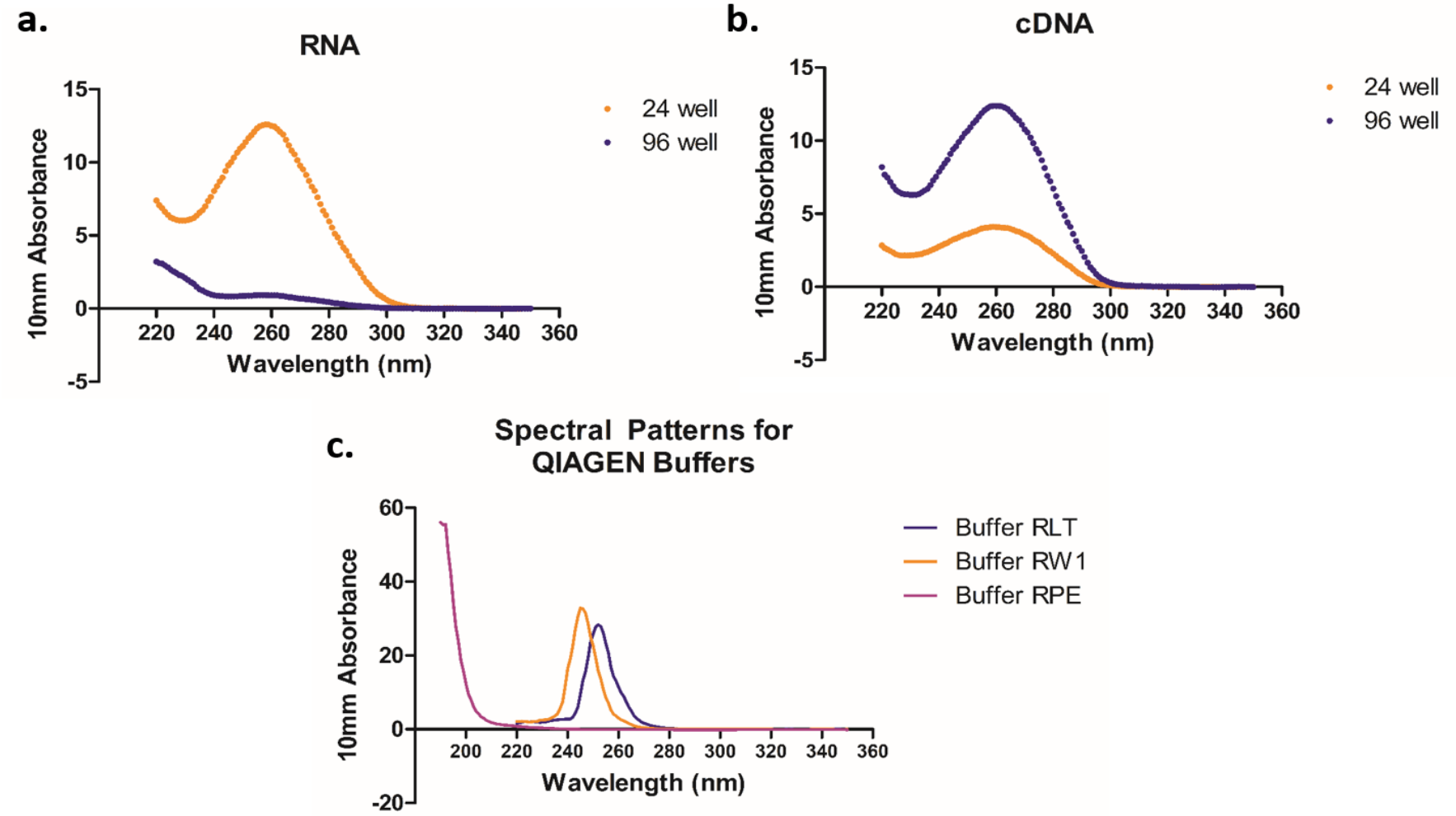



*Figure 6.2. Spectral patterns for SW620 cell RNA and cDNA.*

*(a) harvested RNA and the (b) reverse transcribed cDNA, both from SW620 cells grown in 24 and 96-well plates. Part (c) shows the spectral patterns for the buffers included with the QIAGEN RNeasy Mini Kit.*

#### 6.2.1.3 Amplification Efficiency

Calibration curves were performed to calculate the amplification efficiency for a selection of genes to assess the suitability of experimental conditions and RNA harvesting conditions for future assays. Due to limited resources and with 24 different genes to test, it was unfortunately not possible to perform calibration curves for all genes in all conditions. However, the reference genes (18S, β-actin, Tubulin and GAPDH) were all tested to ensure good amplification efficiency of these genes. In addition, the amplification efficiency of TGFβ was assessed as a test gene. The amplification efficiencies of cDNA samples from RNA harvested from different sized wells were compared to additionally assess the impact of potential differences in cDNA quality on amplification efficiencies of genes, as detailed below.

A serial dilution of cDNA harvested from 24- and 96-well plates was performed to assess whether the different well surface areas had an impact on the quality of cDNA amplification in qPCR reactions. In addition, the amplification efficiency of two primers were assessed to determine whether the primer design, qPCR conditions, melting temperatures etc. were optimal before proceeding with the bulk of qPCR reactions. Well-designed primers with optimal qPCR conditions have an amplification efficiency of 90-110% with an $R^2$ value >0.99 (Taylor *et al.*, 2019).

Calibration curves comparing 24-well and 96-well plates for 18S and Actin primers are displayed in Figure 6.3 and the amplification efficiencies and $R^2$ values can be found in Table 6.3. The $R^2$ values for the 96-well sample were poor and it can be seen on both calibration curves that the graph reaches a saturation point between 7-70 pg cDNA suggesting these values are outside the dynamic range of the primer pairs. The 7 pg cDNA samples were removed from the sample sets and the amplification efficiency and $R^2$ value recalculated with a dynamic range of 70 pg-70 ng cDNA. This gives an improved $R^2$ value in both cases, although the amplification efficiencies are still not within the acceptable range (Table 6.3).

The following calculation was used for amplification efficiency (Bustin *et al.*, 2009):

$$Efficiency = 10^{\left(\frac{-1}{slope}\right)} - 1.$$

*Table 6.3. Amplification efficiencies and $R^2$ values for 24- and 96-well cDNA samples*

*Based on the calibration curves in Figure 6.2. Values in red are outside of the normally accepted range for primer pairs in qPCR reactions (90-110% amplification efficiency and $R^2$ value >0.99).*

| Primer | Sample | Amplification Efficiency (%) | $R^2$ |
|---|---|---|---|
| 18S | 24-well | 89.5 | 0.991 |
| | 96-well (7pg included) | 111.6 | 0.953 |
| | 96-well (7pg excluded) | 86.7 | 0.988 |
| Actin | 24-well | 98.56 | 0.999 |
| | 96-well (7pg included) | 96.88 | 0.954 |
| | 96-well (7pg excluded) | 75.2 | 0.991 |

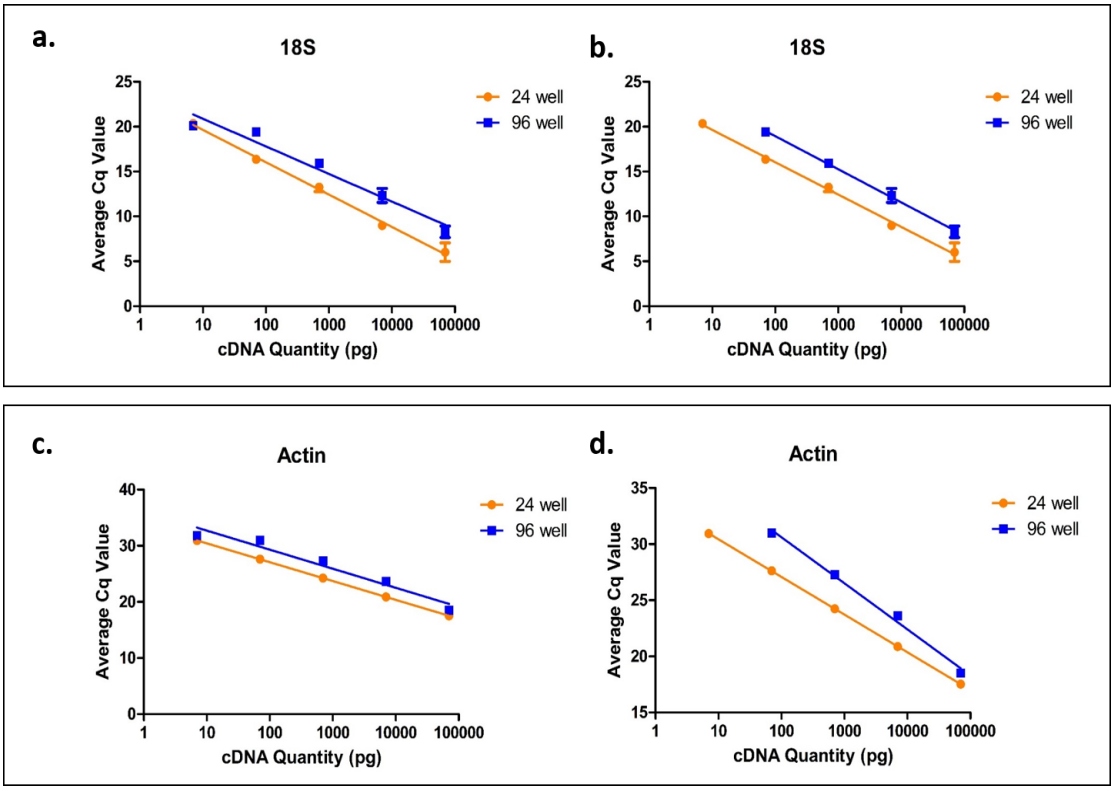



*Figure 6.3. Calibration curve for 18S and Actin primer pairs against cDNA obtained from 24- and 96- well plates.*

*Error bars are SD based on n=3. a) 18S primer with all points included. b) 18S primer with 7 pg point excluded as this point appeared to be outside the dynamic range. c) Actin primer with all points included. d) Actin primer with 7ng point excluded as this point appeared to be outside the dynamic range. The R-values for these calibration curves can be found in Table 6.3.*

### *6.2.1.4 Investigating the effect of 96- to 384-well miniaturisation and the effect of treating with resazurin prior to RNA harvest*

To allow further use of the limited protein available, the feasibility of the harvest of 3 x 384 wells and 3 x 384 wells previously treated with resazurin were assessed and compared to harvesting 1 x 96 well. For this investigation, three biological repeats were collected for each sample of BxPC-3 cells to allow assessment of biological variability amongst the samples. A summary of the concentration and quality (represented by the A260/280 and A260/230 ratios) of the RNA and cDNA are shown in Table 6.4. The spectral patterns for the RNA and cDNA samples are shown in Figure 6.4. It was suspected that the concentration readings of the RNA were being affected by contamination, leading to misleading readings. This is supported by the poor A260/230 readings (Table 6.4) and abnormal spectral patterns of the RNA (Figure 6.4). The measured RNA concentration is plotted against the cDNA concentration in Figure 6.4(c) and no correlation is seen between the two values ($R^2$=0.055).

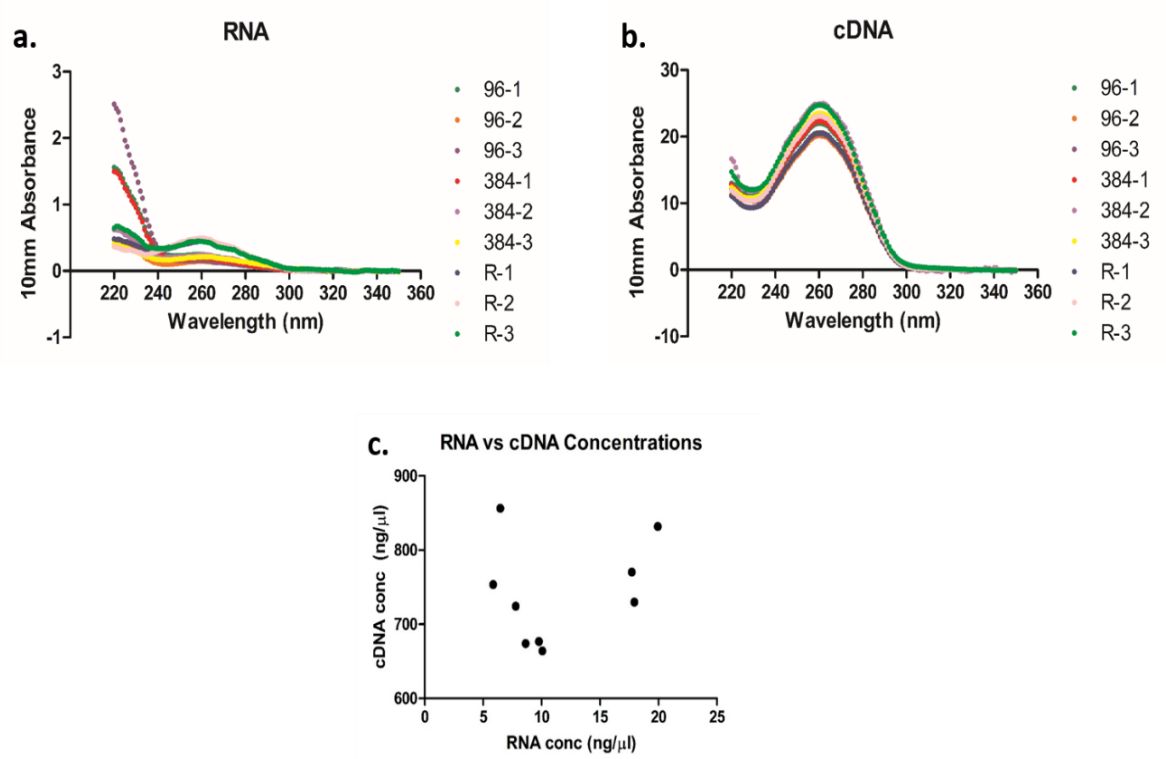



*Figure 6.4. Spectral patterns for the harvested RNA and DNA from 96 and 384-well plates.*

*(a) RNA and the (b) reverse transcribed cDNA from SW620 cells grown in 96-well, 384-well and R-treated plates. (c) RNA concentration plotted against cDNA concentration. The $R^2$ value for RNA vs CDNA is 0.05523 showing no correlation between the two values.*

*Table 6.4. Concentration and nucleotide quality ratios were assessed for the RNA and cDNA from the 96-well, 384-well and resazurin-treated (R) well samples.*

*Values in red lies outside the normal range (A260/280 ratio of ~2.0 for RNA and ~1.8 for DNA and A260/230 ratios of 2-2.2 for pure nucleic acids).*

| Sample | Type of nucleic acid | Concentration (ng/µl) | A260/280 | A260/230 |
|---|---|---|---|---|
| 96-well sample 1 | RNA | 9.77 | 1.82 | 0.24 |
| 96-well sample 2 | RNA | 5.85 | 1.71 | 0.49 |
| 96-well sample 3 | RNA | 6.46 | 1.85 | 0.11 |
| 384-well sample 1 | RNA | 7.78 | 1.93 | 0.21 |
| 384-well sample 2 | RNA | 10.07 | 1.82 | 0.56 |
| 384-well sample 3 | RNA | 8.63 | 1.61 | 0.74 |
| R sample 1 | RNA | 17.94 | 1.81 | 1.17 |
| R sample 2 | RNA | 19.94 | 1.95 | 1.70 |
| R sample 3 | RNA | 17.72 | 1.78 | 0.86 |
| 96-well sample 1 | cDNA | 724.37 | 1.86 | 1.98 |
| 96-well sample 2 | cDNA | 664.11 | 1.86 | 2.17 |
| 96-well sample 3 | cDNA | 674.24 | 1.84 | 1.98 |
| 384-well sample 1 | cDNA | 729.77 | 1.86 | 2.07 |
| 384-well sample 2 | cDNA | 831.92 | 1.79 | 2.09 |
| 384-well sample 3 | cDNA | 770.26 | 1.85 | 2.18 |
| R sample 1 | cDNA | 676.81 | 1.86 | 2.17 |
| R sample 2 | cDNA | 753.70 | 1.86 | 2.21 |
| R sample 3 | cDNA | 856.32 | 1.85 | 2.10 |

#### *6.2.1.5 Amplification efficiencies for 96-,384- and R-treated cells*

Both the 18S and Actin primers used in the 24/96-well assessment did not cross intron/exon boundaries. In addition, Actin gave a sub-optimal amplification efficiency and/or $R^2$ value even when cells were harvested from a 24-well plate and the purpose of this optimisation was to see whether cDNA from different sized plates influences amplification efficiency when other variables are controlled. For the next round of calibration curves, the 18S primer pair was kept but Actin was replaced. TGF-β was chosen as a suitable marker that amplifies at a relatively low Cq value and has the correct sized product and number of bands.

The calculated amplification efficiencies for each biological replicate and the amplification efficiency of all biological repeats were averaged and re-plotted, and the corresponding $R^2$ values are shown in Table 6.5. Amplification efficiency should be between 90-110% and the $R^2$ value should be >0.99 to be within acceptable limits. Many of the values, including the averaged values are outside of these parameters and indicate unsuitability of using 'miniaturised' assays such as 96- and 384- well plates. It is unclear whether treatment with resazurin has a negative effect on amplification efficiency or otherwise adversely affects downstream qPCR. Calibration curves can be found in Figure 6.5.

*Table 6.5. Summary of amplification efficiencies and $R^2$ values for 18-S and TGF-β primer pairs with cDNA from 96-, 384- and R-treated plates.*

*Samples 1, 2 and 3 represent the biological repeats for each condition. Average amplification efficiency and average $R^2$ values were calculated. Values in red indicate data which are outside of the acceptable range for primer pairs in qPCR reactions (90-110% amplification efficiency and $R^2$ value >0.99).*

| Primer | Sample | Individual Amplification Efficiency (%) | Individual $R^2$ | Average Amplification Efficiency (%) | Average $R^2$ |
|---|---|---|---|---|---|
| 18S | 96-1 | 86.82 | 0.997 | 86.39 | 0.999 |
| | 96-2 | 85.96 | 1.000 | | |
| | 96-3 | n/a | n/a | | |
| | 384-1 | 83.54 | 0.999 | 89.37 | 0.999 |
| | 384-2 | 91.36 | 0.999 | | |
| | 384-3 | 93.21 | 0.998 | | |
| | R-1 | 81.31 | 0.999 | 86.69 | 0.999 |
| | R-2 | 86.80 | 0.999 | | |
| | R-3 | 91.90 | 0.999 | | |
| TGF-β | 96-1 | 73.64 | 0.990 | 87.36 | 0.993 |
| | 96-2 | 101.08 | 0.996 | | |
| | 96-3 | n/a | n/a | | |
| | 384-1 | 104.69 | 0.983 | 112.6 | 0.961 |
| | 384-2 | 111.68 | 0.908 | | |
| | 384-3 | 121.44 | 0.993 | | |
| | R-1 | 91.90 | 0.999 | 98.62 | 0.993 |
| | R-2 | 87.30 | 0.999 | | |
| | R-3 | 116.66 | 0.980 | | |

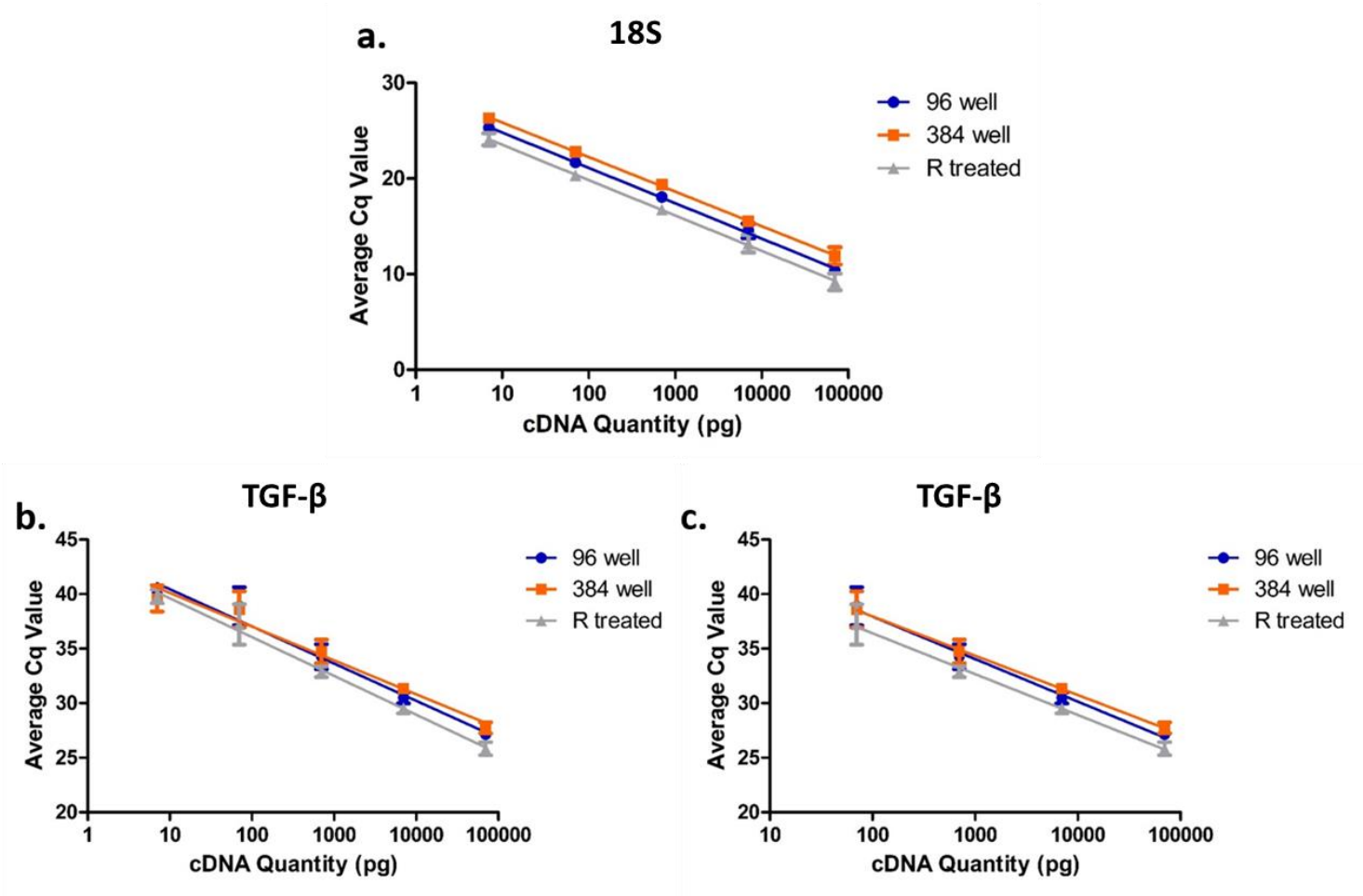



*Figure 6.5. Calibration curve of 18S and TGF-β primer pairs from 7 pg-70 ng cDNA*

*(a) 18S and (b) TGF-β primer pairs against cDNA obtained from 96-, 384-well and R-treated plates. Error bars are SD based on n=3. (c) TGF-β data replotted with 7 ng data point excluded.*

The TGF-β graph begins to level out at 7 ng due to the number of cycles being limited to 40 and thus anything not amplified was displayed as having a Cq value of 40. This may have skewed the data and so the amplification efficiency and $R^2$ values were re-calculated after removing the 7 ng data point (graph in Figure 6.5(c), data not shown). However, this was not enough to bring the amplification efficiency or $R^2$ values into the acceptable range.

#### *6.2.1.6 Comparison of Biological and Technical Repeats*

Due to scarcity of remaining 2D venom fraction samples, it was not possible for the qPCR study to perform both biological and technical repeats. Biological repeats were compared to assess whether technical repeats could be performed on a single biological sample for the pilot 2D venom fraction study. Biological repeats were taken from three different, independently harvested and reverse transcribed cell samples, whilst technical repeats were taken from the same biological sample, but assayed three times to compare pipetting, incomplete mixing etc.

The data from section 6.2.1.4 - *Investigating the effect of 96- to 384-well miniaturisation and the effect of treating with resazurin prior to RNA harvest* - was used to compare biological and technical repeats. From this data, it was assessed whether suitable conclusions may be inferred from technical repeats from one set of biological data. Biological and technical repeats are compared for 18S in Figure 6.6 and TGF-β in Figure 6.7.

In the 18S data (Figure 6.6), both biological and technical repeats appear reproducible (except 96-well plate sample 3 which had some technical problems and is excluded from Figure 6.6a). The TGF-β data (Figure 6.7) shows much more technical variation than the 18S data. This is likely due to the higher Cq values in the TGF-β data leading to increased sensitivity to initial cDNA copy number. The biological repeats appear similar for each of the replicated samples.

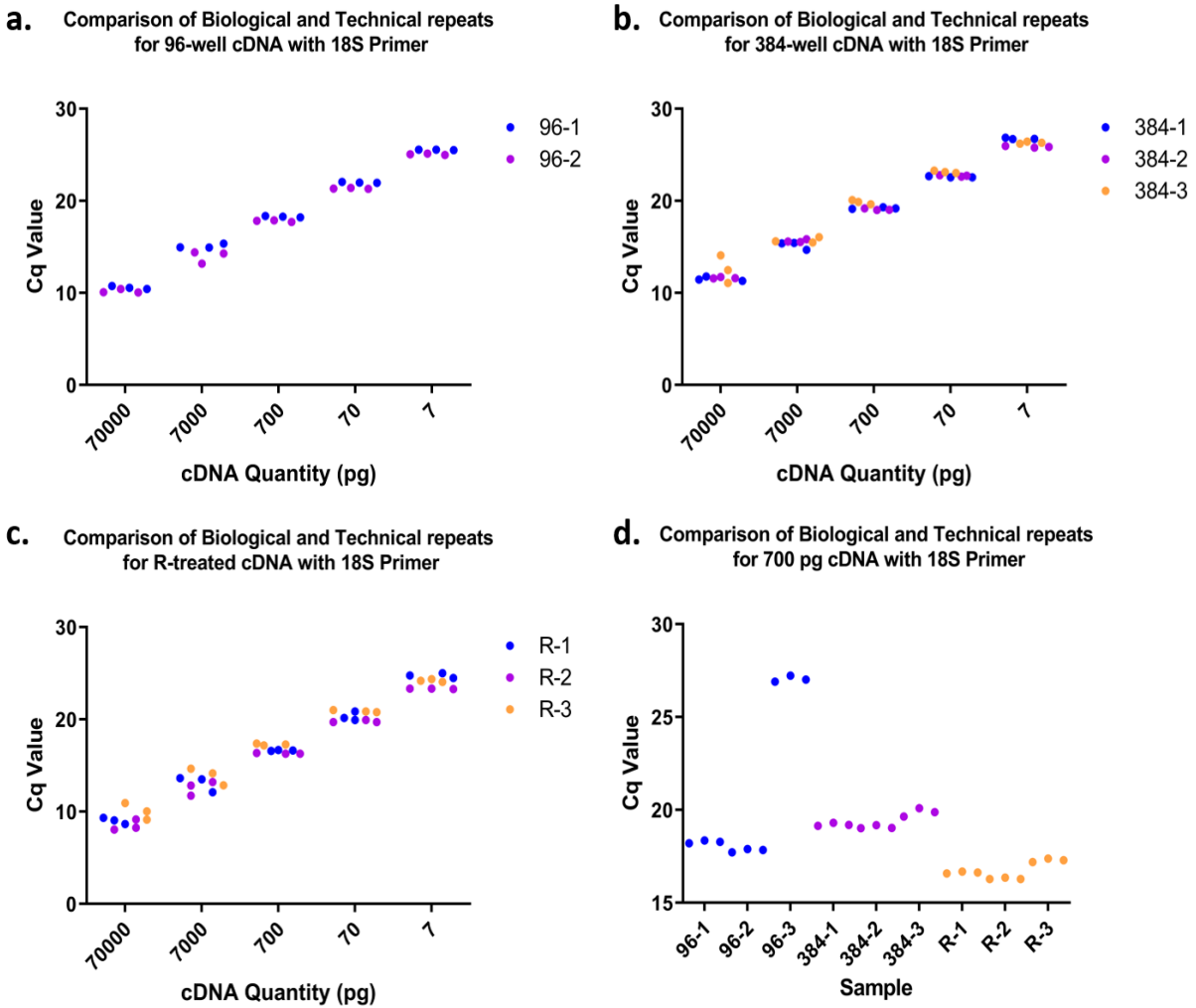



*Figure 6.6. Comparison of biological and technical repeats for different cDNA samples with 18S primer.*

*cDNA harvested from a) 96-well (96-3 data excluded from graph), b) 384-well, c) resazurin treated 384-well plate. d) Comparison of all biological and technical repeats at 700 pg concentration.*

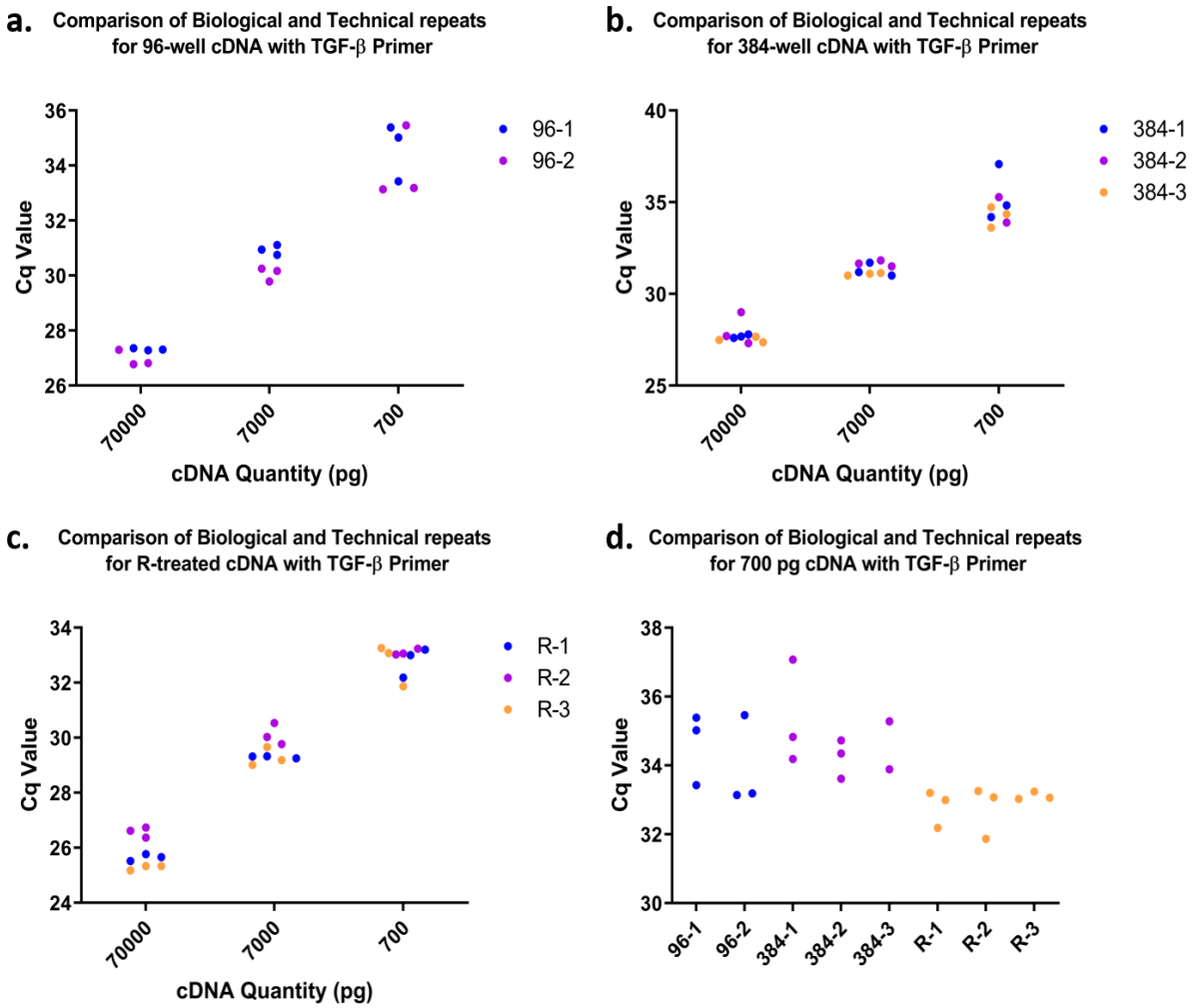



*Figure 6.7. Comparison of biological and technical repeats for different cDNA samples with TGFß primer.*

*cDNA harvested from a) 96-well (96-3 data excluded from graph), b) 384-well, c) resazurin treated 384-well plate. d) Comparison of all biological and technical repeats.*

#### *6.2.1.7 Selection of Appropriate Reference Genes for qPCR Analysis*

The Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin *et al.*, 2009) is a report that describes the factors of a qPCR experiment that must be controlled or optimised to ensure good quality, reproducible results. In this report, it is suggested that more than one reference gene should be used to compare the 'test genes' against to assess changes in expression level. Four reference genes were selected for this purpose – 18S, GAPDH, β-Actin and Tubulin. GAPDH and β-Actin were placed on every qPCR plate to also act as inter-plate calibrator genes.

To ensure appropriateness for use as reference genes, the selected genes were validated. The sequences for the 18S, GAPDH and Actin primer pairs were taken from the RTPrimerDB (found at rtprimerdb.org). The available Tubulin primer pairs on RTPrimerDB that use SYBR green dye did not fulfil the primer design criteria laid out in section 2.10.1.1. Therefore a primer pair for Tubulin was designed using the NCBI primer design tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/).

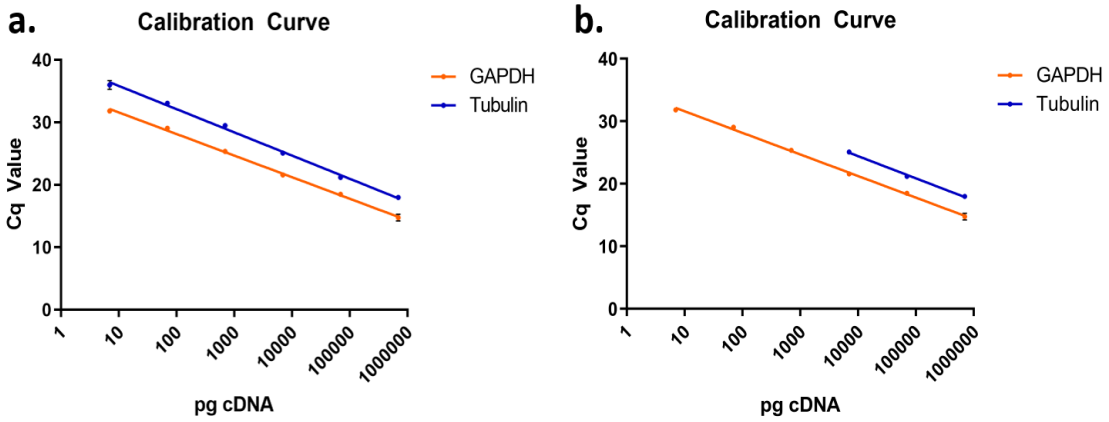
All potential reference genes amplified the correct sized product when run on an agarose gel (see Figure 6.1). Actin, Tubulin and GAPDH had consistently clear negative controls. 18S appeared to regularly have amplification in at least one of the NTCs approximately 20 cycles after amplification of the 10 ng cDNA template. This suggests either intermittent amplification of 18S from contaminating genomic DNA at approximately 1/1,000,000 the concentration of the 10 ng cDNA, or the presence of primer-dimers amplifying at approximately 33 cycles when added at a 300 nM concentration of each primer.

The amplification efficiencies of the housekeeping genes were tested at various points throughout the multi-well plate optimisation process. Amplification efficiency details for 18S and Actin may be found in section 6.2.1.3. Additional 18S amplification efficiency data may be found in section 6.2.1.4. Amplification efficiencies and $R^2$ values for GAPDH and Tubulin are presented in Table 6.6 and Figure 6.8. Data were produced from the six-hour time point of untreated BxPC-3 cells grown in 24-well plates. The amplification efficiency for Tubulin was slightly out of the optimal range of 90-110% when all cDNA concentrations were included. Excluding the 7, 70 and 700 pg points improved the amplification efficiency to be within the optimal range and offered the Tubulin primer pair an effective dynamic range of 7 -700 ng of cDNA per reaction (Figure 6.8).

*Table 6.6. Summary of amplification efficienceis and $R^2$ values for the four potential reference genes.*

*Values in red indicate data which are outside of the acceptable range for primer pairs in qPCR reactions (90-110% amplification efficiency and $R^2$ value >0.99).*

| Gene | Amplification Efficiency | $R^2$ Value |
|---|---|---|
| 18S | 89.50 | 0.991 |
| Actin | 98.56 | 0.999 |
| GAPDH | 94.80 | 0.997 |
| Tubulin | 85.9 | 0.995 |
| Tubulin (7 - 700 ng) | 91.3 | 0.996 |




c.

| Gene | Figure Part | Dynamic Range (ng) | Amplification efficiency (%) | $R^2$ |
|---|---|---|---|---|
| GAPDH | a) & b) | 0.007 - 700 | 94.8 | 0.997 |
| Tubulin | a) | 0.007 - 700 | 85.9 | 0.995 |
| Tubulin | b) | 7 – 700 | 91.3 | 0.996 |

*Figure 6.8. Calibration curve with line of best fit showing the amplification efficiency of GAPDH and Tubulin.*

*Error bars show SD based on n=3, part a) shows full calibration curves for both genes, $R^2$ = 0.997 (GAPDH, orange) and 0.995 (Tubulin, blue), part b) shows full calibration curve for GAPDH and a range of just 7-700 ng for Tubulin. This is so that the amplification efficiency for Tubulin falls within an acceptable range (90-110%). $R^2$ = 0.997 (GAPDH) and 0.996 (Tubulin).*

#### 6.2.1.8 Potential Interference by Venom Proteins

Using the BioRad CFX manager software (Version 3.1, Bio-Rad Laboratories Inc., Hercules, USA), the potential reference genes were selected individually as the single reference gene for the gene study for each time point. To be suitable, a reference gene will leave the majority of the 'test' genes with an expression value of approximately 1. If most genes appear to be up- or down-regulated, it is likely that the reference gene itself is being up- or down-regulated and is misrepresenting the 'base' expression levels of the other genes. A gene which causes most of the other genes to appear to be up/down-regulated is unsuitable as a reference gene. This is based on the hypothesis that most gene are unaffected by the test compound. Suitability of the selected control genes are investigated in Figure 6.9 (Actin), Figure 6.10 (Tubulin), Figure 6.11 (18S) and Figure 6.12 (GAPDH).

In the BxPC-3 cell line, using Actin as a reference gene causes all other gene values to appear downregulated. This effect is also seen when Tubulin is used as a control gene in the BxPC-3 cells, but only at the 24h time point. This makes both genes unsuitable for use as reference genes for this cell line. 18S and GAPDH genes demonstrate more stable expression levels at both time points.

To quantitatively assess whether the expression of reference genes was being affected by exposure to venom proteins, two sets of double-delta Cq analysis were performed with 18S and GAPDH as the 'control' genes and the other three (GAPDH, Actin and Tubulin or 18S, Actin and Tubulin, respectively) as the 'test' genes.

The results are displayed in Figure 6.13. Since the thresholds for biological significance were decided to be 2-fold and 0.5-fold, it was decided that Actin and Tubulin would be unsuitable as reference genes in this instance due to the >2-fold increase in their apparent expression when compared to GAPDH and 18S.

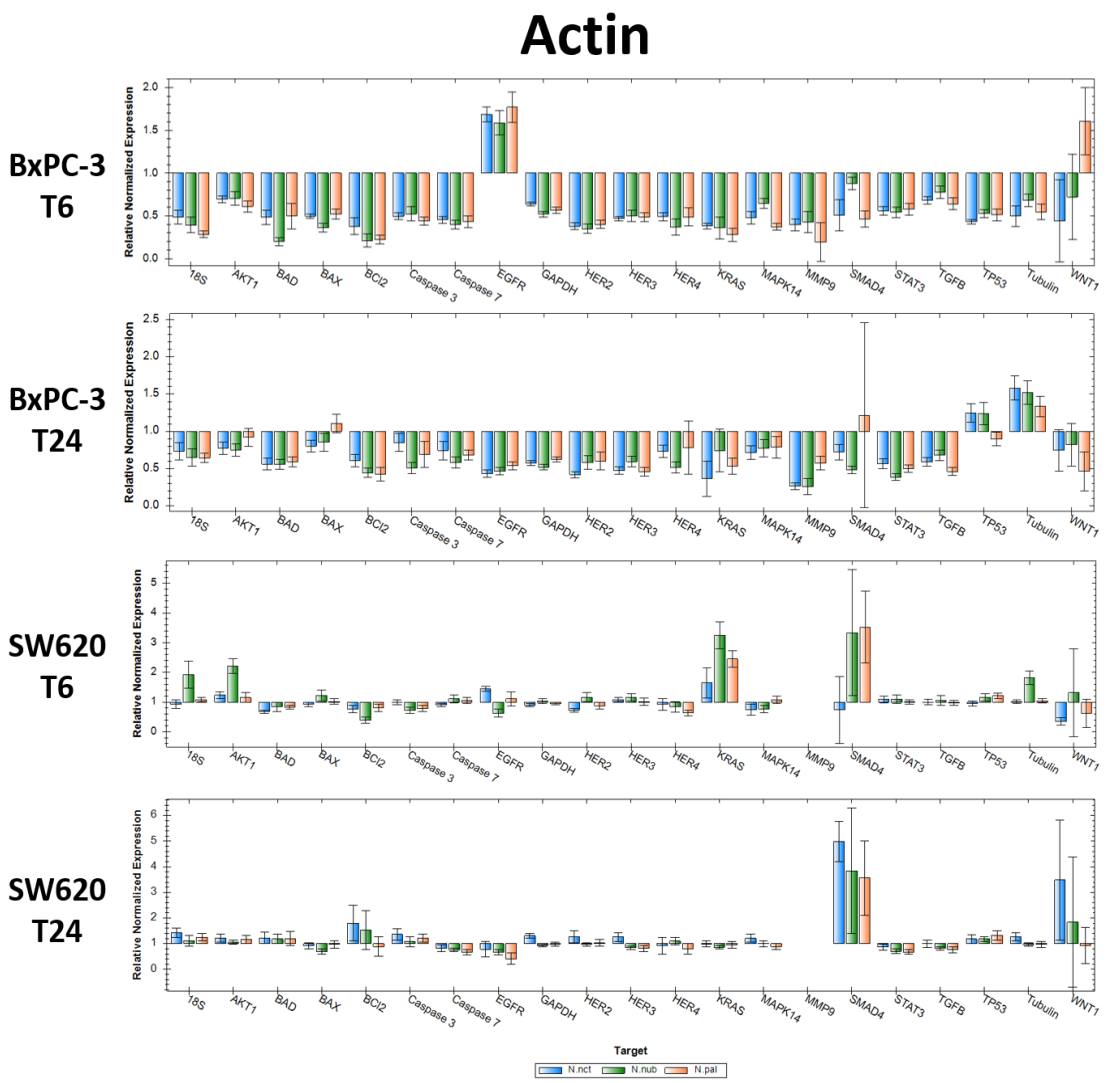



*Figure 6.9. Gene expression levels of a panel of cancer genes normalised against the Actin gene.*

*The charts show BxPC-3 (top) and SW620 (bottom) cell lines treated with N.nct_i18r2 (blue), N.nub_i17r2 (green) and N.pal_i17r2 (orange) venom fractions at 6h (T6) and 24h (T24) time points, normalised against the Actin gene. The expression level of each of the 21 genes (18 test genes plus the other 3 reference genes) relative to the expression of Actin is represented by the height of each bar where a value of 1 indicates gene expression level is equal to that of Actin. Error bars indicate standard deviation based on n=3.*

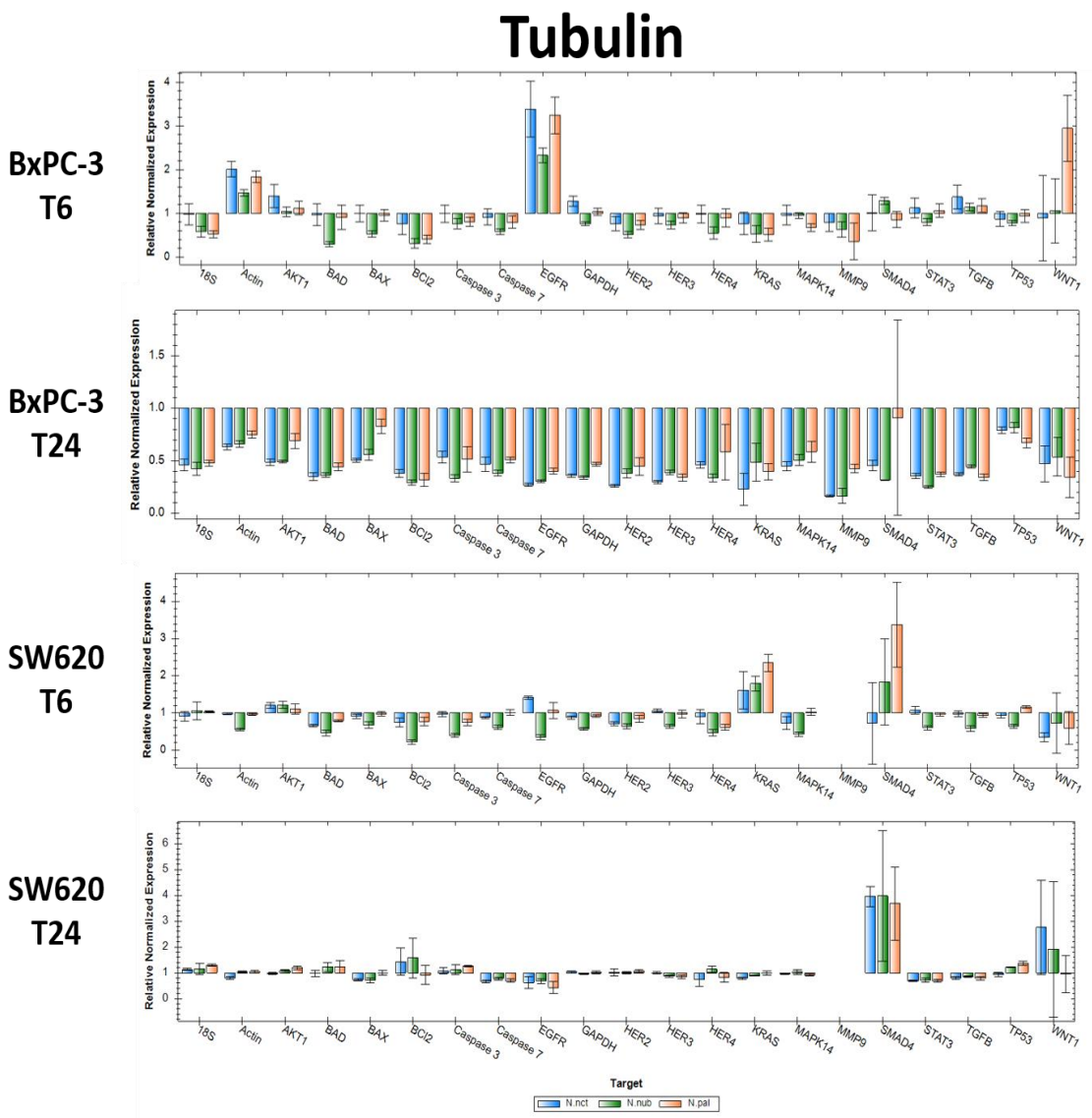



*Figure 6.10. Gene expression levels of a panel of cancer genes normalised against the Tubulin gene.*

*The charts show BxPC-3 (top) and SW620 (bottom) cell lines treated with N.nct_i18r2 (blue), N.nub_i17r2 (green) and N.pal_i17r2 (orange) venom fractions at 6h (T6) and 24h (T24) time points, normalised against the Tubulin gene. The expression level of each of the 21 genes (18 test genes plus the other 3 reference genes) relative to the expression of Tubulin is represented by the height of each bar where a value of 1 indicates gene expression level is equal to that of Tubulin. Error bars indicate standard deviation based on n=3.*

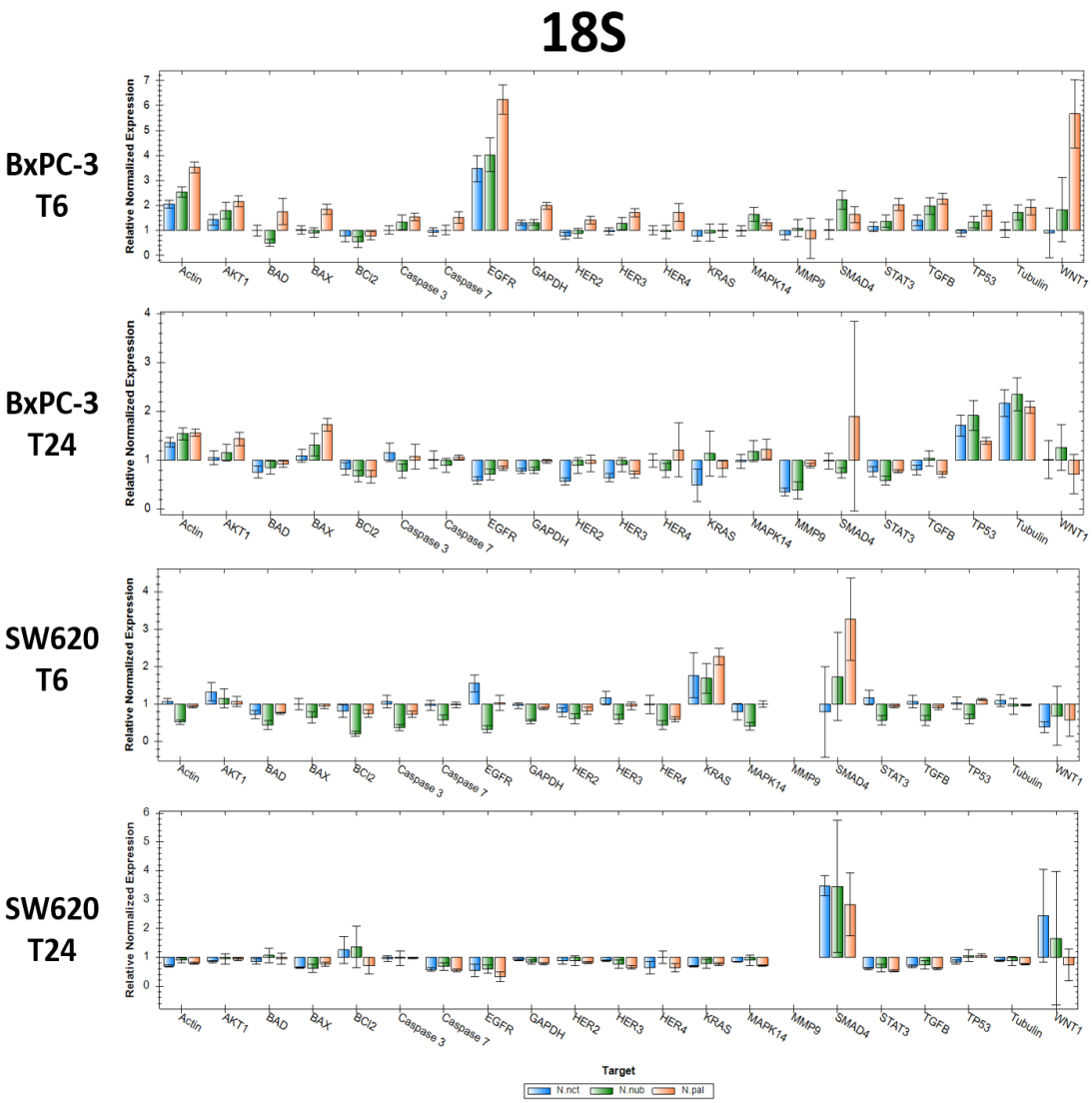



*Figure 6.11. Gene expression levels of a panel of cancer genes normalised against the 18S gene.*

*The charts show BxPC-3 (top) and SW620 (bottom) cell lines treated with N.nct_i18r2 (blue), N.nub_i17r2 (green) and N.pal_i17r2 (orange) venom fractions at 6h (T6) and 24h (T24) time points, normalised against the 18S gene. The expression level of each of the 21 genes (18 test genes plus the other 3 reference genes) relative to the expression of 18S is represented by the height of each bar where a value of 1 indicates gene expression level is equal to that of 18S. Error bars indicate standard deviation based on n=3.*

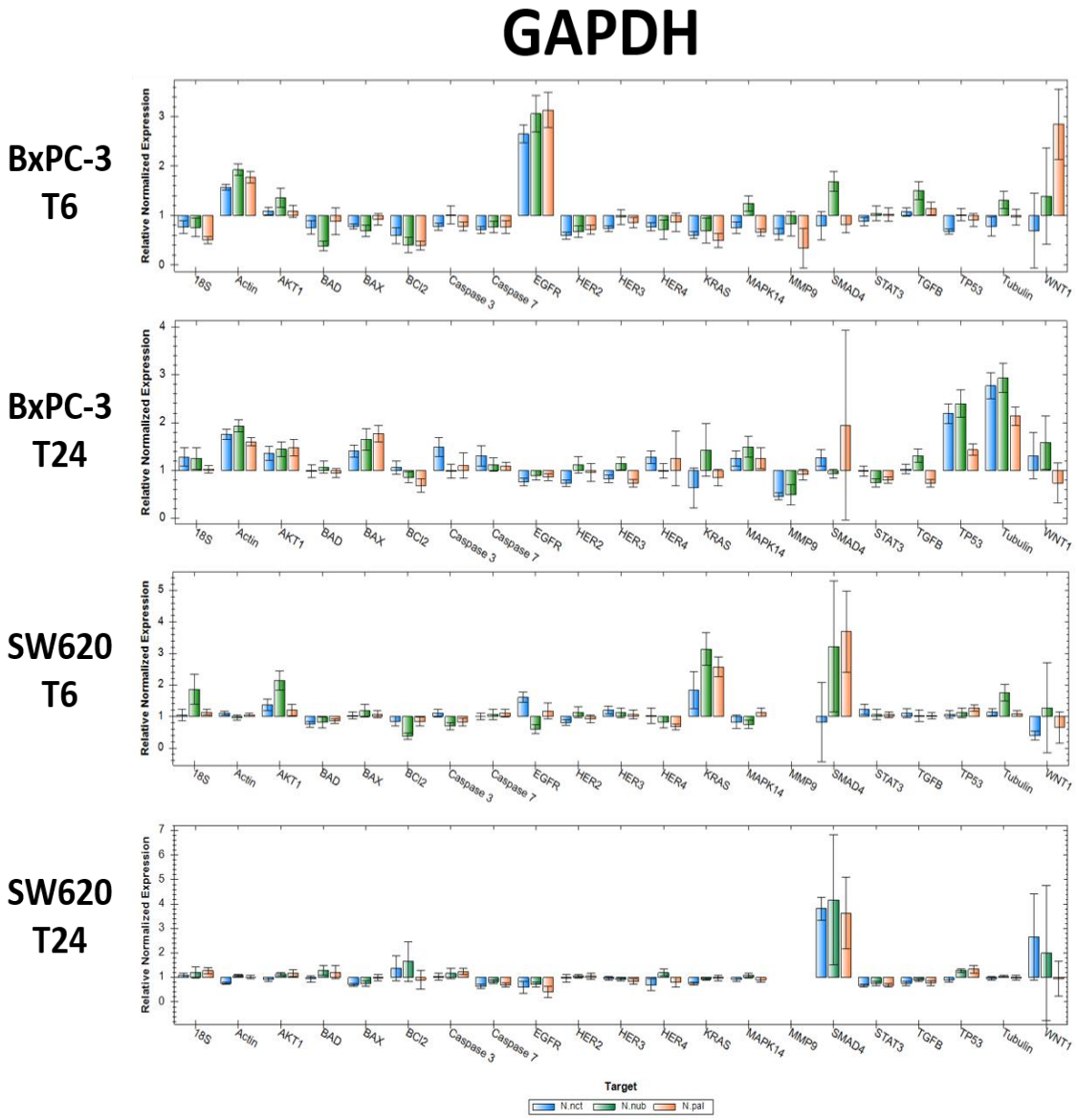



*Figure 6.12. Gene expression levels of a panel of cancer genes normalised against the GAPDH gene.*

*The charts show BxPC-3 (top) and SW620 (bottom) cell lines treated with N.nct_i18r2 (blue), N.nub_i17r2 (green) and N.pal_i17r2 (orange) venom fractions at 6h (T6) and 24h (T24) time points, normalised against the GAPDH gene. The expression level of each of the 21 genes (18 test genes plus the other 3 reference genes) relative to the expression of GAPDH is represented by the height of each bar where a value of 1 indicates gene expression level is equal to that of GAPDH. Error bars indicate standard deviation based on n=3.*

**a.**

$2^{\Delta\Delta Cq}$

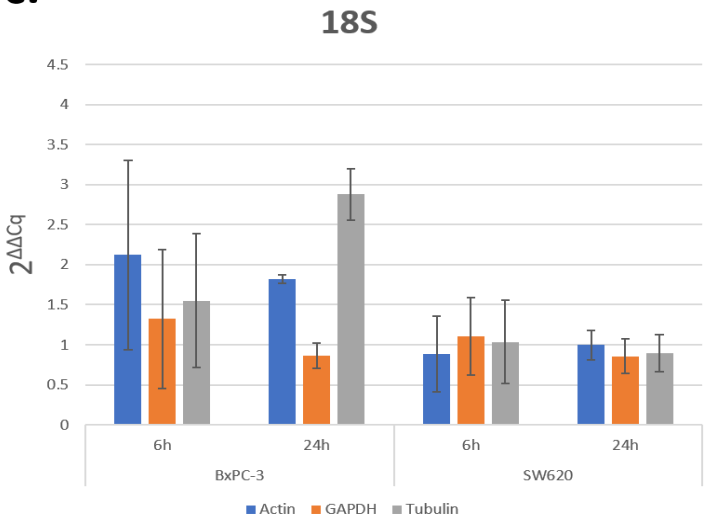
| 18S | | BxPC-3 | | SW620 | |
|---|---|---|---|---|---|
| | | 6h | 24h | 6h | 24h |
| GAPDH | N.nub | 1.33 | 0.82 | 1.02 | 0.83 |
| | N.pal | 1.86 | 0.97 | 1.09 | 0.97 |
| | N.nct | 1.10 | 0.88 | 0.94 | 0.80 |
| Actin | N.nub | 2.23 | 1.67 | 0.79 | 0.98 |
| | N.pal | 2.70 | 1.91 | 1.08 | 0.70 |
| | N.nct | 1.61 | 1.56 | 1.33 | 0.93 |
| Tubulin | N.nub | 1.73 | 2.35 | 0.94 | 0.87 |
| | N.pal | 1.92 | 2.08 | 0.97 | 0.77 |
| | N.nct | 1.14 | 2.17 | 1.10 | 0.88 |

**b.**

$2^{\Delta\Delta Cq}$

| GAPDH | | BxPC-3 | | SW620 | |
|---|---|---|---|---|---|
| | | 6h | 24h | 6h | 24h |
| 18S | N.nub | 0.75 | 1.22 | 0.98 | 1.20 |
| | N.pal | 0.54 | 1.03 | 0.91 | 1.03 |
| | N.nct | 0.91 | 1.14 | 1.06 | 1.25 |
| Actin | N.nub | 1.68 | 2.04 | 0.77 | 1.18 |
| | N.pal | 1.45 | 1.96 | 0.99 | 0.72 |
| | N.nct | 1.47 | 1.78 | 1.40 | 1.16 |
| Tubulin | N.nub | 1.30 | 2.87 | 0.92 | 1.04 |
| | N.pal | 1.03 | 2.14 | 0.88 | 0.79 |
| | N.nct | 1.04 | 2.48 | 1.16 | 1.10 |

**c.**




**d.**

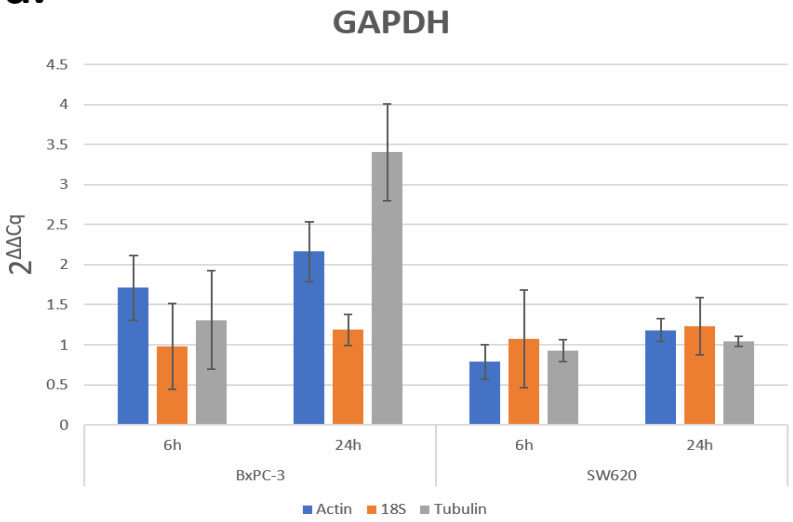



*Figure 6.13. Comparison of double delta Cq ($2^{\Delta\Delta Cq}$) analysis using 18S and GAPDH primer pairs as reference genes.*

*Parts (a) and (b) compare the $2^{\Delta\Delta Cq}$ results from 18S and GAPDH respectively for BxPC-3 and SW620 cell lines at 6 and 24 h time points against each of the three venom fractions (N.nub_i17r2, N.pal_i17r2 and N.nct_i18r2). The replicates for each gene for each sample at each time point were averaged before $2^{\Delta\Delta Cq}$ analysis was performed. Parts (c) and (d) show 18S and GAPDH respectively as reference genes with $2^{\Delta\Delta Cq}$ plotted for N.nub_i17r2 sample only for illustrative purposes. The$2^{\Delta\Delta Cq}$ values for these graphs were calculated by comparing the first replicate of each gene with the first replicate of the reference gene, then the second replicate of each gene with the second replicate of the reference gene, and so forth. Replicate comparison values were averaged and plotted with standard deviation plotted as error bars.*

Both 18S and GAPDH appeared appropriate as potential reference genes for the qPCR study. GAPDH demonstrated a better amplification efficiency than 18S (see Figure 6.8 for GAPDH and Table 6.5 for 18S) and so GAPDH was further investigated for suitability as a reference gene. Figure 6.14 shows that GAPDH has a single, correctly sized amplification product, a single peak on the melt curve and an appropriate amplification efficiency calculated from its calibration curve. The amplification efficiency and $R^2$ value can be found in Figure 6.5.

The results of the GAPDH temperature gradient analysis can be found in Figure 6.15. All primers were designed to have an annealing temperature of 60°C. Although the optimal temperature for the GAPDH primer was 62.9°C, an annealing temperature of 60°C gave a low Cq value and produced a single, correctly sized amplicon band with a sharp melt peak and good repeatability and so this temperature was selected to be in line with all other primer pair sets. Once the temperature reaches 67.3°C, an additional product appears to be amplified and can be seen on the melt curve for this temperature. Additional bands can be seen on the agarose gel at this temperature for both cell lines. This is likely to be non-specific amplification due to sub-optimal primer annealing temperature conditions.

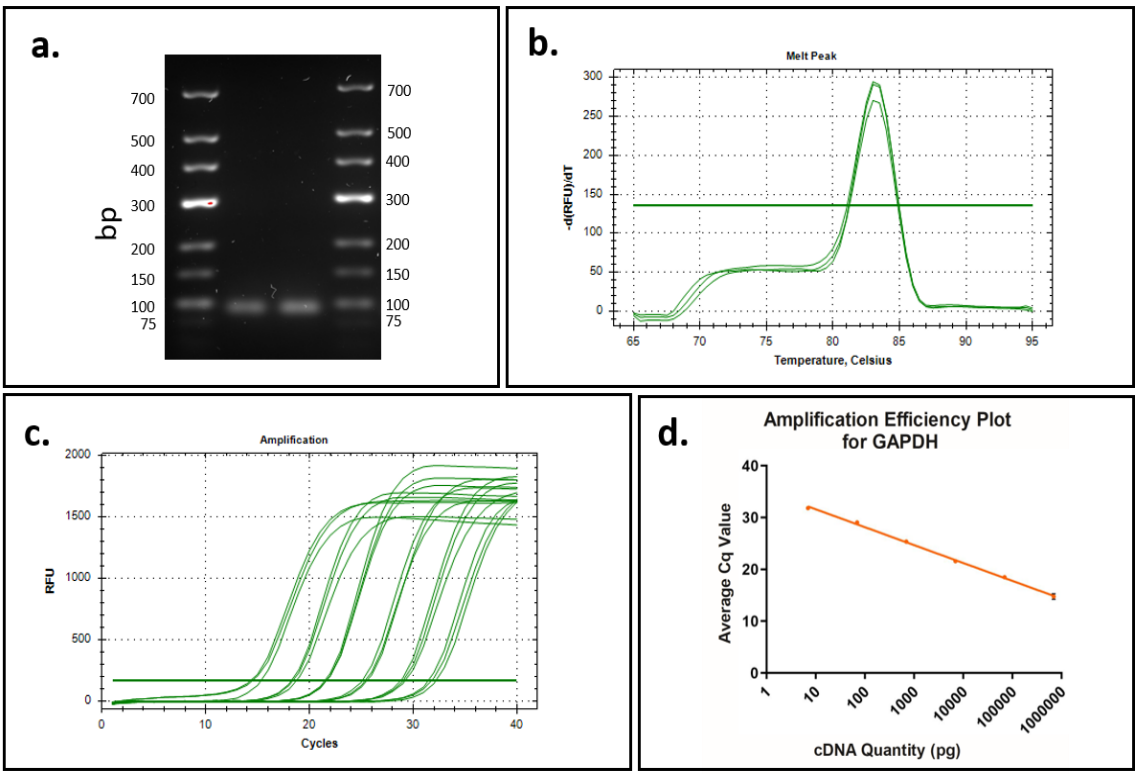



*Figure 6.14. GAPDH summary.*

*(a) agarose gel showing a single band at the anticipated 87 bp for both cell lines (b) Melt curve showing single peak at 83.5°C, (c) qPCR amplification graph for 10-fold serial dilution of cDNA from 7pg to 700 ng, (d) calibration curve used to calculate the amplification efficiency and $R^2$ values from average Cq values from 7 pg- 700 ng serial dilution. Plotted semi-logarithmically with SD error bars based on n=3.*

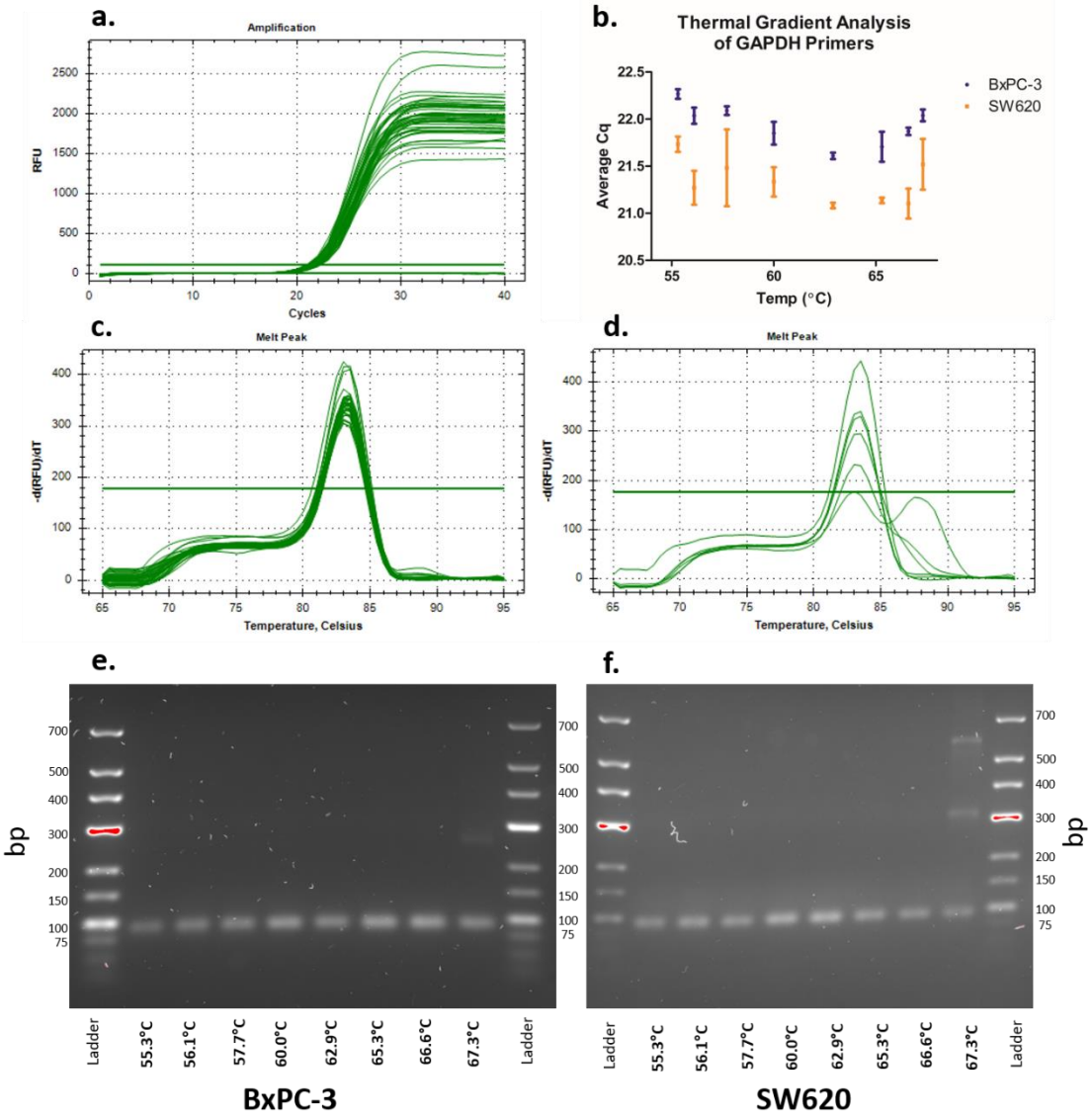



*Figure 6.15. GAPDH thermal gradient analysis.*

*(a) amplification plot for all temperatures, (b) average Cq plotted for each temperature with SD error bars based on n=3, (c) melt curve for all temperatures except 67.3°C, (d) melt curve for 67.3°C, agarose gels for the qPCR product for each temperature for (e) BxPC-3 cells,( f) SW620 cells.*

### 6.2.2 Investigating the Effect of Second Dimension Venom Fractions on Gene Regulation

A pilot study was performed using three of the cytotoxins identified in Chapter 5 (N.pal_i17r2, N.nct_i18r2 and N.nub_i17r2). BxPC-3 and SW620 cells were treated with the purified cytotoxins as detailed in section 2.10.3. The experimental factors used for the assay and the name for each RNA/cDNA sample are detailed in Table 6.7. Spectral patterns for the harvested RNA and reverse transcribed cDNA are shown in Figure 6.16 and details of RNA/cDNA concentration and quality (A260/280 and A260/230 ratios) are shown in Table 6.8. Except for SW_24_N.pal and Bx_24_N.nct, all RNA quality ratios were within the expected range. All cDNA quality ratios were within the expected range.

*Table 6.7. Experimental factors for the 2D fraction treated cells.*

*Each combination of cell line, treatment time and 2D fraction treatment has a unique identifier, as shown by the "Sample Name".*

| Sample Name | Cell Line | Treatment Time (Hours) | 2D Fraction Treatment |
|---|---|---|---|
| Bx_6_neg.con | BxPC-3 | 6 | Untreated |
| Bx_6_N.nub | BxPC-3 | 6 | N.nub_i17r2 |
| Bx_6_N.pal | BxPC-3 | 6 | N.pal_i17r2 |
| Bx_6_N.nct | BxPC-3 | 6 | N.nct_i18r2 |
| Bx_24_neg.con | BxPC-3 | 24 | Untreated |
| Bx_24_N.nub | BxPC-3 | 24 | N.nub_i17r2 |
| Bx_24_N.pal | BxPC-3 | 24 | N.pal_i17r2 |
| Bx_24_N.nct | BxPC-3 | 24 | N.nct_i18r2 |
| SW_6_neg.con | SW620 | 6 | Untreated |
| SW_6_N.nub | SW620 | 6 | N.nub_i17r2 |
| SW_6_N.pal | SW620 | 6 | N.pal_i17r2 |
| SW_6_N.nct | SW620 | 6 | N.nct_i18r2 |
| SW_24_neg.con | SW620 | 24 | Untreated |
| SW_24_N.nub | SW620 | 24 | N.nub_i17r2 |
| SW_24_N.pal | SW620 | 24 | N.pal_i17r2 |
| SW_24_N.nct | SW620 | 24 | N.nct_i18r2 |

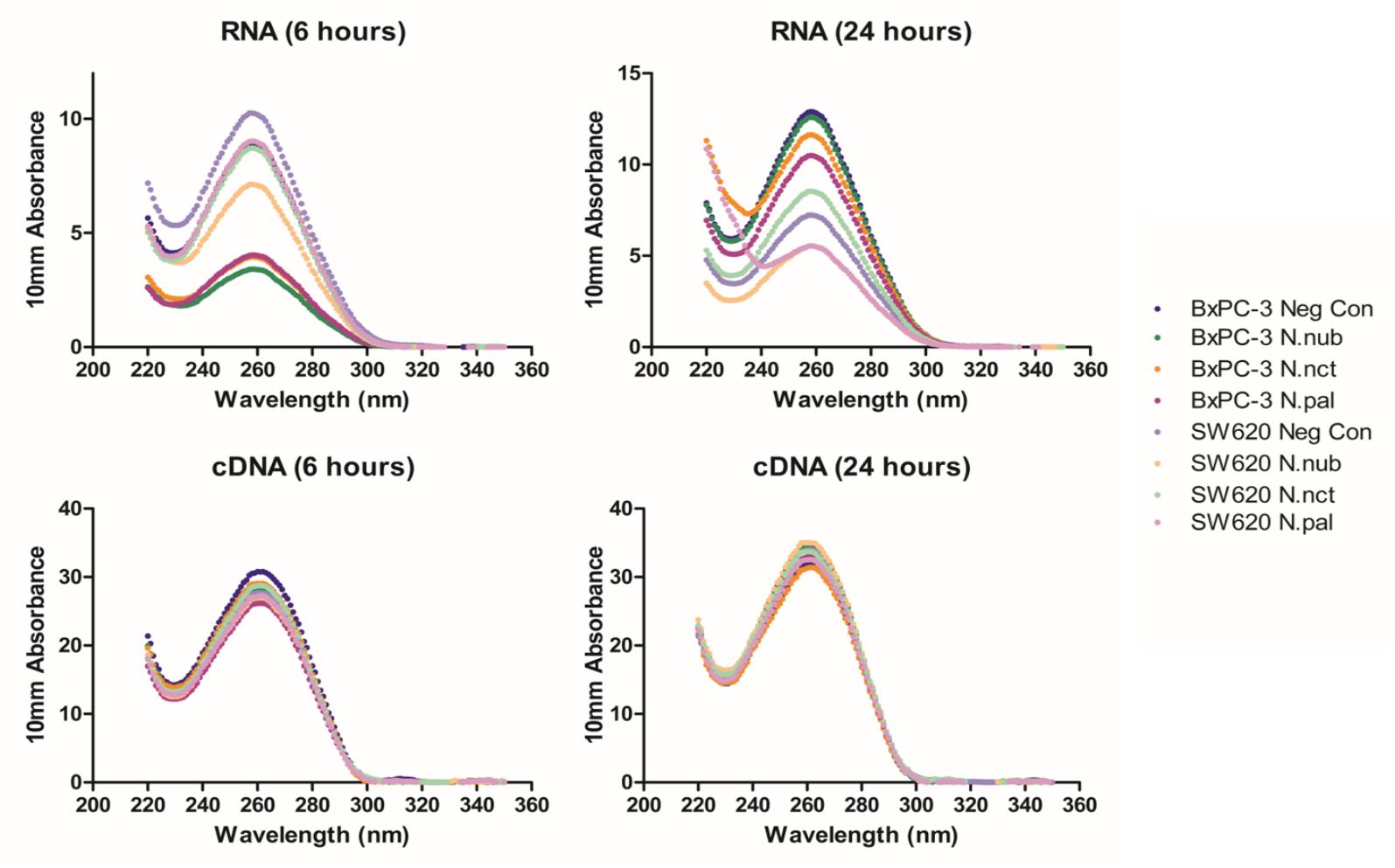



*Figure 6.16. Spectral patterns for the harvested RNA and reverse transcribed cDNA at 6h and 24h time points.*

*All spectral patterns were as anticipated for nucleic acids except for SW_24_N.pal (SW620 cells treated with N.pal_i17r2 cytotoxin at the 24 hour time point) which has a much higher peak at around 220 nm compared to its 260 nm peak.*

*Table 6.8. Concentration and quality of purified RNA and reverse transcribed cDNA from the 2D venom fraction pilot study.*

*Values in red lies outside the normal range (A260/280 ratio of ~2.0 for RNA and ~1.8 for DNA and A260/230 ratios of 2-2.2 for pure nucleic acids).*

| Type of nucleic acid | Sample | Concentration (ng/µl) | A260/280 | A260/230 |
|---|---|---|---|---|
| RNA | Bx_6_neg.con | 358.87 | 2.09 | 2.16 |
| | Bx_6_N.nub | 159.82 | 2.10 | 1.87 |
| | Bx_6_N.pal | 160.76 | 2.08 | 2.12 |
| | Bx_6_N.nct | 138.89 | 2.13 | 1.87 |
| | Bx_24_neg.con | 514.25 | 2.11 | 2.16 |
| | Bx_24_N.nub | 501.12 | 2.11 | 2.16 |
| | Bx_24_N.pal | 418.69 | 2.09 | 2.05 |
| | Bx_24_N.nct | 464.27 | 2.10 | 1.46 |
| | SW_6_neg.con | 409.93 | 2.09 | 2.06 |
| | SW_6_N.nub | 283.87 | 2.11 | 1.90 |
| | SW_6_N.pal | 358.35 | 2.09 | 2.24 |
| | SW_6_N.nct | 352.40 | 2.09 | 2.26 |
| | SW_24_neg.con | 287.32 | 2.08 | 2.06 |
| | SW_24_N.nub | 220.84 | 2.09 | 2.16 |
| | SW_24_N.pal | 219.14 | 2.11 | 0.78 |
| | SW_24_N.nct | 340.54 | 2.09 | 2.16 |
| cDNA | Bx_6_neg.con | 1012.12 | 1.82 | 2.14 |
| | Bx_6_N.nub | 937.94 | 1.84 | 2.11 |
| | Bx_6_N.pal | 871.66 | 1.84 | 2.12 |
| | Bx_6_N.nct | 957.44 | 1.86 | 2.13 |
| | Bx_24_neg.con | 1053.24 | 1.82 | 2.21 |
| | Bx_24_N.nub | 1032.38 | 1.85 | 2.16 |
| | Bx_24_N.pal | 1100.91 | 1.87 | 2.03 |
| | Bx_24_N.nct | 1128.67 | 1.87 | 2.14 |
| | SW_6_neg.con | 917.79 | 1.85 | 2.14 |
| | SW_6_N.nub | 884.58 | 1.86 | 2.14 |
| | SW_6_N.pal | 897.19 | 1.87 | 2.14 |
| | SW_6_N.nct | 955.21 | 1.86 | 2.14 |
| | SW_24_neg.con | 1135.90 | 1.87 | 2.17 |
| | SW_24_N.nub | 1158.06 | 1.85 | 2.13 |
| | SW_24_N.pal | 1076.25 | 1.85 | 2.14 |
| | SW_24_N.nct | 1101.73 | 1.86 | 2.13 |

### 6.2.3 Summary of Gene Regulation

Gene regulation was determined using the $2^{-\Delta\Delta Ct}$ method detailed in section 2.10.2.7. The results of the gene regulation of each cell line following exposure to the different venom samples at different time points are shown in Table 6.9 and a matrix summarising the gene regulation is shown in Figure 6.17.

| | ***BxPC-3*** | | | | | | ***SW620*** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***6h*** | | | ***24h*** | | | ***6h*** | | | ***24h*** | | |
| | ***N.nct*** | ***N.nub*** | ***N.pal*** | ***N.nct*** | ***N.nub*** | ***N.pal*** | ***N.nct*** | ***N.nub*** | ***N.pal*** | ***N.nct*** | ***N.nub*** | ***N.pal*** |
| ***18S*** | | | | | | | | | | | | |
| ***Actin*** | | | | | | | | | | | | |
| ***Tubulin*** | | | | | | | | | | | | |
| ***AKT1*** | | | | | | | | | | | | |
| ***KRAS*** | | | | | | | | | | | | |
| ***MAPK14*** | | | | | | | | | | | | |
| ***SMAD4*** | | | | | | | | | | | | |
| ***TGFB*** | | | | | | | | | | | | |
| ***TP53*** | | | | | | | | | | | | |
| ***Caspase 3*** | | | | | | | | | | | | |
| ***Caspase 7*** | | | | | | | | | | | | |
| ***MMP9*** | | | | | | | | | | | | |
| ***STAT3*** | | | | | | | | | | | | |
| ***BAX*** | | | | | | | | | | | | |
| ***BAD*** | | | | | | | | | | | | |
| ***BCL2*** | | | | | | | | | | | | |
| ***WNT1*** | | | | | | | | | | | | |
| ***NOTCH1*** | | | | | | | | | | | | |
| ***EGFR*** | | | | | | | | | | | | |
| ***HER2*** | | | | | | | | | | | | |
| ***HER3*** | | | | | | | | | | | | |

*Figure 6.17. Matrix showing gene regulation for each gene tested in the study in response to the three venom fractions tested for both cell lines at both time points (6 and 24 hours).*

*Green colour indicates an upregulation (≥ 2-fold control) of the relevant gene in response to treatment. Orange colour indicates a down regulation (≤ 0.5-fold control) of the relevant gene in response to treatment. Genes shown in light grey indicate no significant change in gene regulation (2^-ΔΔCt between 0.5 to 2-fold compared to the untreated controls) All values were calculated through the 2^-ΔΔCt calculation using GAPDH as the reference gene and comparing the test condition (venom treatment) against the untreated control for each time point. Genes in dark grey were excluded from analysis as they were expressed in an extremely low quantity (EGFR in SW620 cell line) or not at all (SMAD4 in BxPC-3 and MMP9 in SW620 cell lines).*

*Table 6.9. Summary of gene regulation changes in response to the three different cytotoxins at 6h and 24h time points for each cell line.*

*Values represent the ΔΔCq value and indicate gene regulation compared to the untreated control cells (ΔΔCq = 1 indicates no change in gene regulation compared to the untreated control cells). Gene which underwent upregulation (ΔΔCq ≥ 2) are highlighted in green and genes which underwent downregulation (ΔΔCq ≤ 0.5) are highlighted in red. Genes not expressed or with extremely low expression in a cell line are highlighted in grey. N.nub = cytotoxin from Naja nubiae venom, N.pal = cytotoxin from Naja pallida venom, N.nct = cytotoxin from Naja nigricincta venom.*

| Gene of Interest | BxPC-3 | | | | | | Gene of Interest | SW620 | | | | | | Gene of Interest |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N.nct | | N.nub | | N.pal | | | N.nct | | N.nub | | N.pal | | |
| | 6h | 24h | 6h | 24h | 6h | 24h | | 6h | 24h | 6h | 24h | 6h | 24h | |
| 18S | 0.91 | 1.14 | 0.75 | 1.22 | 0.54 | 1.03 | 18S | 1.06 | 1.25 | 0.98 | 1.20 | 0.91 | 1.03 | 18S |
| Actin | 1.78 | 1.84 | 2.52 | 1.91 | 2.30 | 1.35 | Actin | 1.00 | 0.65 | 1.02 | 1.01 | 1.05 | 1.20 | Actin |
| Tubulin | 1.04 | 2.48 | 1.30 | 2.87 | 1.03 | 2.14 | Tubulin | 1.16 | 1.10 | 0.92 | 1.04 | 0.88 | 0.79 | Tubulin |
| AKT1 | 1.31 | 1.20 | 1.35 | 1.41 | 1.16 | 1.47 | AKT1 | 1.40 | 1.06 | 1.12 | 1.13 | 0.98 | 0.95 | AKT1 |
| KRAS | 0.72 | 0.56 | 0.69 | 1.39 | 0.53 | 0.85 | KRAS | 1.87 | 0.87 | 1.65 | 0.93 | 2.08 | 0.78 | KRAS |
| MAPK14 | 0.71 | 1.37 | 0.92 | 1.65 | 0.72 | 1.31 | MAPK14 | 0.91 | 0.85 | 0.87 | 1.00 | 1.13 | 0.95 | MAPK14 |
| SMAD4 | | | | | | | SMAD4 | 0.90 | 3.51 | 3.75 | 3.87 | 3.70 | 3.82 | SMAD4 |
| TGFB | 1.01 | 1.13 | 1.11 | 1.45 | 1.24 | 0.76 | TGFB | 1.21 | 0.70 | 1.19 | 0.82 | 1.02 | 0.81 | TGFB |
| TP53 | 0.64 | 2.40 | 0.75 | 2.66 | 0.99 | 1.49 | TP53 | 1.17 | 0.83 | 1.31 | 1.19 | 1.27 | 1.42 | TP53 |
| Caspase 3 | 0.75 | 1.20 | 0.99 | 0.87 | 0.74 | 1.04 | Caspase 3 | 0.99 | 1.11 | 0.81 | 1.28 | 0.86 | 1.66 | Caspase 3 |
| Caspase 7 | 0.69 | 1.04 | 0.75 | 0.99 | 0.72 | 1.03 | Caspase 7 | 0.89 | 0.67 | 1.25 | 0.90 | 1.16 | 0.92 | Caspase 7 |
| MMP9 | 0.60 | 0.37 | 0.81 | 0.43 | 0.32 | 0.86 | MMP9 | | | | | | | MMP9 |
| STAT3 | 0.85 | 0.79 | 1.03 | 0.65 | 0.97 | 0.75 | STAT3 | 1.09 | 0.71 | 1.23 | 0.83 | 1.10 | 0.91 | STAT3 |
| BAD | 0.85 | 1.03 | 0.57 | 1.02 | 0.90 | 0.94 | BAD | 0.79 | 0.88 | 0.98 | 1.34 | 0.91 | 1.22 | BAD |
| BAX | 0.88 | 1.47 | 1.03 | 1.58 | 0.95 | 1.76 | BAX | 1.09 | 0.66 | 1.43 | 0.78 | 1.14 | 0.99 | BAX |
| BCL2 | 0.67 | 1.11 | 0.60 | 0.82 | 0.40 | 0.68 | BCL2 | 0.89 | 1.30 | 0.46 | 1.73 | 0.89 | 0.91 | BCL2 |
| WNT1 | 0.78 | 1.37 | 2.08 | 1.51 | 2.92 | 0.73 | WNT1 | 0.42 | 2.50 | 1.54 | 2.10 | 0.69 | 0.95 | WNT1 |
| NOTCH1 | 0.98 | 0.81 | 0.46 | 0.91 | 0.94 | 0.83 | NOTCH1 | 1.13 | 0.81 | 1.40 | 0.91 | 1.36 | 0.83 | NOTCH1 |
| EGFR | 2.16 | 0.92 | 2.85 | 0.98 | 2.77 | 0.88 | EGFR | | | | | | | EGFR |
| HER2 | 0.48 | 0.89 | 0.63 | 1.22 | 0.63 | 0.97 | HER2 | 0.77 | 0.92 | 1.31 | 0.99 | 1.03 | 0.93 | HER2 |
| HER3 | 0.60 | 1.01 | 0.90 | 1.25 | 0.76 | 0.74 | HER3 | 1.15 | 0.91 | 1.30 | 0.87 | 1.19 | 0.73 | HER3 |
| HER4 | 0.63 | 1.55 | 0.67 | 1.08 | 0.76 | 1.27 | HER4 | 0.98 | 0.67 | 0.94 | 1.13 | 0.75 | 0.71 | HER4 |

#### 6.2.3.1 Reference Genes

The 18S gene shows no significant changes at any time point and is acting as a good control gene showing that the qPCR is amplifying as anticipated. The Actin gene shows an increase at the 6 hour time point for *Naja nubiae* and *Naja pallida* venom samples in BxPC-3 cell line. No change is observed in the SW620 cell line. An increase in expression levels is also observed for the Tubulin gene at 24h for all venoms tested in the BxPC-3 cell line. No change is observed in the SW620 cell line.

#### 6.2.3.2 Changes in Genes of Interest

The significance threshold for changes in gene regulation was deemed to be changes in $2^{-\Delta\Delta Ct}$ values of less than 0.5 and more than 2 times normal gene expression levels. The expression of most of the genes tested were not significantly affected by the addition of venom samples using these thresholds but a few genes did show a significant change in gene regulation.

The Tp53 gene was significantly upregulated when BxPC-3 cells were treated with *Naja nubiae* and *Naja nigricincta* venom samples for 24 hours. EGFR was significantly upregulated by all three venoms in the BxPC-3 cell line following 6 hours of exposure.

There was a significant increase in SMAD4 expression levels in SW620 cells when exposed to all venom samples at all time points except *Naja nigricincta* sample following 6 hours of exposure. Raw Cq values for untreated cells were above 35 cycles for SW620 cells, indicating that 35 cycles of PCR replication were required before the SMAD4 gene was detectable in these samples. This suggests a very low expression of SMAD4. BxPC-3 cells have a homozygous deletion in SMAD4 and therefore do not express the SMAD4 gene so were not tested in qPCR. WNT1 was upregulated in BxPC-3 cells treated with *Naja nubiae* and *Naja pallida* for 6 hours and in SW620 cells when exposed to *Naja nigricincta* and *Naja nubiae* venom for 24 hours. This gene was also downregulated following 6 hours of exposure to *Naja nigricincta* venom.

In the BxPC-3 cell line, MMP-9 was significantly downregulated after 6 hours of exposure to the *Naja* pallida venom sample. Although this effect was not seen after 24 hours exposure to the same venom sample, it should be noted that Cq values could not be collected for two out of three samples during the first experiment. Following 24h exposure to both *Naja nigricincta* and *Naja nubiae* venom samples, a similar downregulation of MMP-9 was observed in BxPC-3 cells. The SW620 cell line does not express the MMP-9 gene and so was omitted from analysis.

A significant decrease was observed in the Bcl-2 gene following exposure of BxPC-3 cells to *Naja pallida* venom sample for 6 hours and SW620 cells to *Naja nubiae* venom sample for 6 hours.

## 6.3 Discussion

### 6.3.1 Assay Miniaturisation

The selected second dimension fractions had very low quantities of protein available for cell treatment prior to RNA harvesting. Therefore, an assay was run to test the feasibility of harvesting RNA samples from 96 or 384 well format plates rather than the standard 24 well plate (approximate growth areas of Corning 24 well plate: 1.9cm$^2$ , 96-well plate: 0.32 cm$^2$, 384 well plate: 0.056 cm$^2$ (https://www.corning.com/catalog/cls/documents/application-notes/CLS-AN-209.pdf (accessed 21/01/2020)). 24-well plates have approximately six times the growth area compared to 96-well plates and 96-well plates have approximately six times the growth are compared to 384-well plates. There is therefore approximately 36 times the growth area in 24-well compared to 384-well format and a corresponding 36-fold reagent requirement.

Growing and treating cells in 24-well format requires additional cells and, crucially, a larger quantity of 2D venom fraction to provide the same concentration of sample on the cells. Miniaturising the assay format would reduce the reagents and cells needed and allow the limited 2D venom fractions to go further. The quality of cDNA produced from cells pre-treated with resazurin was assessed to determine the feasibility of venom-treating cells for the resazurin dose response assay, then washing and harvesting the same cells for qPCR analysis. If possible, this would allow enough venom sample to complete both assays and would provide qPCR samples from the exact same cells on which the dose response was performed, mitigating any potential assay to assay variability.

### 6.3.2 Poor Quality RNA

When the assay was miniaturised from 24- to 96-well and from 96- to 384- well format, low A260/230 ratios were observed suggesting the presence of a contaminant(s) with an absorbance value of 230 nm. Guanidine isothiocyanate is present in the QIAGEN buffer RLT RNA extraction buffer and low levels of contamination have been shown to have a large effect on A260/230 ratio. A 0.5% contamination level can cause the A260/230 ratio to drop below 0.5 (https://www.sigmaaldrich.com/technical-documents/articles/biology/sample-purification-and-quality-assessment.html#nucliec (accessed 10/01/2020)). This could also be due to the presence of a contaminant produced by the cells themselves as the confluency level of the cells increased as the well size decreased. The resazurin treated 384-well samples appear to be the least affected by the poor RNA ratio and low concentration readings. This is most likely due to the cells receiving a PBS wash after removing the resazurin dye, potentially removing some of the contaminant(s).

Once the RNA was reverse transcribed into cDNA, the A260/230 ratio appeared to rectify itself, supporting the suspicion that contaminants were present in the RNA samples and were lost or broken down during the reverse transcription process. The poor-quality RNA did not appear to be read at an accurate concentration and this explains why there is no direct correlation between RNA and cDNA quantities when the same quantity was reverse transcribed.

Amplifications efficiencies of less than 90% suggest suboptimal test conditions such as primer annealing temperature or primer concentrations. Amplification efficiencies above 100% suggest polymerase inhibition. This may be due to excessive amounts of nucleic acid or contaminants from the original sample such as heparin, haemoglobin or sodium acetate although this is less likely from a 2D cell culture environment. More likely is contamination from the RNA isolation step such as phenol, ethanol, guanidine hydrochloride or SDS or from the use of some plasticware (Fleige and Pfaffl, 2006).

### 6.3.3 Potential Reference Gene Interference by Venom Proteins

The $2^{-\Delta\Delta Ct}$ method was used to analyse the qPCR data. This is a relative quantification method which gives a value of gene transcription of a treatment group relative to the transcription of a reference gene (Livak and Schmittgen, 2001). It is therefore crucial that genes used as reference genes in qPCR studies are stably transcribed. Using an unstable reference gene - one for which the relative quantity changes between biological groups or following treatment – are unsuitable for qPCR as they cause misleading results when the relative expressions of 'test' genes are calculated. It was investigated whether one or more of the reference genes selected for this study were affected by treatment with any of the cardiotoxin samples. It has been observed previously that Actin may be affected by treatment of theraphosid venoms (Reed, 2012), neurotoxic peptides and $PLA_2$ present in the venom of *Crotalus durissus terrificus* pit viper (Sampaio *et al.*, 2006) and BJcuL, a lectin found in the venom of *Bothrops jararacussu* pit viper (Nolte *et al.*, 2012). Ahluwalia *et al.* (2015) suggest both an increase in actin and a decrease in tubulin expression levels in MDA-MB-231 breast cancer cells following exposure to drCT-I, a toxic peptide isolated from *Daboia russelli* venom. Tubulin expression has also been shown to be reduced following exposure from *Rhinella* spp. toad poison (Abdel-Ghani *et al.*, 2019) and downregulated by *Naja atra* cobra venom (Yan *et al.*, 2017). Conversely, tubulin levels have also been shown to increase following exposure to *Chilobrachys jingzhao* spider venom (Li *et al.*, 2018b). These examples demonstrate the ability of venom (or other compounds) to interfere with the relative expression of genes which are normally considered to be stably transcribed. It is for this reason that actin and tubulin were not selected as control genes for this study.

Out of the remaining potential reference genes, GAPDH showed the best amplification efficiency. 18S showed amplification efficiencies of between 86-90% whilst GAPDH showed 94.8%, indicating the better suitability of this gene as a reference gene from which to calculate the $2^{-\Delta\Delta Ct}$ values of the test genes. 18S still acted as an effective baseline gene, with its $2^{-\Delta\Delta Ct}$ values always ranging from 0.5-2.

### 6.3.4 The Effect on Genes of Interest

The following genes were significantly upregulated:

#### *6.3.4.1 SMAD4 and WNT1*

The SW620 cell line has an extremely low intrinsic SMAD4 expression as is the nature of CRC. The $2^{-\Delta\Delta Ct}$ values of SMAD4 increased in SW620 cells after 24 hours of *Naja pallida* venom sample exposure, suggesting that the venom sample is causing an upregulation in the expression of the SMAD4 gene. The *Naja nigricincta* and *Naja nubiae* samples caused upregulation at both time points suggesting this is both a fast onset and sustained gene regulation effect.

The SMAD4 signalling pathway is linked to other cellular signalling pathways including the WNT pathway (Zhao *et al.*, 2018). An upregulation of both the SMAD4 and the WNT1 gene is observed in the SW620 cell line after 24h when exposed to *Naja nigricincta* and *Naja nubiae* venoms. The only venom/time combination that does not experience an upregulation in SMAD4 expression is 6 hour *Naja nigricincta* venom treated cells. This venom at 6 hours exposure causes a down regulation in WNT1, which could be related to this lack of SMAD4 response. The increase in SMAD4 could influence the expression of WNT1, or vice versa.

Patients whose tumours express SMAD4 have a significantly better prognosis than those who have SMAD4 deficient tumours, and thus the SMAD4 restorative nature of these venom proteins could prove beneficial in treatment of patients who have SMAD4 deficient tumours (Liu, 2001).

The BxPC-3 cell line has a homozygous deletion in the SMAD4 gene and thus does not produce functional SMAD4 protein and was therefore not tested in qPCR (Deer *et al.*, 2010).

#### *6.3.4.2 Tp53*

Mutations in the Tp53 gene are extremely common in both PC and CRC, usually causing reduction in the apoptosis mechanism (Tuveson and Neoptolemos, 2012, Gong *et al.*, 2017). Tp53 was significantly upregulated in BxPC-3 cells when exposed to *Naja nigricincta* and *Naja nubiae* venom samples for 24h. This suggests a potential utility of these venom samples in the restoration of function of the Tp53 gene. Tp53 restoration can lead to a decrease in cell proliferation, arrest of cell cycle and apoptosis in tumour cells (Lozano, 2019). This would have large value in the potential treatment of both PC and CRC which contain Tp53 inactivation and

mutations leading to dysfunctional protein expression respectively (Liu *et al.*, 2018, Malhotra *et al.*, 2021).

#### *6.3.4.3 EGFR*

An upregulation of the EGFR gene was observed in the BxPC-3 cell line following 6 hours of treatment. This increased expression level could be caused by activation of stress pathways following exposure to potentially cytotoxic stressors (Tan *et al.*, 2016). This could have the unintended effect of protecting the cells against growth suppression and apoptosis that could potentially be induced by changes in regulation of other important cancer related genes (Cao *et al.*, 2011). This upregulation is only observed after 6 hour treatment time, so appears to be an acute stress response rather than a long-lived one. The other members of the ErbB family of genes appear mostly unaffected by exposure to the venom proteins except for downregulation of the HER2 gene in the BxPC-3 cell line when exposed to the *Naja nigricincta* venom sample for 6 hours. If proved to downregulate HER2 in follow-up studies, this effect could prove beneficial to other cancers types that are caused by an overexpression of HER2, such as breast cancers and some gastric, lung and bladder cancers (Oh and Bang, 2020).

The following genes were significantly downregulated:

#### *6.3.4.4 MMP9*

All venom protein samples caused down regulation in BxPC-3 cells. The effect of the N.pal_i17r2 sample can be observed in the 6-hour sample, whilst the N.nub_i17r2 and N.nct_i18r2 samples can be observed in the 24-hour sample.

The *Naja pallida* venom sample did not generate a Cq value in 2 out of 3 samples in the first experiment and in all samples in the repeat experiment, This could suggest that by the 24 hour time point, the MMP9 gene had been downregulated so extensively that it could no longer be detected, although this would need to be investigated further by repeating the qPCR experiment (using a higher number of cycles to maximise chance of detecting gene expression) and using an orthogonal assay. Suitable alternative assay formats for assessing gene expression include Northern blotting (for RNA), Southern blotting (for DNA), next generation sequencing or DNA microarray. Protein-based methods such as Western blotting or enzyme-linked immunosorbent assay (ELISA) could also be used to compare the generation of the MMP9 protein in treated samples compared to untreated control cells.

MMP9 inactivation through the p38 MAPK and PI3K/AKT pathways inhibit migration and invasion of breast cancer cells (Jokhio and Ansari, 2005). In addition, transfection of a non-invasive CRC cell line with MMP9 lead to increased migration and invasion, demonstrating the

relationship between MMP9 level and metastatic potential (Murray *et al.*, 2004). This suggests that downregulation of the MMP9 gene could result in decreased invasiveness and reduce the likelihood of metastases in PC cells.

Other researchers have investigated the effect of venom of MMP expression levels. Al-Asmari *et al*. (2016) researched the ability of three undisclosed scorpion venoms to inhibit cell motility and formation of colonies in colorectal cancer cell lines, possibly through a reduction in MMP signalling. Chlorotoxin from scorpion venom has been shown to reduce the expression of MMP2 and inhibit its enzymatic effect, reducing its invasive ability (Deshane *et al.*, 2003). Interestingly, the same effect was not observed for MMP-1, -3 or -9. Cohen *et al*. (2018) investigated the effect of the peptide Chlorotoxin utilised from *Leiurus quinquestriatus* venom on glioblastomas. They found that Chlorotoxin specifically acted on MMP-2, decreasing its expression and inhibiting its activity, reducing the invasiveness of the glioma cells. It has been suggested that tubulin is downregulated following injection of *Naja atra* (Chinese cobra) venom which subsequently leads to upregulation of MMPs (Yan *et al.*, 2017). In this study, tubulin was observed to be upregulated following treatment with cytotoxin samples, and thus it may follow that this is correlated with the downregulation of MMP-9 observed.

#### *6.3.4.5 BCL-2*

Bcl-2 inhibits apoptosis so its downregulation encourages apoptosis and cell death. Bcl-2 downregulation was observed for BxPC-3 cells exposed to *Naja pallida* sample for 6 hours and SW620 cells exposed to *Naja nigricincta* venom for 6 hours. This suggests that these venom proteins induce apoptosis in the tested cell lines. Previous studies have investigated the ability of cardiotoxins to induce apoptosis via ROS- independent mitochondrial dysfunction pathway in a caspase dependent mechanism involving BAX/BCl-2 in various cancer cells (Debnath *et al.*, 2010). This study also found a low intrinsic cytotoxicity on normal human leukocyte cells, suggesting this could be a mechanism specific to cancer cells.

Derakhshani *et al*. (2020) performed a similar qPCR experiment, exposing MCF-7 breast cancer cells to 4 µg/ml of recombinant cytotoxin II originating from *N. oxiana* for 24h. They found that this CT caused a significant upregulation in Caspases 3, 8, 9 and 10 as well as p53 and Bax. A significant reduction in Bcl-2 and MMP-3 and -9 reduction was also observed. This study observed changes in gene regulation for genes which were not observed in this study. Possible reasons for the discrepancies between this study and the literature include the use of a different cell line from a different type of cancer, the CT of interest being recombinantly expressed and from a different species, and that the literature study used a low dose (4 µg/ml) for a prolonged period (24h) whilst this study used an acute higher dose (40 µg/ml for 2h). The additional

exposure time in the literature study may have affected the regulation of genes which are slower to be activated, or those which are further down in a signalling cascade of an earlier affected gene.

It has also been suggested that overexpression of Bcl-2 contributes to chemoresistance, reducing the efficacy of existing drug therapies against various cancers (Mani *et al.*, 2015). Therefore the targeted downregulation of Bcl-2 shown by these venom proteins could provide an interesting adjuvant therapy to complement existing chemotherapeutic treatments for a variety of Bcl-2 overexpressing cancers such as PC or CRC.

#### *6.3.4.6 CDKN2A (excluded)*

SW620 cells have complete, biallelic methylation of the promoter for the CDKN2A gene, repressing gene expression and the subsequent production of $p16^{INK4a}$ protein (Bian *et al.*, 2002, Melcher *et al.*, 2007). BxPC-3 cells do not express the $P16^{INK4a}$ protein, despite being wild type for the CDKN2A gene (Deer *et al.*, 2010). It would have been interesting to investigate whether the venom proteins had an effect on CDKN2A by removal of some promotor methylation or another mechanism, however, it was not possible to validate the CDKN2A gene for qPCR analysis due to the lack of cDNA amplification. To be included in future studies, the CKDN2A primer pair could be validated against a cell line that expresses measurable levels of the gene and then tested in cell lines such as BxPC-3 and SW620 included in this study to see if any gene expression is recovered following drug (or in this case venom protein) treatment.

### 6.3.5 Further Work

It should be emphasised that this was a pilot study. It was not possible to collect multiple biological repeats of the samples to repeat the qPCR and perform statistical analysis on the data. Genes that are highlighted here as interesting should be investigated further by repeating the venom exposure of the cells and performing qPCR again using the primers that are of most interest. At this point, further cell lines including other cancerous cell lines and healthy pancreatic and colorectal control cell lines could be included in the study to compare the gene regulation when exposed to the same venom samples. Modifications could also be made to the protocol and all genes reassessed, such as altering the venom exposure time to be able to assess changes in gene modification not captured here due to exposure time being too short or too long for that particular gene.

Ideally, each gene would have had three independent biological repeats from different runs would have been performed to assess the biological variability as well as the technical variability when testing the same biological sample multiple times. Sadly, access to limited venom samples meant that full biological repeats could not be performed, only technical repeats. In addition, it

was not possible to perform standard curves for every gene in every run. Although standard curves were performed for the reference genes to ensure appropriate amplification efficiencies were achieved, Yuan *et al.* (2006) suggest that amplification efficiency should be calculated for each gene in each sample as there may be subtle differences when different cDNA samples are used. Since the $2^{-\Delta\Delta CT}$ analysis is heavily reliant on samples having comparable amplification efficiencies, this could contribute towards the reason some genes did not show anticipated changes in gene regulation based on similar studies, and why some genes did not show the same changes in gene regulation when the experiment was repeated (Livak and Schmittgen, 2001). If this experiment were to be repeated, samples should be tested more robustly including running standard curves for each gene of interest for each different cDNA sample. Rao *et al*. (2013) suggest a method which calculates individual amplification efficacy corrected values which can be applied to each sample and give a more robust estimation of gene regulation changes between samples. If running standard curves for each gene and cDNA combination as previously suggested, this method could make conclusions about relative gene expression more reliable and improve the understanding of the effect the venom samples may be having on genes of interest. It would also be advisable to go back and re-optimise and/or validate the primers of the genes of interest which did not fully pass the quality control presented in section 6.2.1.1 to ensure they are suitable for the study and provide robust and reliable results.

Attempts were made to determine whether the selected cytotoxins were exerting their effects through an apoptotic mechanism of action by including genes such as Caspase-3 and -7 and members from the Bcl-2 family. No change in gene regulation was observed for either of the caspases or for BAX or BAD, however a significant downregulation of Bcl-2 gene was observed in a small number of venom exposed samples, detailed above. This suggests mitochondrial apoptosis could be involved in the mechanism of action of some of these cardiotoxin proteins (Yang *et al.*, 2020a).

Other studies have suggested that CTs act on apoptotic pathways and have minimal effects on normal cell types (Ebrahim *et al.*, 2015), however future work should attempt to further determine whether the venom components are acting through an apoptotic or necrotic pathway. This could be achieved usually specific dyes such as by Abcam (https://www.abcam.com/apoptosis-necrosis-assay-kit-blue-green-red-ab176749.html, accessed 26/06/2021) using flow cytometry and fluorescence microscopy, or a more high throughput method such as Promega's RealTimeGlo™ Annexin V assay (https://www.promega.co.uk/products/cell-health-assays/apoptosis-assays/realtime-glo-

annexin-v-apoptosis-assay, accessed 26/06/2021) which utilises both fluorescence and luminescence read outs in small volume microtiter plates.

Overall, this pilot study suggests some interesting starting points for further investigation of the effects the selected venom samples are having on PC and CRC cell lines. Further work includes repeating the qPCR using genes which showed changes in regulation, collecting multiple biological repeats, and including additional cell lines to include non-cancerous pancreatic and colorectal cell lines.

# Chapter 7 - General Discussion

The aim of this thesis was to screen a selection of cobra venoms and assess the ability of selected venom proteins to inhibit the growth of BxPC-3 pancreatic cancer and SW620 colorectal cancer cells.

Firstly, a panel of 19 different cobra venoms were screened for cytotoxicity against the two cell lines of interest using a resazurin-based cytotoxicity screen optimised in Chapter 3. 18 of these cobra species originated from the *Naja* genus, split into four categories depending on their location and whether they had developed the 'spitting' trait - African spitting, African non-spitting, Asian spitting and Asian non-spitting - plus *Ophiophagus hannah*, the "King cobra" as the 19$^{th}$ species. At high concentrations, cobra venoms from all four of the different categories caused significant levels of cell inhibition in both cell lines. When screened at a lower concentration, however, African spitting cobra venoms were the only of the four categories which caused significant levels of inhibition in the SW620 cell line. This difference in cell growth inhibition is likely due to the relative abundance and increased potency of cytotoxins present in spitting cobra venom compared to non-spitting cobra venom (Feofanov *et al.*, 2004). Spitting cobra venom contains a high proportion of three-finger toxins and $PLA_2$ enzymes, as well as other cytotoxic and neurotoxic components potentially contributing to the observed activity (Mendez *et al.*, 2011). Despite both African and Asian species convergently evolving the spitting trait, African species were exposed to the selective pressure of predatory hominids for a significantly longer time period, whereby fitness increasing mutations in the gene pool had a greater chance of manifesting (Kazandjian *et al.*, 2021, Wüster *et al.*, 2007). African spitting cobra venoms in particular have been studied to assess their toxin composition (Petras *et al.*, 2011, Mendez *et al.*, 2011, Katali *et al.*, 2020, Pollakova *et al.*, 2021) and how they compare to non-spitting cobras and those from Asia (Feofanov *et al.*, 2004, Wüster and Broadley, 2004, Pollakova *et al.*, 2021, Adamude *et al.*, 2021).

To investigate which venom components were responsible for the selective cell growth inhibition effects, a two-dimensional HPLC process was optimised and run in Chapter 4. HPLC separation was performed on the five African spitting cobras of interest: *Naja mossambica, Naja nigricincta, Naja nigricollis, Naja nubiae* and *Naja pallida*, resulting in 303 purified venom fractions. Each of the collected HPLC fractions were again tested on the two cancer cell lines, this time in a miniaturised 384-well screening format, to assess which of the fractions conferred the desired activity. Six venom fractions - N.mos_i26r2, N.nct_i18r2, N.nub_i17r2, N.pal_i15r4,

N.pal_i17r2 and N.pal_i18r2 - were selected to run as dose response curves as they had significant cytotoxic activity against the BxPC-3 cell line and had a large enough protein yield (at least 80 µg total protein) to be able to perform the dose response curve and all downstream assays.

In Chapter 5, four of the active fractions were partially identified using protein mass spectroscopy (MS) and concluded to each be cytotoxins (CTs). CTs are 60 amino acid proteins found in high abundance in snake venoms, particularly cobras (Konshina *et al.*, 2017). They contain 8 conserved cysteine residues, forming 4 disulphide bridges which provide structural stability and provide the rigid 'finger-like' β-hairpin loops that protrude and can interact with cell membranes (Dubovskii and Utkin, 2015, Konshina *et al.*, 2012). There is a large amount of variation between different CT isoforms leading to a wide range of toxicities between venoms, even from closely related species. These cytotoxins have evolved alongside colourful marking, hooding and, crucially, spitting trait as defensive mechanisms to warn and protect against potential predators (Panagides *et al.*, 2017).

This form of high-throughput cobra venom component screening involving HPLC followed by MS has been previously performed to identify venom proteins contributing specifically to coagulation (Slagboom *et al.*, 2020). It is also described as the basis of activity-based venom peptide discovery which may then include either chemical synthesis or recombinant expression of the identified protein (Prashanth *et al.*, 2017). This form of screening allows many purified venom toxins to be rapidly assessed for the treatment of snakebite or as potential therapeutic agents for a range of human and animal diseases. The technique can be applied to the wider field of natural products such as plant matter, animal and plant poisons, algae, microbial toxins, etc. It should be noted that although the venom of snakes of medically importance have been the focus of the majority of studies to date, components from venom from less harmful snakes still provide a large library of biomolecules potentially useful for treatment of medical conditions that may be present in lower quantities and are easily overlooked (Vonk *et al.*, 2011). There are multiple studies describing the effectiveness of cobra venoms and their components on a variety of cancer types, many of which are described in the introductory chapter.

Natural products have long been a rich source of potential drugs with 1,881 drugs for a variety of conditions receiving approval between 1981-2019 (Newman and Cragg, 2020). Indications for drug entities derived from natural products include anticancer, antihypertensives, antidiabetics, antidepressants and analgesics to name a few, as well as a range of antimicrobials such as antibiotics, antifungals, antiparasitics and antivirals (Newman and Cragg, 2020). The natural world provides a wealth of bioactive components evolved over millions of years by a huge

variety of organisms to thrive in their specific environments. With the gradual movement away from small molecules as the primary source of new drug molecules, there has been a movement back towards natural products and their components (Harvey *et al.*, 2015).

There is an abundant but almost entirely unexplored library of bioactive components found in venoms which could be utilised for a range of diseases and medical conditions. Many of these components are already perfectly evolved to bind to biological targets of interest involved in these medical conditions. In fact, from the discovery of morphine in poppies to isolation of penicillin from mould, the pharmaceutical industry itself was born out of discoveries from the natural world (Rishton, 2008). Study of natural products saw a decrease in popularity in the previous two to three decades, possibly due to concerns about the reproducibility of different isolations of natural extracts and concerns about the storage and handling of potentially unstable compounds contained within natural products when performing HTS (Harvey *et al.*, 2015).

Natural products in drug discovery are, however, making a recent resurgence with the use of phenotypic screening and genomic analysis (Harvey *et al.*, 2015). This resurgence is complimented by the US National Cancer Institute (NCI) Natural Product Repository housing almost a quarter of a million plant, microbial and marine extracts available to researchers worldwide (Thornburg *et al.*, 2018). The organisation has also tried to facilitate handling and HTS by providing a free, publicly accessible library of over one million fractions derived from these natural product extracts, pre-plated in 384-well format (Thornburg *et al.*, 2018), (https://ccr.cancer.gov/news/horizons/article/natural-products).

One key factor to be considered when dealing with natural products, in particular peptides and proteins, is their eventual delivery method into the body. If these venom proteins were to be developed into potential drug molecules, these factors would need to be borne in mind. Oral delivery of peptide- and protein-based drugs such as venoms can be a challenge due to pharmacokinetic (absorption, distribution, metabolism and excretion, or 'ADME') difficulties. Peptides are subject to harsh digestive processes including exposure to proteolytic enzymes and are not readily absorbed across the intestinal mucosa. Following this, they undergo first-pass metabolism in the liver before entering circulation which may inactivate any remaining viable peptide molecules. Injected venom overcomes some of the difficulties usually associated with protein-based drug delivery, however proteinaceous venom components are still susceptible to degradation by proteolysis. Study of ADME properties is very important during drug discovery to improve drug efficacy and safety profiles (Tibbitts *et al.*, 2016). Animal venoms specifically have evolved mechanisms, such as networks of disulphide bridges, including cysteine knots,

which confer greater protein stability and the ability to bypass some of the ADME processes due to the presence of proteases in many prey animals and in the venom itself (Herzig and King, 2015). Many venom peptides possess cell penetrating properties which allow membrane penetration and cell uptake, or this can often be engineered in to molecules that do not already have these properties (Radis-Baptista, 2021). Some ADME limitations for use of proteins as drugs can be overcome by encapsulating the drug molecules in protective delivery systems, such as liposomes. Liposomal delivery of contortrostatin, a disintegrin from *Agkistrodon contortrix* (southern copperhead) venom has been shown to increase the circulating half-life of the drug, help to avoid immune detection and allow the drug to accumulate at the tumour site (Swenson *et al.*, 2004).

The next factor to consider when working with proteinaceous natural products is sustainably obtaining sufficient quantities of protein for development, testing and supply. Traditionally, venoms are either obtained by 'milking' the venomous animal, or by sacrificing the animal and harvesting the venom glands. In the majority of cases, venom can be harvested in a recoverable way such that the animal is not sacrificed, however the principle of the 3Rs of animal research - replacement, reduction and refinement - serve to ensure that the use of animals in scientific research is avoided as much as possible and only used when suitable alternatives are lacking (Wurbel, 2017). In addition to these ethical considerations when it comes to the laboratory, there is also an ongoing crisis threatening wild venomous species due to the popular trade of exotic pets and illegal scorpion farms (Zamani *et al.*, 2021). As more species are being protected by the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) list (https://checklist.cites.org/#/en accessed 27/06/2021), it will become more difficult to access species whose venom could be valuable to medical research. A sustainable method for producing large quantities of venom to be used for medical research and venomics studies is needed if this field of study is to have a lasting future.

Recombinantly produced venom provides an excellent alternative to repeated 'milking' or venom gland removal to obtain venom. There have been several studies showing recombinantly produced venom using *E. coli*-based expression systems. Venoms expressed this way may be post-translationally modified (Shulepko *et al.*, 2017) or have fusion tags added to the DNA sequence to assist correct folding during expression which may be enzymatically cleaved later (Turchetto *et al.*, 2017). Other efforts in this field include the production of *in-vitro* venom gland expression systems which provide long-term culture primary cells which secrete up to 63 µg/ml of venom in culture medium (Yamanouye *et al.*, 2006). More recently, venom gland organoid systems have been established, able to secrete functionally active venoms displaying the same

biological activity of the crude venom from the snakes the venom glands originated from (Post *et al.*, 2020). These organoids also secrete disulphide isomerase, and essential enzyme for ensuring the correct disulphide folding of the venom proteins. Although these venom gland culture examples require sacrificing snakes in order to obtain the venom glands required for initiating the culture, they reflect exciting advancements in the field which, once fully established, could significantly reduce the requirement for animal use in the field of venomics and for the medical study of venoms and antivenom production.

Finally, chapter 6 was a pilot study to investigate the effect of the selected venom cytotoxins on gene regulation of some key genes involved in PC and CRC. This study found that SMAD4, Tp53, WNT1 and EGFR genes were upregulated following addition of venom and MMP9 was downregulated. Future work would involve investigating these affected genes further by re-testing using qPCR as well as orthogonal techniques to confirm these findings. If confirmed in repeated experiments, the effects shown in this study against the genes mentioned above could give a promising starting point for the development of a novel cancer therapy, diagnostic tool or pharmacological tool compound to aid further research in this area. Cancer therapy using novel compounds is particularly important in the light of emerging acquired drug resistance (Michaelis *et al.*, 2019).

## 7.1 Limitations and Further Work

The main limitation of this study was in the choice of cell line studied. Unfortunately, following STR profiling, it became apparent that the Mia PaCa-2 PC cells selected for its distinct genetic profile to complement the BxPC-3 PC cells were in fact mislabelled SW620 CRC cells. This meant that the study was performed with two different cancer types, rather than distinct subtypes of the same cancer. Cell line authentication has become more widespread in the scientific community in recent years due to stricter requirements for grant funding and publication to improve transparency and rigor in scientific research and to ensure reproducibility of findings (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-103.html accessed 30/06/2021). The use of misauthenticated cell lines in scientific research is a problem with significant consequences as is demonstrated in this study, however STR profiling is expensive, time consuming and needs to be repeated regularly to ensure that no mislabelling, contamination or genetic drift has occurred since the last authentication (Mzurikwao *et al.*, 2020). The use of simple to use, affordable cell line identification that may be used regularly within the lab to monitor cell lines in use would provide a valuable complement to STR profiling, helping to limit the negative impact caused by performing large volumes of work on incorrect cell lines. Recent developments of image-based deep learning methodologies suggest this form of routine cell line

monitoring could soon be possible (Mzurikwao *et al.*, 2020). Another recent method utilising MALDI-ToF MS is able to distinguish different cell lines originating from the same organism and modified cell lines from original cell lines from the same genetic origin, such as those modified for drug resistance studies (Povey *et al.*, 2019).

Despite these authentication difficulties, the SW620 cells provided a valuable control cell line function and showed some interesting differences in gene regulation. The findings related to this cell line could be applied to the development of the potential treatment of CRC as well as PC and this application could prove just as valuable as the findings relating to PC. Likewise, other cancer types could be investigated with respect to the findings in this study to assess the effects of the selected venom fractions on other cancer types. A non-cancerous pancreatic control cell line could have provided a valuable comparison and given insight into the specificity of the venom components on the PC cells. Resource limitations prevented obtaining non-cancerous cell lines or another PC comparison cell line which could have greatly enhanced the impact of this study.

Resazurin was used as the main choice of assay for this study to screen multiple venoms and their fractionated components. Resazurin assays provide an effective method for phenotypically screening cells in 2D culture and provide a measure of overall cytotoxicity when exposed to compounds such as venom components. Resazurin has been used for several venom studies previously including assessing the effects of jellyfish venom (Lau *et al.*, 2019), toad poison (Abdelfatah *et al.*, 2019, Ferreira *et al.*, 2013), scorpion venoms (Cota-Arce *et al.*, 2020) and protective effects of molecules against the effects of spider venoms (Paixao-Cavalcante *et al.*, 2007). Resazurin assays do not, however, provide information about mechanism of action of compounds or any effects on downstream processes. Further experiments are required to deconvolute the biological effect.

Target-based screening could have been used as an alternative method to phenotypic screening if a particular protein known to contribute to cancer had been investigated. This would have enabled the use of biochemical or biophysical methods to assess competitive binding and protein-protein interactions (PPIs). If purified protein was expressed with an appropriate tag, techniques such as time-resolved fluorescence resonance energy transfer (TR-FRET) or fluorescence polarisation (FP) could have been used. TR-FRET is a cheap and effective way of assessing binding events. It measures the energy transfer from a protein conjugated ‘donor’ molecule, most commonly a lanthanide such as Terbium, to an ‘acceptor’ molecule which is usually a fluorescent probe conjugated to a molecule known to bind at the site of interest on the target protein. When a compound binds to the target protein, the ‘acceptor’ probe is

displaced and fluorescence signal decreases (Zhang *et al.*, 2020). FP is another biochemical method used to assess binding to a protein of interest. This method measures the changes in the polarisation of light when a fluorescent probe is excited and uses this to determine protein binding (Simon *et al.*, 2020, Lee *et al.*, 2021). There have been some reference made to the use of FRET and FP in venom screening (Herzig *et al.*, 2020, Prashanth *et al.*, 2017).

There was a limitation to throughput of the resazurin based assays, HPLC and qPCR due to the lack of automation used in this study. To reach higher throughput levels, high precision liquid handling units are required. In industry, 1536-well microtiter plates are routinely used for drug screening, whilst this study was limited to 384-well plate format due to the use of hand-pipetting. This manual process also has a negative impact on assay precision and relies heavily on care taken by the experimenter. Laboratory automation has several advantages including reduced running costs, better efficiency and improved testing quality, however it is expensive to set up and requires space and engineering support (Lippi and Da Rin, 2019).

The mass spectrometry performed in this study was unable to form full matches to the current protein databases. This could be due to limitations in the applied methodology itself but is more likely to be due to the lack of completeness of the databases searched. These proteins may be novel and further investigation of their sequence is warranted. Other cobra venom studies have suffered from the same problem where unknown proteins are isolated and cannot be identified without further experimentation (Malih *et al.*, 2014). Other venom proteins have been further investigated using Edman degradation of the proteins, or next-generation sequencing (NGS) of the transcriptome to determine their novel sequences (Robinson *et al.*, 2017). There is even a Python script now available to assist with precisely annotating snake venom gland transcriptomes (Nachtigall *et al.*, 2021). Other potentially useful methods for venom protein identification are NMR and protein crystallography. NMR has proven to be effective for structural determination of smaller proteins and peptides, however this technique struggles to deconvolute larger, more complex protein structures, particularly those with multiple disulphide bonds present, although this is possible (Walewska *et al.*, 2008). X-ray crystallography has been used to identify numerous venom proteins including scorpion (Dang *et al.*, 2016), sea anemone (Dang *et al.*, 2017) and even cobra cytotoxins (Lee *et al.*, 2014).

Resources were limited when performing the qPCR study and the results were therefore unable to be repeated. Amplification efficiencies could not be assessed for every gene and therefore the $2^{-\Delta\Delta Ct}$ analysis applied may not be entirely accurate. Standard curves should be performed for each gene being tested to improve the robustness of the qPCR analysis. Other techniques could be applied following repeated qPCR results to assess the effect on downstream protein

production. Such techniques include Western Blotting and ELISA which are both regularly employed in venomics and venom studies (Escalante *et al.*, 2021, Mccullough *et al.*, 2020, Paniagua *et al.*, 2020).

## 7.2 Concluding Remarks

The methods employed in this study were able to effectively categorise African spitting cobra venom as significantly more cytotoxic to the selected cancer cell lines compared to Asian spitting and all non-spitting cobra species. The venom of five African spitting cobras were then fractionated by HPLC and made into a screening panel, from which several fractions were selected from the non-cytotoxic venom components. This illustrates how venom can effectively be utilised as a screening library for HTS against potential therapeutic targets in a drug-discovery context. Even if found to be inappropriate as potential drug candidates, venom proteins could have great utility as pharmacological tool compounds, acting as positive controls for other studies to allow the detection of potential therapies.

The selected venom components were then partially identified by MS and found to be cobra cytotoxins. The fact that these components could not be fully matched to the UniProt database suggests these are novel proteins that demonstrate some sequence homology to the cytotoxins in the database but with unique sequence sections. This highlights the diversity of venom components and illustrates that many more venom protein sequences remain to be discovered and characterised.

The cobra cytotoxins detailed in this study have potential utility in the treatment of PC and/or CRC. They showed significant effects against SMAD4, Tp53, WNT1, EGFR and MMP9 genes in qPCR and these findings warrant further investigation to confirm their activity and potential utility in cancer therapy. As both PC and CRC are common and aggressive cancer types with poor morbidity and mortality, novel therapies such as these are desperately required. It is hoped that venom proteins such as those identified in this study can make a small but significant contribution to the treatment of these devastating diseases to help patients and their families in the near future.

# References


Abdel-Ghani, L. M., Rahmy, T. R., Tawfik, M. M., Kaziri, I., Al-Obaidi, A., Rowan, E. G., Plevin, R. & Abdel-Rahman, M. A. 2019. Cytotoxicity of Nubein6.8 peptide isolated from the snake venom of Naja nubiae on melanoma and ovarian carcinoma cell lines. *Toxicon,* 168**,** 22-31.

Abdelfatah, S., Lu, X., Schmeda-Hirschmann, G. & Efferth, T. 2019. Cytotoxicity and antimitotic activity of Rhinella schneideri and Rhinella marina venoms. *J Ethnopharmacol,* 242**,** 112049.

Adamude, F. A., Dingwoke, E. J., Abubakar, M. S., Ibrahim, S., Mohamed, G., Klein, A. & Sallau, A. B. 2021. Proteomic analysis of three medically important Nigerian Naja (Naja haje, Naja katiensis and Naja nigricollis) snake venoms. *Toxicon,* 197**,** 24-32.

Ahluwalia, S., Sawant, M. G., Shah, N. & Chowdhary, A. 2015. *RE: Experimental Evaluation of Antitumor Effect of Russell' s Viper Venom on Breast Cancer Cell Line - MDA-MB-231.*

Ahmed, S., Bradshaw, A. D., Gera, S., Dewan, M. Z. & Xu, R. 2017. The TGF-beta/Smad4 Signaling Pathway in Pancreatic Carcinogenesis and Its Clinical Significance. *J Clin Med,* 6.

Al-Asmari, A. K., Islam, M. & Al-Zahrani, A. M. 2016. In vitro analysis of the anticancer properties of scorpion venom in colorectal and breast cancer cell lines. *Oncol Lett,* 11**,** 1256-1262.

Al-Hajeili, M., Azmi, A. S. & Choi, M. 2014. Nab-paclitaxel: potential for the treatment of advanced pancreatic cancer. *OncoTargets and therapy,* 7**,** 187-92.

Al-Nasiry, S., Geusens, N., Hanssens, M., Luyten, C. & Pijnenborg, R. 2007. The use of Alamar Blue assay for quantitative analysis of viability, migration and invasion of choriocarcinoma cells. *Hum Reprod,* 22**,** 1304-9.

Albury, T. M., Pandey, V., Gitto, S. B., Dominguez, L., Spinel, L. P., Talarchek, J., Klein-Szanto, A. J., Testa, J. R. & Altomare, D. A. 2015. Constitutively active Akt1 cooperates with KRas(G12D) to accelerate in vivo pancreatic tumor onset and progression. *Neoplasia,* 17**,** 175-82.

Ali, S. A., Hamid, F., Abbasi, A., Zaidi, Z. H. & Shehnaz, D. 1999. Pharmacological effects of the leaf-nosed viper snake (Eristocophis macmahoni) venom and its HPLC fractions. *Toxicon,* 37**,** 1095-107.

Alqahtani, M. S., Kazi, M., Alsenaidy, M. A. & Ahmad, M. Z. 2021. Advances in Oral Drug Delivery. *Front Pharmacol,* 12**,** 618411.

Alsubai, J., Matters, G. L., Mcgovern, C. O., Liao, J., Gilius, E. L. & Smith, J. P. 2016. Germline Mutation of the CCK Receptor: A Novel Biomarker for Pancreas Cancer. *Clinical and Translational Gastroenterology,* 7**,** e134-e134.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-10.

Andrews, P. 1965. The gel-filtration behaviour of proteins related to their molecular weights over a wide range. *Biochem J,* 96**,** 595-606.

Anoopkumar-Dukie, S., Carey, J. B., Conere, T., O'sullivan, E., Van Pelt, F. N. & Allshire, A. 2005. Resazurin assay of radiation response in cultured cells. *Br J Radiol,* 78**,** 945-7.

Antunes, T. C., Yamashita, K. M., Barbaro, K. C., Saiki, M. & Santoro, M. L. 2010. Comparative analysis of newborn and adult Bothrops jararaca snake venoms. *Toxicon,* 56**,** 1443-58.

Ariston Gabriel, A. N., Wang, F., Jiao, Q., Yvette, U., Yang, X., Al-Ameri, S. A., Du, L., Wang, Y. S. & Wang, C. 2020. The involvement of exosomes in the diagnosis and treatment of pancreatic cancer. *Mol Cancer,* 19**,** 132.

Arslan, A. A., Helzlsouer, K. J., Kooperberg, C., Shu, X.-O., Steplowski, E., Bueno-De-Mesquita, H. B., Fuchs, C. S., Gross, M. D., Jacobs, E. J., Lacroix, A. Z., Petersen, G. M.,

Stolzenberg-Solomon, R. Z., Zheng, W., Albanes, D., Amundadottir, L., Bamlet, W. R., Barricarte, A., Bingham, S. A., Boeing, H., Boutron-Ruault, M.-C., Buring, J. E., Chanock, S. J., Clipp, S., Gaziano, J. M., Giovannucci, E. L., Hankinson, S. E., Hartge, P., Hoover, R. N., Hunter, D. J., Hutchinson, A., Jacobs, K. B., Kraft, P., Lynch, S. M., Manjer, J., Manson, J. E., Mctiernan, A., Mcwilliams, R. R., Mendelsohn, J. B., Michaud, D. S., Palli, D., Rohan, T. E., Slimani, N., Thomas, G., Tjønneland, A., Tobias, G. S., Trichopoulos, D., Virtamo, J., Wolpin, B. M., Yu, K., Zeleniuch-Jacquotte, A. & Patel, A. V. 2010. Anthropometric measures, body mass index, and pancreatic cancer: a pooled analysis from the Pancreatic Cancer Cohort Consortium (PanScan). *Archives of internal medicine,* 170**,** 791-802.

Attarde, S. S. & Pandit, S. V. 2017. Cytotoxic activity of NN-32 toxin from Indian spectacled cobra venom on human breast cancer cell lines. *BMC Complement Altern Med,* 17**,** 503.

Attri, J., Srinivasan, R., Majumdar, S., Radotra, B. D. & Wig, J. 2005. Alterations of tumor suppressor gene p16INK4a in pancreatic ductal carcinoma. *BMC gastroenterology,* 5**,** 22-22.

Aubert, L., Nandagopal, N., Steinhart, Z., Lavoie, G., Nourreddine, S., Berman, J., Saba-El-Leil, M. K., Papadopoli, D., Lin, S., Hart, T., Macleod, G., Topisirovic, I., Gaboury, L., Fahrni, C. J., Schramek, D., Meloche, S., Angers, S. & Roux, P. P. 2020. Copper bioavailability is a KRAS-specific vulnerability in colorectal cancer. *Nat Commun,* 11**,** 3701.

Avella, I., Barajas-Ledesma, E., Casewell, N. R., Harrison, R. A., Rowley, P. D., Crittenden, E., Wüster, W., Castiglia, R., Holland, C. & Van Der Meijden, A. 2021. Unexpected lack of specialisation in the flow properties of spitting cobra venom. *J Exp Biol,* 224.

Babaoglu, K., Simeonov, A., Irwin, J., Nelswon, M., Feng, B., Thomas, C., Cancian, L., Costi, M., Maltby, D., Jadhav, A., Inglese, J., Austin, C. & Shoichet, B. 2008. Comprehensive Mechanistic Analysis of Hits from HighThroughput and Docking Screens against β-Lactamase. *J Med Chem,* 51**,** 2502-2511.

Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A.-M., Gingras, M.-C., Miller, D. K., Christ, A. N., Bruxner, T. J. C., Quinn, M. C., Nourse, C., Murtaugh, L. C., Harliwong, I., Idrisoglu, S., Manning, S., Nourbakhsh, E., Wani, S., Fink, L., Holmes, O., Chin, V., Anderson, M. J., Kazakoff, S., Leonard, C., Newell, F., Waddell, N., Wood, S., Xu, Q., Wilson, P. J., Cloonan, N., Kassahn, K. S., Taylor, D., Quek, K., Robertson, A., Pantano, L., Mincarelli, L., Sanchez, L. N., Evers, L., Wu, J., Pinese, M., Cowley, M. J., Jones, M. D., Colvin, E. K., Nagrial, A. M., Humphrey, E. S., Chantrill, L. A., Mawson, A., Humphris, J., Chou, A., Pajic, M., Scarlett, C. J., Pinho, A. V., Giry-Laterriere, M., Rooman, I., Samra, J. S., Kench, J. G., Lovell, J. A., Merrett, N. D., Toon, C. W., Epari, K., Nguyen, N. Q., Barbour, A., Zeps, N., Moran-Jones, K., Jamieson, N. B., Graham, J. S., Duthie, F., Oien, K., Hair, J., Grützmann, R., Maitra, A., Iacobuzio-Donahue, C. A., Wolfgang, C. L., Morgan, R. A., Lawlor, R. T., Corbo, V., Bassi, C., Rusev, B., Capelli, P., Salvia, R., Tortora, G., Mukhopadhyay, D., Petersen, G. M., Munzy, D. M., Fisher, W. E., Karim, S. A., Eshleman, J. R., Hruban, R. H., Pilarsky, C., Morton, J. P., Sansom, O. J., Scarpa, A., Musgrove, E. A., Bailey, U.-M. H., Hofmann, O., Sutherland, R. L., Wheeler, D. A., Gill, A. J., Gibbs, R. A., Pearson, J. V., Waddell, N., et al. 2016. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature,* 531**,** 47-52.

Bairoch, A. & Apweiler, R. 1996. The SWISS-PROT protein sequence data bank and its new supplement TREMBL. *Nucleic Acids Res,* 24**,** 21-5.

Bakhle, Y. S. 2020. How ACE inhibitors transformed the renin-angiotensin system. *Br J Pharmacol,* 177**,** 2657-2665.

Bartel, D. P. 2004. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell,* 116**,** 281-297.

Bartholomeusz, C., Yamasaki, F., Saso, H., Kurisu, K., Hortobagyi, G. N. & Ueno, N. T. 2011. Gemcitabine Overcomes Erlotinib Resistance in EGFR-Overexpressing Cancer Cells through Downregulation of Akt. *Journal of Cancer,* 2**,** 435-42.

Bartsch, D., Barth, P., Bastian, D., Ramaswamy, A., Gerdes, B., Chaloupka, B., Deiss, Y., Simon, B. & Schudy, A. 1999. Higher frequency of DPC4/Smad4 alterations in pancreatic cancer cell lines than in primary pancreatic adenocarcinomas. *Cancer Lett,* 139**,** 43-9.

Batista, C. V., Del Pozo, L., Zamudio, F. Z., Contreras, S., Becerril, B., Wanke, E. & Possani, L. D. 2004. Proteomics of the venom from the Amazonian scorpion Tityus cambridgei and the role of prolines on mass spectrometry analysis of toxins. *J Chromatogr B Analyt Technol Biomed Life Sci,* 803**,** 55-66.

Bayat, A. 2002. Science, medicine, and the future: Bioinformatics. *BMJ,* 324**,** 1018-22.

Bazaa, A., Luis, J., Srairi-Abid, N., Kallech-Ziri, O., Kessentini-Zouari, R., Defilles, C., Lissitzky, J. C., El Ayeb, M. & Marrakchi, N. 2009. MVL-PLA2, a phospholipase A2 from Macrovipera lebetina transmediterranea venom, inhibits tumor cells adhesion and migration. *Matrix Biol,* 28**,** 188-93.

Bazaa, A., Marrakchi, N., El Ayeb, M., Sanz, L. & Calvete, J. J. 2005. Snake venomics: Comparative analysis of the venomproteomes of the Tunisian snakes Cerastes cerastes,Cerastes vipera and Macrovipera lebetina. *Proteomics,* 5**,** 4223-4235.

Belov, A. M., Kozole, J., Bean, M. F., Machutta, C. A., Zhang, G., Gao, E. N., Ghislain, L., Datwani, S. S., Leveridge, M. & Annan, R. S. 2020. Acoustic Mist Ionization-Mass Spectrometry: A Comparison to Conventional High-Throughput Screening and Compound Profiling Platforms. *Anal Chem,* 92**,** 13847-13854.

Berlin, J. D., Catalano, P., Thomas, J. P., Kugler, J. W., Haller, D. G. & Benson, A. B. 2002. Phase III study of gemcitabine in combination with fluorouracil versus gemcitabine alone in patients with advanced pancreatic carcinoma: Eastern Cooperative Oncology Group Trial E2297. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 20**,** 3270-5.

Berman, H., Henrick, K. & Nakamura, H. 2003. Announcing the worldwide Protein Data Bank. *Nat Struct Biol,* 10**,** 980.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. 2000. The Protein Data Bank. *Nucleic Acids Res,* 28**,** 235-42.

Bian, Y. S., Osterheld, M. C., Fontolliet, C., Bosman, F. T. & Benhattar, J. 2002. p16 inactivation by methylation of the CDKN2A promoter occurs early during neoplastic progression in Barrett's esophagus. *Gastroenterology,* 122**,** 1113-21.

Bobaly, B., Mikola, V., Sipko, E., Marta, Z. & Fekete, J. 2015. Recovery of Proteins Affected by Mobile Phase Trifluoroacetic Acid Concentration in Reversed-Phase Chromatography. *J Chromatogr Sci,* 53**,** 1078-83.

Bobály, B., Tóth, E., Drahos, L., Zsila, F., Visy, J., Fekete, J. & Vékey, K. 2014. Influence of acid-induced conformational variability on protein separation in reversed phase high performance liquid chromatography. *Journal of Chromatography A,* 1325**,** 155-162.

Boice, A. & Bouchier-Hayes, L. 2020. Targeting apoptotic caspases in cancer. *Biochim Biophys Acta Mol Cell Res,* 1867**,** 118688.

Bonfim, V. L., Toyama, M. H., Novello, J. C., Hyslop, S., Oliveira, C. R., Rodrigues-Simioni, L. & Marangoni, S. 2001. Isolation and enzymatic characterization of a basic phospholipase A2 from Bothrops jararacussu snake venom. *J Protein Chem,* 20**,** 239-45.

Borkow, G., Gutierrez, J. M. & Ovadia, M. 1993. Isolation and characterization of synergistic hemorrhagins from the venom of the snake Bothrops asper. *Toxicon,* 31**,** 1137-50.

Borra, R. C., Lotufo, M. A., Gagioti, S. M., Barros, F. D. M. & Andrade, P. M. 2009. A simple method to measure cell viability in proliferation and cytotoxicity assays. *Brazilian oral research,* 23**,** 255-262.

Botes, D. P. & Viljoen, C. C. 1976. The amino acid sequence of three non-curarimimetictoxins from Naja nivea venom. *Biochim Biophys Acta,* 446**,** 1-9.

Botteri, E., Iodice, S., Bagnardi, V., Raimondi, S., Lowenfels, A. B. & Maisonneuve, P. 2008. Smoking and colorectal cancer: a meta-analysis. *JAMA,* 300**,** 2765-78.

Bougis, P., Tessier, M., Van Rietschoten, J., Rochat, H., Faucon, J. F. & Dufourcq, J. 1983. Are interactions with phospholipids responsible for pharmacological activities of cardiotoxins? *Mol Cell Biochem,* 55**,** 49-64.

Bougis, P. E., Marchot, P. & Rochat, H. 1987. In vivo synergy of cardiotoxin and phospholipase A2 from the elapid snake Naja mossambica mossambica. *Toxicon,* 25**,** 427-31.

Bournet, B., Buscail, C., Muscari, F., Cordelier, P. & Buscail, L. 2016. Targeting KRAS for diagnosis, prognosis, and treatment of pancreatic cancer: Hopes and realities. *Eur J Cancer,* 54**,** 75-83.

Bouwmeester, R., Martens, L. & Degroeve, S. 2019. Comprehensive and Empirical Evaluation of Machine Learning Algorithms for Small Molecule LC Retention Time Prediction. *Anal Chem,* 91**,** 3694-3703.

Bramhall, S. R., Schulz, J., Nemunaitis, J., Brown, P. D., Baillet, M. & Buckels, J. a. C. 2002. A double-blind placebo-controlled, randomised study comparing gemcitabine and marimastat with gemcitabine and placebo as first line therapy in patients with advanced pancreatic cancer. *British journal of cancer,* 87**,** 161-7.

Brand, R. 2001. The diagnosis of pancreatic cancer. *Cancer journal (Sudbury, Mass.),* 7**,** 287-97.

Breckeridge, R. & Dufton, M. J. 1987. The structural evolution of cobra venom cytotoxins. *Journal of Molecular Evolution,* 26**,** 274-283.

Brideau, C., Gunter, B., Pikounis, B. & Liaw, A. 2003. Improved statistical methods for hit selection in high-throughput screening. *J Biomol Screen,* 8**,** 634-47.

Bruno, W., Andreotti, V., Bisio, A., Pastorino, L., Fornarini, G., Sciallero, S., Bianchi-Scarra, G., Inga, A. & Ghiorzo, P. 2017. Functional analysis of a CDKN2A 5'UTR germline variant associated with pancreatic cancer development. *PLoS One,* 12**,** e0189123.

Buchsbaum, D. J., Bonner, J. A., Grizzle, W. E., Stackhouse, M. A., Carpenter, M., Hicklin, D. J., Bohlen, P. & Raisch, K. P. 2002. Treatment of pancreatic cancer xenografts with Erbitux (IMC-C225) anti-EGFR antibody, gemcitabine, and radiation. *International journal of radiation oncology, biology, physics,* 54**,** 1180-93.

Burin, S. M., Cacemiro, M. D. C., Cominal, J. G., Grandis, R. A., Machado, A. R. T., Donaires, F. S., Cintra, A. C. O., Ambrosio, L., Antunes, L. M. G., Sampaio, S. V. & De Castro, F. A. 2020. Bothrops moojeni L-amino acid oxidase induces apoptosis and epigenetic modulation on Bcr-Abl(+) cells. *J Venom Anim Toxins Incl Trop Dis,* 26**,** e20200123.

Burin, S. M., Ghisla, S., Ouchida, A. T., Aissa, A. F., Coelho, M. G., Costa, T. R., Marsola, A. P., Pinto-Simoes, B., Antunes, L. M., Curti, C., Sampaio, S. V. & De Castro, F. A. 2016. CR-LAAO antileukemic effect against Bcr-Abl(+) cells is mediated by apoptosis and hydrogen peroxide. *Int J Biol Macromol,* 86**,** 309-20.

Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Dalenberg, K., Duarte, J. M., Dutta, S., Feng, Z., Ghosh, S., Goodsell, D. S., Green, R. K., Guranovic, V., Guzenko, D., Hudson, B. P., Kalro, T., Liang, Y., Lowe, R., Namkoong, H., Peisach, E., Periskova, I., Prlic, A., Randle, C., Rose, A., Rose, P., Sala, R., Sekharan, M., Shao, C., Tan, L., Tao, Y. P., Valasatava, Y., Voigt, M., Westbrook, J., Woo, J., Yang, H., Young, J., Zhuravleva, M. & Zardecki, C. 2019. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res,* 47**,** D464-D474.

Burris, H. A., 3rd, Moore, M. J., Andersen, J., Green, M. R., Rothenberg, M. L., Modiano, M. R., Cripps, M. C., Portenoy, R. K., Storniolo, A. M., Tarassoff, P., Nelson, R., Dorr, F. A., Stephens, C. D. & Von Hoff, D. D. 1997. Improvements in survival and clinical benefit with gemcitabine as first- line therapy for patients with advanced pancreas cancer: a randomized trial. *J. Clin. Oncol.,* 15**,** 2403-2413.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J. & Wittwer, C. T. 2009. The MIQE

guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem,* 55**,** 611-22.

Caldas, C., Hahn, S. A., Da Costa, L. T., Redston, M. S., Schutte, M., Seymour, A. B., Weinstein, C. L., Hruban, R. H., Yeo, C. J. & Kern, S. E. 1994. Frequent somatic mutations and homozygous deletions of the p16 (MTS1) gene in pancreatic adenocarcinoma. *Nat Genet,* 8**,** 27-32.

Canavan, C., Abrams, K. R. & Mayberry, J. 2006. Meta-analysis: colorectal and small bowel cancer risk in patients with Crohn's disease. *Aliment Pharmacol Ther,* 23**,** 1097-104.

Cancer Research, U. K. 2018a. *Bowel Cancer Statistics* [Online]. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer. [Accessed].

Cancer Research, U. K. 2018b. *RE: Pancreatic cancer mortality statistics.*

Cao, X., Zhu, H., Ali-Osman, F. & Lo, H. W. 2011. EGFR and EGFRvIII undergo stress- and EGFR kinase inhibitor-induced mitochondrial translocalization: a potential mechanism of EGFR-driven antagonism of apoptosis. *Mol Cancer,* 10**,** 26.

Carlsson, F. H. H. & Joubert, F. J. 1974. Snake venom toxins The isolation and purification of three cytotoxin homologues from the venom of the forest cobra (Naja melanoleuca) and the complete amino acid sequence of toxin VII1. *Biochimica et Biophysica Acta (BBA) - Protein Structure,* 336**,** 453-469.

Cascinu, S., Berardi, R., Sobrero, A., Bidoli, P., Labianca, R., Siena, S., Ferrari, D., Barni, S., Aitini, E., Zagonel, V., Caprioni, F., Villa, F., Mosconi, S., Faloppi, L., Tonini, G., Boni, C., Conte, P., Di Costanzo, F. & Cinquini, M. 2014. Sorafenib does not improve efficacy of chemotherapy in advanced pancreatic cancer: A GISCAD randomized phase II study. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver,* 46**,** 182-6.

Casewell, N. R., Wüster, W., Vonk, F. J., Harrison, R. A. & Fry, B. G. 2013. *RE: Complex cocktails: The evolutionary novelty of venoms.*

Chakravarthy, A. B., Tsai, C. J., O'brien, N., Lockhart, A. C., Chan, E., Parikh, A., Berlin, J. D. & Merchant, N. 2012. A phase I study of cetuximab in combination with gemcitabine and radiation for locally advanced pancreatic cancer. *Gastrointest Cancer Res,* 5**,** 112-8.

Chapeaurouge, A., Silva, A., Carvalho, P., Mccleary, R. J. R., Modahl, C. M., Perales, J., Kini, R. M. & Mackessy, S. P. 2018. Proteomic Deep Mining the Venom of the Red-Headed Krait, Bungarus flaviceps. *Toxins (Basel),* 10**,** 373.

Che, P., Wang, L. & Li, Q. 2009. The development, optimization and validation of an assay for high throughput antiviral drug screening against Dengue virus. *Int J Clin Exp Med,* 2**,** 363-73.

Chen, J. L., Steele, T. W. & Stuckey, D. C. 2015. Modeling and Application of a Rapid Fluorescence-Based Assay for Biotoxicity in Anaerobic Digestion. *Environ Sci Technol,* 49**,** 13463-71.

Chen, K. C., Liu, W. H. & Chang, L. S. 2010. Taiwan cobra phospholipase A2-elicited JNK activation is responsible for autocrine fas-mediated cell death and modulating Bcl-2 and Bax protein expression in human leukemia K562 cells. *J Cell Biochem,* 109**,** 245-54.

Chen, X. S., Li, L. Y., Guan, Y. D., Yang, J. M. & Cheng, Y. 2016. Anticancer strategies based on the metabolic profile of tumor cells: therapeutic targeting of the Warburg effect. *Acta Pharmacol Sin,* 37**,** 1013-9.

Chen, Y., Muller, J. D., Tetin, S. Y., Tyner, J. D. & Gratton, E. 2000. Probing ligand protein binding equilibria with fluorescence fluctuation spectroscopy. *Biophys J,* 79**,** 1074-84.

Cheng, X., Xu, X., Chen, D., Zhao, F. & Wang, W. 2019. Therapeutic potential of targeting the Wnt/beta-catenin signaling pathway in colorectal cancer. *Biomed Pharmacother,* 110**,** 473-481.

Chien, K. Y., Chiang, C. M., Hseu, Y. C., Vyas, A. A., Rule, G. S. & Wu, W. 1994. Two distinct types of cardiotoxin as revealed by the structure and activity relationship of their interaction with zwitterionic phospholipid dispersions. *J Biol Chem,* 269**,** 14473-83.

Cicenas, J., Kvederaviciute, K., Meskinyte, I., Meskinyte-Kausiliene, E., Skeberdyte, A. & Cicenas, J. 2017. KRAS, TP53, CDKN2A, SMAD4, BRCA1, and BRCA2 Mutations in Pancreatic Cancer. *Cancers (Basel),* 9**,** 42.

Cohen, G., Burks, S. R. & Frank, J. A. 2018. Chlorotoxin-A Multimodal Imaging Platform for Targeting Glioma Tumors. *Toxins (Basel),* 10.

Collisson, E. A., Sadanandam, A., Olson, P., Gibb, W. J., Truitt, M., Gu, S., Cooc, J., Weinkle, J., Kim, G. E., Jakkula, L., Feiler, H. S., Ko, A. H., Olshen, A. B., Danenberg, K. L., Tempero, M. A., Spellman, P. T., Hanahan, D. & Gray, J. W. 2011. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nature medicine,* 17**,** 500-3.

Coma, I., Clark, L., Diez, E., Harper, G., Herranz, J., Hofmann, G., Lennon, M., Richmond, N., Valmaseda, M. & Macarron, R. 2009. Process validation and screen reproducibility in high-throughput screening. *J Biomol Screen,* 14**,** 66-76.

Conlon, J. M., Attoub, S., Musale, V., Leprince, J., Casewell, N. R., Sanz, L. & Calvete, J. J. 2020. Isolation and characterization of cytotoxic and insulin-releasing components from the venom of the black-necked spitting cobra Naja nigricollis (Elapidae). *Toxicon X,* 6**,** 100030.

Conroy, T., Desseigne, F., Ychou, M., Bouché, O., Guimbaud, R., Bécouarn, Y., Adenis, A., Raoul, J.-L., Gourgou-Bourgade, S., De La Fouchardière, C., Bennouna, J., Bachet, J.-B., Khemissa-Akouz, F., Péré-Vergé, D., Delbaldo, C., Assenat, E., Chauffert, B., Michel, P., Montoto-Grillot, C. & Ducreux, M. 2011. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *The New England journal of medicine,* 364**,** 1817-25.

Cota-Arce, J. M., Zazueta-Favela, D., Diaz-Castillo, F., Jimenez, S., Bernaldez-Sarabia, J., Caram-Salas, N. L., Dan, K. W. L., Escobedo, G., Licea-Navarro, A. F., Possani, L. D. & De Leon-Nava, M. A. 2020. Venom components of the scorpion Centruroides limpidus modulate cytokine expression by T helper lymphocytes: Identification of ion channel-related toxins by mass spectrometry. *Int Immunopharmacol,* 84**,** 106505.

Cowan, R. W., Maitra, A. & Rhim, A. D. 2015. *RE: A New Scalpel for the Treatment of Pancreatic Cancer: Targeting Stromal-Derived STAT3 Signaling.*

Cross, A. J., Boca, S., Freedman, N. D., Caporaso, N. E., Huang, W. Y., Sinha, R., Sampson, J. N. & Moore, S. C. 2014. Metabolites of tobacco smoking and colorectal cancer risk. *Carcinogenesis,* 35**,** 1516-22.

Cunningham, D., Chau, I., Stocken, D. D., Valle, J. W., Smith, D., Steward, W., Harper, P. G., Dunn, J., Tudur-Smith, C., West, J., Falk, S., Crellin, A., Adab, F., Thompson, J., Leonard, P., Ostrowski, J., Eatock, M., Scheithauer, W., Herrmann, R. & Neoptolemos, J. P. 2009. Phase III randomized comparison of gemcitabine versus gemcitabine plus capecitabine in patients with advanced pancreatic cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 27**,** 5513-8.

Cvetkovic, R. S. & Plosker, G. L. 2007. Exenatide: a review of its use in patients with type 2 diabetes mellitus (as an adjunct to metformin and/or a sulfonylurea). *Drugs,* 67**,** 935-54.

Dandapani, S., Rosse, G., Southall, N., Salvino, J. & Thomas, C. 2012. Selecting, Acquiring, and Using Small Molecule Libraries for High-Throughput Screening. *Current protocols in Chemical Biology,* 4**,** 177-191.

Dang, B., Kubota, T., Mandal, K., Correa, A. M., Bezanilla, F. & Kent, S. B. 2016. Elucidation of the Covalent and Tertiary Structures of Biologically Active Ts3 Toxin. *Angew Chem Int Ed Engl,* 55**,** 8639-42.

Dang, B., Shen, R., Kubota, T., Mandal, K., Bezanilla, F., Roux, B. & Kent, S. B. 2017. Inversion of the Side-Chain Stereochemistry of Indvidual Thr or Ile Residues in a Protein Molecule:

Impact on the Folding, Stability, and Structure of the ShK Toxin. *Angew Chem Int Ed Engl,* 56**,** 3324-3328.

Danovi, S. A., Wong, H. H. & Lemoine, N. R. 2008. Targeted therapies for pancreatic cancer. *British medical bulletin,* 87**,** 97-130.

De Hoffmann, E. 2005. Mass Spectrometry. *Kirk-Othmer Encyclopedia of Chemical Technology.*

Debnath, A., Saha, A., Gomes, A., Biswas, S., Chakrabarti, P., Giri, B., Biswas, A. K., Gupta, S. D. & Gomes, A. 2010. A lethal cardiotoxic-cytotoxic protein from the Indian monocellate cobra (Naja kaouthia) venom. *Toxicon,* 56**,** 569-79.

Deer, E. L., Gonzalez-Hernandez, J., Coursen, J. D., Shea, J. E., Ngatia, J., Scaife, C. L., Firpo, M. A. & Mulvihill, S. J. 2010. Phenotype and genotype of pancreatic cancer cell lines. *Pancreas,* 39**,** 425-35.

Deplanque, G., Demarchi, M., Hebbar, M., Flynn, P., Melichar, B., Atkins, J., Nowara, E., Moyé, L., Piquemal, D., Ritter, D., Dubreuil, P., Mansfield, C. D., Acin, Y., Moussy, A., Hermine, O. & Hammel, P. 2015. A randomized, placebo-controlled phase III trial of masitinib plus gemcitabine in the treatment of advanced pancreatic cancer. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO,* 26**,** 1194-200.

Deplanque, G., Demarchi, M., Hebbar, M., Flynn, P., Melichar, B., Atkins, J., Ruszniewski, P., Raymond, E., Hermine, O. & Hammel, P. 2013. Masitinib in nonresectable pancreatic cancer: Results of a phase III randomized placebo-controlled trial. *Journal of Clinical Oncology,* 31.

Derakhshani, A., Silvestris, N., Hajiasgharzadeh, K., Mahmoudzadeh, S., Fereidouni, M., Paradiso, A. V., Brunetti, O., Atarod, D., Safarpour, H. & Baradaran, B. 2020. Expression and characterization of a novel recombinant cytotoxin II from Naja naja oxiana venom: A potential treatment for breast cancer. *Int J Biol Macromol,* 162**,** 1283-1292.

Deshane, J., Garner, C. C. & Sontheimer, H. 2003. Chlorotoxin inhibits glioma cell invasion via matrix metalloproteinase-2. *J Biol Chem,* 278**,** 4135-44.

Doley, R. & Kini, R. M. 2009. Protein complexes in snake venom. *Cell Mol Life Sci,* 66**,** 2851-71.

Dragovich, T., Burris, H., Loehrer, P., Von Hoff, D. D., Chow, S., Stratton, S., Green, S., Obregon, Y., Alvarez, I. & Gordon, M. 2008. Gemcitabine plus celecoxib in patients with advanced or metastatic pancreatic adenocarcinoma: results of a phase II trial. *American journal of clinical oncology,* 31**,** 157-62.

Du, G., Fang, Q. & Den Toonder, J. 2016. Microfluidics for cell-based high throughput screening platforms—A review. *Analytica Chimica Acta,* 903**,** 36-50.

Duan, J. X., Jiao, H., Kaizerman, J., Stanton, T., Evans, J. W., Lan, L., Lorente, G., Banica, M., Jung, D., Wang, J., Ma, H., Li, X., Yang, Z., Hoffman, R. M., Ammons, W. S., Hart, C. P. & Matteucci, M. 2008. Potent and highly selective hypoxia-activated achiral phosphoramidate mustards as anticancer drugs. *Journal of Medicinal Chemistry,* 51**,** 2412-2420.

Dubovskii, P. V., Konshina, A. G. & Efremov, R. G. 2014. Cobra cardiotoxins: membrane interactions and pharmacological potential. *Curr Med Chem,* 21**,** 270-87.

Dubovskii, P. V., Lesovoy, D. M., Dubinnyi, M. A., Konshina, A. G., Utkin, Y. N., Efremov, R. G. & Arseniev, A. S. 2005. Interaction of three-finger toxins with phospholipid membranes: comparison of S- and P-type cytotoxins. *Biochem J,* 387**,** 807-15.

Dubovskii, P. V. & Utkin, Y. N. 2015. Antiproliferative activity of cobra venom cytotoxins. *Current topics in medicinal chemistry,* 15**,** 638-48.

Duffy, M. J., Synnott, N. C., O'grady, S. & Crown, J. 2020. Targeting p53 for the treatment of cancer. *Semin Cancer Biol*.

Dufton, M. J. & Hider, R. C. 1988. Structure and pharmacology of elapid cytotoxins. *Pharmacol Ther,* 36**,** 1-40.

Dutta, S., Chanda, A., Kalita, B., Islam, T., Patra, A. & Mukherjee, A. K. 2017. Proteomic analysis to unravel the complex venom proteome of eastern India Naja naja: Correlation of venom composition with its biochemical and pharmacological properties. *J Proteomics*.

Eaden, J. A., Abrams, K. R. & Mayberry, J. F. 2001. The risk of colorectal cancer in ulcerative colitis: a meta-analysis. *Gut,* 48**,** 526-35.

Ebrahim, K., Shirazi, F. H., Mirakabadi, A. Z. & Vatanpour, H. 2015. Cobra venom cytotoxins; apoptotic or necrotic agents? *Toxicon,* 108**,** 134-40.

Ebrahim, K., Vatanpour, H., Zare, A., Shirazi, F. H. & Nakhjavani, M. 2016. Anticancer Activity a of Caspian Cobra (Naja naja oxiana) snake Venom in Human Cancer Cell Lines Via Induction of Apoptosis. *Iran J Pharm Res,* 15**,** 101-112.

El-Naggar, A. K., Lai, S., Clayman, G., Lee, J. K., Luna, M. A., Goepfert, H. & Batsakis, J. G. 1997. Methylation, a major mechanism of p16/CDKN2 gene inactivation in head and neck squamous carcinoma. *Am J Pathol,* 151**,** 1767-74.

El-Rayes, B. F., Zalupski, M. M., Shields, A. F., Ferris, A. M., Vaishampayan, U., Heilbrun, L. K., Venkatramanamoorthy, R., Adsay, V. & Philip, P. A. 2005. A Phase II study of celecoxib, gemcitabine, and cisplatin in advanced pancreatic cancer. *Investigational New Drugs,* 23**,** 583-590.

Emter, R. & Natsch, A. 2015. A fast Resazurin-based live viability assay is equivalent to the MTT-test in the KeratinoSens assay. *Toxicol In Vitro,* 29**,** 688-93.

Erb, R. E. & Ehlers, M. H. 1950. Resazurin Reducing Time as an Indicator of Bovine Semen Fertilizing Capacity. *Journal of Dairy Science,* 33**,** 853-864.

Escalante, T., Saravia-Otten, P., Gastaldello, S., Hernandez, R., Marin, A., Garcia, G., Garcia, L., Estrada, E., Rucavado, A. & Gutierrez, J. M. 2021. Changes in basement membrane components in an experimental model of skeletal muscle degeneration and regeneration induced by snake venom and myotoxic phospholipase A2. *Toxicon,* 192**,** 46-56.

Escoubas, P. & King, G. F. 2009. Venomics as a drug discovery platform. *Expert Rev. Proteomics,* 6**,** 221-224.

Esparza-Molto, P. B. & Cuezva, J. M. 2020. Reprogramming Oxidative Phosphorylation in Cancer: A Role for RNA-Binding Proteins. *Antioxid Redox Signal*.

Ezechias, M. & Cajthaml, T. 2018. Receptor partial agonism and method to express receptor partial activation with respect to novel Full Logistic Model of mixture toxicology. *Toxicology,* 393**,** 26-33.

Falasca, M., Kim, M. & Casari, I. 2016. Pancreatic cancer: Current research and future directions. *Biochimica et biophysica acta*.

Falconer, R. J. & Collins, B. M. 2011. Survey of the year 2009: applications of isothermal titration calorimetry. *J Mol Recognit,* 24**,** 1-16.

Favreau, P., Menin, L., Michalet, S., Perret, F., Cheneval, O., Stocklin, M., Bulet, P. & Stocklin, R. 2006. Mass spectrometry strategies for venom mapping and peptide sequencing from crude venoms: case applications with single arthropod specimen. *Toxicon,* 47**,** 676-87.

Fenton, T. R., Garrett, M. D., Wass, M. N. & Michaelis, M. 2018. What really matters - response and resistance in cancer therapy. *Cancer Drug Resistance,* 1**,** 200-203.

Feofanov, A. V., Sharonov, G. V., Astapova, M. V., Rodionov, D. I., Utkin, Y. N. & Arseniev, A. S. 2005. Cancer cell injury by cytotoxins from cobra venom is mediated through lysosomal damage. *Biochem J,* 390**,** 11-8.

Feofanov, A. V., Sharonov, G. V., Dubinnyi, M. A., Astapova, M. V., Kudelina, I. A., Dubovskii, P. V., Rodionov, D. I., Utkin, Y. N. & Arseniev, A. S. 2004. Comparative study of structure and activity of cytotoxins from venom of the cobras Naja oxiana, Naja kaouthia, and Naja haje. *Biochemistry (Mosc),* 69**,** 1148-57.

Ferreira, I. G., Pucca, M. B., Oliveira, I. S., Cerni, F. A., Jacob, B. & Arantes, E. C. 2021. Snake venom vascular endothelial growth factors (svVEGFs): Unravelling their molecular structure, functions, and research potential. *Cytokine Growth Factor Rev*.

Ferreira, P. M., Lima, D. J., Debiasi, B. W., Soares, B. M., Machado Kda, C., Noronha Jda, C., Rodrigues Dde, J., Sinhorin, A. P., Pessoa, C. & Vieira, G. M., Jr. 2013. Antiproliferative

activity of Rhinella marina and Rhaebo guttatus venom extracts from Southern Amazon. *Toxicon,* 72**,** 43-51.
Fleige, S. & Pfaffl, M. W. 2006. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med,* 27**,** 126-39.
Franca, S. C., Kashima, S., Roberto, P. G., Marins, M., Ticli, F. K., Pereira, J. O., Astolfi-Filho, S., Stabeli, R. G., Magro, A. J., Fontes, M. R., Sampaio, S. V. & Soares, A. M. 2007. Molecular approaches for structural characterization of Bothrops L-amino acid oxidases with antiprotozoal activity: cDNA cloning, comparative sequence analysis, and molecular modeling. *Biochem Biophys Res Commun,* 355**,** 302-6.
Frank, A. M., Savitski, M. M., Nielsen, M. L., Zubarev, R. A. & Pevzner, P. A. 2007. De novo peptide sequencing and identification with precision mass spectrometry. *J Proteome Res,* 6**,** 114-23.
Friess, H., Langrehr, J. M., Oettle, H., Raedle, J., Niedergethmann, M., Dittrich, C., Hossfeld, D. K., Stöger, H., Neyns, B., Herzog, P., Piedbois, P., Dobrowolski, F., Scheithauer, W., Hawkins, R., Katz, F., Balcke, P., Vermorken, J., Van Belle, S., Davidson, N., Esteve, A. A., Castellano, D., Kleeff, J., Tempia-Caliera, A. A., Kovar, A. & Nippgen, J. 2006. A randomized multi-center phase II trial of the angiogenesis inhibitor Cilengitide (EMD 121974) and gemcitabine compared with gemcitabine alone in advanced unresectable pancreatic cancer. *BMC cancer,* 6**,** 285-285.
Fuchs, C. S., Azevedo, S., Okusaka, T., Van Laethem, J. L., Lipton, L. R., Riess, H., Szczylik, C., Moore, M. J., Peeters, M., Bodoky, G., Ikeda, M., Melichar, B., Nemecek, R., Ohkawa, S., Świeboda-Sadlej, A., Tjulandin, S. A., Van Cutsem, E., Loberg, R., Haddad, V., Gansert, J. L., Bach, B. A. & Carrato, A. 2015. A phase 3 randomized, double-blind, placebo-controlled trial of ganitumab or placebo in combination with gemcitabine as first-line therapy for metastatic adenocarcinoma of the pancreas: the GAMMA trial. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO,* 26**,** 921-7.
Fulda, S., Gorman, A. M., Hori, O. & Samali, A. 2010. Cellular stress responses: cell survival and cell death. *Int J Cell Biol,* 2010**,** 214074.
Gao, B., Dalziel, J., Tanzi, S. & Zhu, S. 2018. Meucin-49, a multifunctional scorpion venom peptide with bactericidal synergy with neurotoxins. *Amino Acids,* 50**,** 1025-1043.
Gao, R., Zhang, Y. & Gopalakrishnakone, P. 2008. Purification and N-terminal sequence of a serine proteinase-like protein (BMK-CBP) from the venom of the Chinese scorpion (Buthus martensii Karsch). *Toxicon,* 52**,** 348-53.
Gasanov, S. E., Shrivastava, I. H., Israilov, F. S., Kim, A. A., Rylova, K. A., Zhang, B. & Dagda, R. K. 2015. Naja naja oxiana Cobra Venom Cytotoxins CTI and CTII Disrupt Mitochondrial Membrane Integrity: Implications for Basic Three-Fingered Cytotoxins. *PLoS One,* 10**,** e0129248.
Gebrim, L. C., Marcussi, S., Menaldo, D. L., De Menezes, C. S. R., Nomizo, A., Hamaguchi, A., Silveira-Lacerda, E. P., Homsi-Brandeburgo, M. I., Sampaio, S. V., Soares, A. M. & Rodrigues, V. M. 2009. Antitumor effects of snake venom chemically modified Lys49 phospholipase A2-like BthTX-I and a synthetic peptide derived from its C-terminal region. *Biologicals : journal of the International Association of Biological Standardization,* 37**,** 222-9.
Geurts, J., Evans, D. & Tsai, S. 2015. Genetic screening for patients with pancreatic cancer: Frequency of high-risk mutations. *Journal of Clinical Oncology,* 33.
Gheyi, T. & Molina-Martin, M. 2018. Chapter 5 Mass Spectrometry in Biophysics: from High Throughput Screening to Structural Biology. *Biophysical Techniques in Drug Discovery.* The Royal Society of Chemistry.
Gillen, S., Schuster, T., Meyer Zum Büschenfelde, C., Friess, H. & Kleeff, J. 2010. Preoperative/neoadjuvant therapy in pancreatic cancer: a systematic review and

meta-analysis of response and resection percentages. *PLoS medicine,* 7**,** e1000267-e1000267.

Goldstein, D., El-Maraghi, R. H., Hammel, P., Heinemann, V., Kunzmann, V., Sastre, J., Scheithauer, W., Siena, S., Tabernero, J., Teixeira, L., Tortora, G., Van Laethem, J.-L., Young, R., Penenberg, D. N., Lu, B., Romano, A. & Von Hoff, D. D. 2015. nab-Paclitaxel plus gemcitabine for metastatic pancreatic cancer: long-term survival from a phase III trial. *Journal of the National Cancer Institute,* 107.

Gonçalves, A., Gilabert, M., François, E., Dahan, L., Perrier, H., Lamy, R., Re, D., Largillier, R., Gasmi, M., Tchiknavorian, X., Esterni, B., Genre, D., Moureau-Zabotto, L., Giovannini, M., Seitz, J. F., Delpero, J. R., Turrini, O., Viens, P. & Raoul, J. L. 2012. BAYPAN study: a double-blind phase III randomized trial comparing gemcitabine plus sorafenib and gemcitabine plus placebo in patients with advanced pancreatic cancer. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO,* 23**,** 2799-805.

Gong, H., Cao, Y., Han, G., Zhang, Y., You, Q., Wang, Y. & Pan, Y. 2017. p53/microRNA-374b/AKT1 regulates colorectal cancer cell apoptosis in response to DNA damage. *Int J Oncol,* 50**,** 1785-1791.

Gong, X., Liang, Z., Yang, Y., Liu, H., Ji, J. & Fan, Y. 2020. A resazurin-based, nondestructive assay for monitoring cell proliferation during a scaffold-based 3D culture process. *Regenerative Biomaterials*.

Gorson, J. & Holford, M. 2016. Small Packages, Big Returns: Uncovering the Venom Diversity of Small Invertebrate Conoidean Snails. *Integr Comp Biol*.

Guha, R. 2013. On exploring structure-activity relationships. *Methods Mol Biol,* 993**,** 81-94.

Gullo, L., Pezzilli, R., Ventrucci, M., Lesi, C., Zoni, L., D'ambrosi, A. & Alvisi, V. 1989. Serum immunoreactive elastase: is it useful for the diagnosis of pancreatic cancer? *Pancreas,* 4**,** 335-8.

Guo, D., Mant, C. T., Taneja, A. K., Parker, J. M. R. & Rodges, R. S. 1986. Prediction of peptide retention times in reversed-phase high-performance liquid chromatography I. Determination of retention coefficients of amino acid residues of model synthetic peptides. *Journal of Chromatography A,* 359**,** 499-518.

Guzmán, E., Harmody, D., Pitts, T., Vera-Diaz, B., Ewinder, P., Yu, Y. & Wright, A. 2017. Inhibition of IL-8 secretion on BxPC-3 and MIA PaCa-2 cells and induction of cytotoxicity in pancreatic cancer cells with marine natural products. *Anti-Cancer Drugs,* 28**,** 153-160.

Hafner, A., Bulyk, M. L., Jambhekar, A. & Lahav, G. 2019. The multiple mechanisms that regulate p53 activity and cell fate. *Nat Rev Mol Cell Biol,* 20**,** 199-210.

Hagel, L. 2001. Gel-filtration chromatography. *Curr Protoc Mol Biol,* Chapter 10**,** Unit 10 9.

Hamid, R., Rotshteyn, Y., Rabadi, L., Parikh, R. & Bullock, P. 2004. Comparison of alamar blue and MTT assays for high through-put screening. *Toxicology in Vitro,* 18**,** 703-710.

Hanahan, D. & Weinberg, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

Harris, D. C. 2010. *Quantitative Chemical Analysis*, W. H. Freeman.

Harvey, A. L. 2014. Toxins and drug discovery. *Toxicon : official journal of the International Society on Toxinology,* 92**,** 193-200.

Harvey, A. L., Edrada-Ebel, R. & Quinn, R. J. 2015. The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discov,* 14**,** 111-29.

Hasan, S., Jacob, R., Manne, U. & Paluri, R. 2019. Advances in pancreatic cancer biomarkers. *Oncol Rev,* 13**,** 410.

Hayakawa, T., Kondo, T., Shibata, T., Kitagawa, M., Katada, N., Kato, K. & Takeichi, M. 1990. Prospective trial for early detection of pancreatic cancer by elevated serum immunoreactive elastase. *Gastroenterologia Japonica,* 25**,** 727-31.

Heinemann, V., Hertel, L. W., Grindey, G. B. & Plunkett, W. 1988. Comparison of the cellular pharmacokinetics and toxicity of 2',2'-difluorodeoxycytidine and 1-beta-D-arabinofuranosylcytosine. *Cancer research,* 48**,** 4024-31.

Heinemann, V., Quietzsch, D., Gieseler, F., Gonnermann, M., Schönekäs, H., Rost, A., Neuhaus, H., Haag, C., Clemens, M., Heinrich, B., Vehling-Kaiser, U., Fuchs, M., Fleckenstein, D., Gesierich, W., Uthgenannt, D., Einsele, H., Holstege, A., Hinke, A., Schalhorn, A. & Wilkowski, R. 2006. Randomized phase III trial of gemcitabine plus cisplatin compared with gemcitabine alone in advanced pancreatic cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 24**,** 3946-52.

Hermann, P. C., Sancho, P., Cañamero, M., Martinelli, P., Madriles, F., Michl, P., Gress, T., De Pascual, R., Gandia, L., Guerra, C., Barbacid, M., Wagner, M., Vieira, C. R., Aicher, A., Real, F. X., Sainz, B. & Heeschen, C. 2014. Nicotine promotes initiation and progression of KRAS-induced pancreatic cancer via Gata6-dependent dedifferentiation of acinar cells in mice. *Gastroenterology,* 147**,** 1119-33.e4.

Hernandez, Y. G. & Lucas, A. L. 2016. MicroRNA in pancreatic ductal adenocarcinoma and its precursor lesions. *World J Gastrointest Oncol,* 8**,** 18-29.

Herzig, V., Cristofori-Armstrong, B., Israel, M. R., Nixon, S. A., Vetter, I. & King, G. F. 2020. Animal toxins - Nature's evolutionary-refined toolkit for basic research and drug discovery. *Biochem Pharmacol,* 181**,** 114096.

Herzig, V. & King, G. F. 2015. The Cystine Knot Is Responsible for the Exceptional Stability of the Insecticidal Spider Toxin omega-Hexatoxin-Hv1a. *Toxins (Basel),* 7**,** 4366-80.

Hidalgo, M., Cascinu, S., Kleeff, J., Labianca, R., Lohr, J. M., Neoptolemos, J., Real, F. X., Van Laethem, J. L. & Heinemann, V. 2015. Addressing the challenges of pancreatic cancer: future directions for improving outcomes. *Pancreatology,* 15**,** 8-18.

Hodges, S. J., Department Of, P., Pharmacology, U. O. S. G., Agbaji, A. S., Department of Chemistry, U. O. E. C., Harvey, A. L., Department Of, P., Pharmacology, U. O. S. G., Hider, R. C. & Department of Chemistry, U. O. E. C. 1987. Cobra cardiotoxins. *The FEBS Journal,* 165**,** 373-383.

Holcomb, J., Spellmon, N., Zhang, Y., Doughan, M., Li, C. & Yang, Z. 2017. Protein crystallization: Eluding the bottleneck of X-ray crystallography. *AIMS Biophys,* 4**,** 557-575.

Hong, S., Giesy, J. P., Lee, J.-S., Lee, J.-H. & Khim, J. S. 2016. Effect-directed analysis: Current status and future challenges. *Ocean Science Journal,* 51**,** 413-433.

Hu, Y., Fu, H., Qiao, H., Sun, S., Zhang, W., Jin, S., Jiang, S., Gong, Y., Xiong, Y. & Wu, Y. 2018. Validation and Evaluation of Reference Genes for Quantitative Real-Time PCR in Macrobrachium Nipponense. *Int J Mol Sci,* 19**,** 2258.

Huang, H. 2018. Matrix Metalloproteinase-9 (MMP-9) as a Cancer Biomarker and MMP-9 Biosensors: Recent Advances. *Sensors (Basel),* 18.

Huang, P., Chubb, S., Hertel, L. W., Grindey, G. B. & Plunkett, W. 1991. Action of 2',2'-difluorodeoxycytidine on DNA synthesis. *Cancer research,* 51**,** 6110-7.

Hus, K. K., Buczkowicz, J., Petrilla, V., Petrillova, M., Lyskowski, A., Legath, J. & Bocian, A. 2018. First Look at the Venom of Naja ashei. *Molecules,* 23.

Ilic, M. & Ilic, I. 2016. Epidemiology of pancreatic cancer. *World J Gastroenterol,* 22**,** 9694-705.

Ismail, M., Al-Bekairi, A. M., El-Bedaiwy, A. M. & Abd-El Salam, M. A. 1993. The ocular effects of spitting cobras: II. Evidence that cardiotoxins are responsible for the corneal opacification syndrome. *J Toxicol Clin Toxicol,* 31**,** 45-62.

Ivanisenko, V. A., Eroshkin, A. M. & Kolchanov, N. A. 2005. WebProAnalyst: an interactive tool for analysis of quantitative structure-activity relationships in protein families. *Nucleic Acids Res,* 33**,** W99-104.

Izeradjene, K., Combs, C., Best, M., Gopinathan, A., Wagner, A., Grady, W. M., Deng, C. X., Hruban, R. H., Adsay, N. V., Tuveson, D. A. & Hingorani, S. R. 2007. Kras(G12D) and

Smad4/Dpc4 haploinsufficiency cooperate to induce mucinous cystic neoplasms and invasive adenocarcinoma of the pancreas. *Cancer Cell,* 11**,** 229-43.

Jang, Y. J., Kim, D. S., Jeon, O. H. & Kim, D. S. 2007. Saxatilin suppresses tumor-induced angiogenesis by regulating VEGF expression in NCI-H460 human lung cancer cells. *J Biochem Mol Biol,* 40**,** 439-43.

Jesupret, C., Baumann, K., Jackson, T. N., Ali, S. A., Yang, D. C., Greisman, L., Kern, L., Steuten, J., Jouiaei, M., Casewell, N. R., Undheim, E. A., Koludarov, I., Debono, J., Low, D. H., Rossi, S., Panagides, N., Winter, K., Ignjatovic, V., Summerhayes, R., Jones, A., Nouwens, A., Dunstan, N., Hodgson, W. C., Winkel, K. D., Monagle, P. & Fry, B. G. 2014. Vintage venoms: proteomic and pharmacological stability of snake venoms stored for up to eight decades. *J Proteomics,* 105**,** 285-94.

Jiao, L., Zhu, J., Hassan, M. M., Evans, D. B., Abbruzzese, J. L. & Li, D. 2007. K-ras mutation and p16 and preproenkephalin promoter hypermethylation in plasma DNA of pancreatic cancer patients: in relation to cigarette smoking. *Pancreas,* 34**,** 55-62.

Jokhio, R. & Ansari, A. F. 2005. Cobra snake venom reduces significantly tissue nucleic acid levels in human breast cancer. *J Pak Med Assoc,* 55**,** 71-3.

Jones, S., Zhang, X., Parsons, D. W., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Kamiyama, H., Jimeno, A., Hong, S. M., Fu, B., Lin, M. T., Calhoun, E. S., Kamiyama, M., Walter, K., Nikolskaya, T., Nikolsky, Y., Hartigan, J., Smith, D. R., Hidalgo, M., Leach, S. D., Klein, A. P., Jaffee, E. M., Goggins, M., Maitra, A., Iacobuzio-Donahue, C., Eshleman, J. R., Kern, S. E., Hruban, R. H., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E. & Kinzler, K. W. 2008. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science,* 321**,** 1801-6.

Jonker, W., Lamoree, M. H., Houtman, C. J. & Kool, J. 2015. Methodologies for Effect-Directed Analysis: Environmental Applications, Food Analysis, and Drug Discovery. *Analyzing Biomolecular Interactions by Mass Spectrometry.*

Joubert, F. & Taljaard, N. 1978. Naja haje (Egyptian cobra) venom. Purification, some properties and the amino acid sequences of four toxins (CM-7, CM-8, CM-9, and CM-10b). *Biochim Biophys Acta,* 534**,** 331-40.

Joubert, F. J. 1976a. Snake venom toxins. The amino-acid sequences of three toxins (CM-8, CM-11 and CM-13a) from Naja haje annulifera (Egyptian cobra) venom. *Eur J Biochem,* 64**,** 219-32.

Joubert, F. J. 1976b. Snake venom toxins. The amino acid sequence of three toxins (CM-2e, CM-4a and CM-) from Naja haje annulifera (Egyptian cobra) venom. *Hoppe Seylers Z Physiol Chem,* 357**,** 1735-50.

Joubert, F. J. 1977. Snake venom toxin. The amino acid sequence of three toxins (CM-2h, CM-4b and CM-6) from Naja haje annulifera (Egyptian cobra) venom. *Hoppe Seylers Z Physiol Chem,* 358**,** 79-96.

Joubert, F. J. & Taljaard, N. 1980. The complete primary structures of three cytotoxins (CM-6, CM-7 and CM-7A) from Naja naja kaouthia (Siamese cobra) snake venom. *Toxicon,* 18**,** 455-67.

Juàrez, P., Wagstaff, S. C., Oliver, J., Sanz, L., Harrison, R. A. & Calvete, J. J. 2006. Molecular Cloning of Disintegrin-like Transcript BA-5A from a Bitis arietans Venom Gland cDNA Library:A Putative Intermediate in the Evolution of the Long-Chain Disintegrin Bitistatin. *Journal of Molecular Evolution,* 63**,** 142-152.

Kalia, J., Milescu, M., Salvatierra, J., Wagner, J., Klint, J. K., King, G. F., Olivera, B. M. & Bosmans, F. 2015. From foe to friend: using animal toxins to investigate ion channel function. *J Mol Biol,* 427**,** 158-175.

Kaneda, N., Sasaki, T. & Hayashi, K. 1977. Primary structures of cardiotoxin analogues II and IV from the venom of Naja jaja atra. *Biochim Biophys Acta,* 491**,** 53-66.

Karapetian, H. 2013. Reptilase time (RT). *Methods Mol Biol,* 992**,** 273-7.

Katali, O., Shipingana, L., Nyarango, P., Paakkonen, M., Haindongo, E., Rennie, T., James, P., Eriksson, J. & Hunter, C. J. 2020. Protein Identification of Venoms of the African Spitting Cobras, Naja mossambica and Naja nigricincta nigricincta. *Toxins (Basel),* 12.

Kazandjian, T. D., Petras, D., Robinson, S. D., Van Thiel, J., Greene, H. W., Arbuckle, K., Barlow, A., Carter, D. A., Wouters, R. M., Whiteley, G., Wagstaff, S. C., Arias, A. S., Albulescu, L. O., Plettenberg Laing, A., Hall, C., Heap, A., Penrhyn-Lowe, S., Mccabe, C. V., Ainsworth, S., Da Silva, R. R., Dorrestein, P. C., Richardson, M. K., Gutierrez, J. M., Calvete, J. J., Harrison, R. A., Vetter, I., Undheim, E. a. B., Wüster, W. & Casewell, N. R. 2021. Convergent evolution of pain-inducing defensive venom components in spitting cobras. *Science,* 371**,** 386-390.

Kerr, J. F., Wyllie, A. H. & Currie, A. R. 1972. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer,* 26**,** 239-57.

Kessentini-Zouari, R., Jebali, J., Taboubi, S., Srairi-Abid, N., Morjen, M., Kallech-Ziri, O., Bezzine, S., Marvaldi, J., El Ayeb, M., Marrakchi, N. & Luis, J. 2010. CC-PLA2-1 and CC-PLA2-2, two Cerastes cerastes venom-derived phospholipases A2, inhibit angiogenesis both in vitro and in vivo. *Lab Invest,* 90**,** 510-9.

Kessler, P., Marchot, P., Silva, M. & Servent, D. 2017. The three-finger toxin fold: a multifunctional structural scaffold able to modulate cholinergic functions. *J Neurochem,* 142 Suppl 2**,** 7-18.

Khunsap, S. 2011. Purification of a phospholipase A2 from Daboia russelii siamensis venom with anticancer effects. *J Venom Res,* 2**,** 42-51.

Kim, D.-S., Jang, Y.-J., Jeon, O.-H. & Kim, D.-S. 2007. Saxatilin, a Snake Venom Disintegrin, Suppresses TNF-α-induced Ovarian Cancer Cell Invasion. *Journal of Biochemistry and molecular biology,* 40**,** 290-294.

Kindler, H. L., Ioka, T., Richel, D. J., Bennouna, J., Létourneau, R., Okusaka, T., Funakoshi, A., Furuse, J., Park, Y. S., Ohkawa, S., Springett, G. M., Wasan, H. S., Trask, P. C., Bycott, P., Ricart, A. D., Kim, S. & Van Cutsem, E. 2011. Axitinib plus gemcitabine versus placebo plus gemcitabine in patients with advanced pancreatic adenocarcinoma: a double-blind randomised phase 3 study. *The Lancet. Oncology,* 12**,** 256-62.

Kindler, H. L., Niedzwiecki, D., Hollis, D., Oraefo, E., Schrag, D., Hurwitz, H., Mcleod, H. L., Mulcahy, M. F., Schilsky, R. L., Goldberg, R. M., Cancer & Leukemia Group, B. 2007. A double-blind, placebo-controlled, randomized phase III trial of gemcitabine (G) plus bevacizumab (B) versus gemcitabine plus placebo (P) in patients (pts) with advanced pancreatic cancer (PC): A preliminary analysis of Cancer and Leukemia Group B (CALGB. *ASCO Meeting Abstracts,* 25**,** 4508-4508.

Kindler, H. L., Wroblewski, K., Wallace, J. A., Hall, M. J., Locker, G., Nattam, S., Agamah, E., Stadler, W. M. & Vokes, E. E. 2012. Gemcitabine plus sorafenib in patients with advanced pancreatic cancer: a phase II trial of the University of Chicago Phase II Consortium. *Investigational new drugs,* 30**,** 382-6.

Kini, R. M. 2002. Molecular moulds with multiple missions: functional sites in three-finger toxins. *Clin Exp Pharmacol Physiol,* 29**,** 815-22.

Kini, R. M. 2003. Excitement ahead: structure, function and mechanism of snake venom phospholipase A2 enzymes. *Toxicon,* 42**,** 827-40.

Kini, R. M. & Doley, R. 2010. Structure, function and evolution of three-finger toxins: mini proteins with multiple targets. *Toxicon,* 56**,** 855-67.

Kishor Roy, N., Bordoloi, D., Monisha, J., Padmavathi, G., Kotoky, J., Golla, R. & B. Kunnumakkara, A. 2017. Specific Targeting of Akt Kinase Isoforms: Taking the Precise Path for Prevention and Treatment of Cancer. *Current Drug Targets,* 18**,** 421-435.

Klaassen, N., Spicer, V. & Krokhin, O. V. 2019. Universal retention standard for peptide separations using various modes of high-performance liquid chromatography. *Journal of Chromatography A,* 1588**,** 163-168.

Klochkov, S. G., Pikhtelev, A. R. & Kozlovskii, V. I. 2008. A new peptide from venom of the East-European hornet Vespa orientalis. Mass spectrometric de novo sequence. *Chemistry of Natural Compounds,* 44**,** 63-66.

Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., Bandukwala, H. S., An, J., Lamperti, E. D., Koh, K. P., Ganetzky, R., Liu, X. S., Aravind, L., Agarwal, S., Maciejewski, J. P. & Rao, A. 2010. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature,* 468**,** 839-43.

Koh, C. Y. & Kini, R. M. 2012. From snake venom toxins to therapeutics--cardiovascular examples. *Toxicon : official journal of the International Society on Toxinology,* 59**,** 497-506.

Konno, K., Rangel, M., Oliveira, J. S., Dos Santos Cabrera, M. P., Fontana, R., Hirata, I. Y., Hide, I., Nakata, Y., Mori, K., Kawano, M., Fuchino, H., Sekita, S. & Neto, J. R. 2007. Decoralin, a novel linear cationic alpha-helical peptide from the venom of the solitary eumenine wasp Oreumenes decoratus. *Peptides,* 28**,** 2320-7.

Konshina, A. G., Boldyrev, I. A., Utkin, Y. N., Omel'kov, A. V. & Efremov, R. G. 2011. Snake cytotoxins bind to membranes via interactions with phosphatidylserine head groups of lipids. *PLoS One,* 6**,** e19064.

Konshina, A. G., Dubovskii, P. V. & Efremov, R. G. 2012. Structure and dynamics of cardiotoxins. *Curr Protein Pept Sci,* 13**,** 570-84.

Konshina, A. G., Krylov, N. A. & Efremov, R. G. 2017. Cardiotoxins: Functional Role of Local Conformational Changes. *J Chem Inf Model,* 57**,** 2799-2810.

Koorstra, J. B. M., Feldmann, G., Habbe, N. & Maitra, A. 2008. Morphogenesis of pancreatic cancer: role of pancreatic intraepithelial neoplasia (PanINs). *Langenbecks Arch Surg,* 393**,** 561-70.

Kopaciewicz, W. & Regnier, F. E. 1983. Mobile phase selection for the high-performance ion-exchange chromatography of proteins. *Anal Biochem,* 133**,** 251-9.

Kumar, S., Stecher, G. & Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol,* 33**,** 1870-4.

Langdon, C. G., Platt, J. T., Means, R. E., Iyidogan, P., Mamillapalli, R., Klein, M., Held, M. A., Lee, J. W., Koo, J. S., Hatzis, C., Hochster, H. S. & Stern, D. F. 2017. Combinatorial Screening of Pancreatic Adenocarcinoma Reveals Sensitivity to Drug Combinations Including Bromodomain Inhibitor Plus Neddylation Inhibitor. *Mol Cancer Ther,* 16**,** 1041-1053.

Larbouret, C., Gaborit, N., Chardes, T., Coelho, M., Campigna, E., Bascoul-Mollevi, C., Mach, J. P., Azria, D., Robert, B. & Pelegrin, A. 2012. In pancreatic carcinoma, dual EGFR/HER2 targeting with cetuximab/trastuzumab is more effective than treatment with trastuzumab/erlotinib or lapatinib alone: implication of receptors' down-regulation and dimers' disruption. *Neoplasia,* 14**,** 121-30.

Larsson, P., Engqvist, H., Biermann, J., Werner Ronnerman, E., Forssell-Aronsson, E., Kovacs, A., Karlsson, P., Helou, K. & Parris, T. Z. 2020. Optimization of cell viability assays to improve replicability and reproducibility of cancer drug sensitivity screens. *Sci Rep,* 10**,** 5798.

Lau, M. T., Manion, J., Littleboy, J. B., Oyston, L., Khuong, T. M., Wang, Q. P., Nguyen, D. T., Hesselson, D., Seymour, J. E. & Neely, G. G. 2019. Molecular dissection of box jellyfish venom cytotoxicity highlights an effective venom antidote. *Nat Commun,* 10**,** 1655.

Lee, J., Jang, K.-T., Ki, C.-S., Lim, T., Park, Y. S., Lim, H. Y., Choi, D.-W., Kang, W. K., Park, K. & Park, J. O. 2007. Impact of epidermal growth factor receptor (EGFR) kinase mutations, EGFR gene amplifications, and KRAS mutations on survival of pancreatic adenocarcinoma. *Cancer,* 109**,** 1561-9.

Lee, S., Abed, D. A., Beamer, L. J. & Hu, L. 2021. Development of a Homogeneous Time-Resolved Fluorescence Resonance Energy Transfer (TR-FRET) Assay for the Inhibition of Keap1-Nrf2 Protein-Protein Interaction. *SLAS Discov,* 26**,** 100-112.

Lee, S. C., Lin, C. C., Wang, C. H., Wu, P. L., Huang, H. W., Chang, C. I. & Wu, W. G. 2014. Endocytotic routes of cobra cardiotoxins depend on spatial distribution of positively charged and hydrophobic domains to target distinct types of sulfated glycoconjugates on cell surface. *J Biol Chem,* 289**,** 20170-81.

Legendre, O., Sookdeo, A. & Foster, D. A. 2014. BxPC3 pancreatic cancer cells express a truncated Smad4 protein upon PI3K and mTOR inhibition. *Oncol Lett.*

Levin, B., Lieberman, D. A., Mcfarland, B., Smith, R. A., Brooks, D., Andrews, K. S., Dash, C., Giardiello, F. M., Glick, S., Levin, T. R., Pickhardt, P., Rex, D. K., Thorson, A., Winawer, S. J., American Cancer Society Colorectal Cancer Advisory, G., Force, U. S. M.-S. T. & American College of Radiology Colon Cancer, C. 2008. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *CA Cancer J Clin,* 58**,** 130-60.

Li, B., Lyu, P., Xi, X., Ge, L., Mahadevappa, R., Shaw, C. & Kwok, H. F. 2018a. Triggering of cancer cell cycle arrest by a novel scorpion venom-derived peptide—Gonearrestide. *Journal of Cellular and Molecular Medicine*.

Li, W., Huang, L., Meng, E., Wang, X., Zhang, D. & Wang, G. 2018b. Effect of crude venom from the spider Chilobrachys jingzhao on the proliferation and differentiation of C17.2 neural stem cells. *Biotechnology & Biotechnological Equipment,* 32**,** 1317-1326.

Li, Y., Wang, J., Gao, X., Han, W., Zheng, Y., Xu, H., Zhang, C., He, Q., Zhang, L., Li, Z. & Zhou, D. 2014. c-Met targeting enhances the effect of irradiation and chemical agents against malignant colon cells harboring a KRAS mutation. *PLoS One,* 9**,** e113186.

Liao, K. F., Lai, S. W., Li, C. I. & Chen, W. C. 2012. Diabetes mellitus correlates with increased risk of pancreatic cancer: a population-based cohort study in Taiwan. *J Gastroenterol Hepatol,* 27**,** 709-13.

Lim, K. T., Zahari, Z., Amanah, A., Zainuddin, Z. & Adenan, M. I. 2016. Development of resazurin-based assay in 384-well format for high throughput whole cell screening of Trypanosoma brucei rhodesiense strain STIB 900 for the identification of potential anti-trypanosomal agents. *Exp Parasitol,* 162**,** 49-56.

Lipinski, C. A. 2016. Rule of five in 2015 and beyond: Target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions. *Adv Drug Deliv Rev,* 101**,** 34-41.

Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev,* 46**,** 3-26.

Lippi, G. & Da Rin, G. 2019. Advantages and limitations of total laboratory automation: a personal overview. *Clin Chem Lab Med,* 57**,** 802-811.

Liu, C. C., Yang, H., Zhang, L. L., Zhang, Q., Chen, B. & Wang, Y. 2014. Biotoxins for cancer therapy. *Asian Pac J Cancer Prev,* 15**,** 4753-8.

Liu, F. 2001. SMAD4/DPC4 and pancreatic cancer survival. Commentary re: M. Tascilar et al., The SMAD4 protein and prognosis of pancreatic ductal adenocarcinoma. Clin. Cancer Res., 7: 4115-4121, 2001. *Clin Cancer Res,* 7**,** 3853-6.

Liu, J., Ji, S., Liang, C., Qin, Y., Jin, K., Liang, D., Xu, W., Shi, S., Zhang, B., Liu, L., Liu, C., Xu, J., Ni, Q. & Yu, X. 2016. Critical role of oncogenic KRAS in pancreatic cancer (Review). *Mol Med Rep,* 13**,** 4943-9.

Liu, R. X., Ma, Y., Hu, X. L., Ren, W. Y., Liao, Y. P., Wang, H., Zhu, J. H., Wu, K., He, B. C. & Sun, W. J. 2018. Anticancer effects of oridonin on colon cancer are mediated via BMP7/p38 MAPK/p53 signaling. *Int J Oncol,* 53**,** 2091-2101.

Livak, K. J. & Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

Logan, R. F., Patnick, J., Nickerson, C., Coleman, L., Rutter, M. D., Von Wagner, C. & English Bowel Cancer Screening Evaluation, C. 2012. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. *Gut,* 61**,** 1439-46.

Louvet, C., Labianca, R., Hammel, P., Lledo, G., Zampino, M. G., André, T., Zaniboni, A., Ducreux, M., Aitini, E., Taïeb, J., Faroux, R., Lepere, C. & De Gramont, A. 2005. Gemcitabine in combination with oxaliplatin compared with gemcitabine alone in locally advanced or metastatic pancreatic cancer: results of a GERCOR and GISCAD phase III trial. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 23**,** 3509-16.

Lozano, G. 2019. Restoring p53 in cancer: the promises and the challenges. *J Mol Cell Biol,* 11**,** 615-619.

Lucena, S. E., Jia, Y., Soto, J. G., Parral, J., Cantu, E., Brannon, J., Lardner, K., Ramos, C. J., Seoane, A. I. & Sanchez, E. E. 2012. Anti-invasive and anti-adhesive activities of a recombinant disintegrin, r-viridistatin 2, derived from the Prairie rattlesnake (Crotalus viridis viridis). *Toxicon,* 60**,** 31-9.

Ma, C., Ren, Y., Yang, J., Ren, Z., Yang, H. & Liu, S. 2018. Improved Peptide Retention Time Prediction in Liquid Chromatography through Deep Learning. *Anal Chem,* 90**,** 10881-10888.

Macarron, R., Banks, M. N., Bojanic, D., Burns, D. J., Cirovic, D. A., Garyantes, T., Green, D. V., Hertzberg, R. P., Janzen, W. P., Paslay, J. W., Schopfer, U. & Sittampalam, G. S. 2011. Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov,* 10**,** 188-95.

Macarron, R. & Hertzberg, R. P. 2011. Design and Implementation of High Throughput Screening Assays. *Mol Biotechnol,* 47**,** 270-285.

Mahajan, U. M., Goni, E., Langhoff, E., Li, Q., Costello, E., Greenhalf, W., Kruger, S., Ormanns, S., Halloran, C., Ganeh, P., Marron, M., Lammerhirt, F., Zhao, Y., Beyer, G., Weiss, F. U., Sendler, M., Bruns, C. J., Kohlmann, T., Kirchner, T., Werner, J., D'haese, J. G., Von Bergwelt-Baildon, M., Heinemann, V., Neoptolemos, J. P., Buchler, M. W., Belka, C., Boeck, S., Lerch, M. M. & Mayerle, J. 2020. Cathepsin D Expression and Gemcitabine Resistance in Pancreatic Cancer. *JNCI Cancer Spectr,* 4**,** pkz060.

Maitra, A., Kern, S. E. & Hruban, R. H. 2006. Molecular pathogenesis of pancreatic cancer. *Best practice & research. Clinical gastroenterology,* 20**,** 211-26.

Malhotra, L., Goyal, H. K. V., Jhuria, S., Dev, K., Kumar, S., Kumar, M., Kaur, P. & Ethayathulla, A. S. 2021. Curcumin rescue p53Y220C in BxPC-3 pancreatic adenocarcinomas cell line: Evidence-based on computational, biophysical, and in vivo studies. *Biochim Biophys Acta Gen Subj,* 1865**,** 129807.

Malih, I., Ahmad Rusmili, M. R., Tee, T. Y., Saile, R., Ghalim, N. & Othman, I. 2014. Proteomic analysis of Moroccan cobra Naja haje legionis venom using tandem mass spectrometry. *J Proteomics,* 96**,** 240-52.

Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J. & Nadon, R. 2006. Statistical practice in high-throughput screening data analysis. *Nat Biotechnol,* 24**,** 167-75.

Mani, J., Vallo, S., Rakel, S., Antonietti, P., Gessler, F., Blaheta, R., Bartsch, G., Michaelis, M., Cinatl, J., Haferkamp, A. & Kogel, D. 2015. Chemoresistance is associated with increased cytoprotective autophagy and diminished apoptosis in bladder cancer cells treated with the BH3 mimetic (-)-Gossypol (AT-101). *BMC Cancer,* 15**,** 224.

Mant, C. T. & Hodges, R. S. 2017. *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation*, CRC Press.

Marek, W. K., Sauer, D., Durauer, A., Jungbauer, A., Piatkowski, W. & Antos, D. 2018. Prediction tool for loading, isocratic elution, gradient elution and scaling up of ion exchange chromatography of proteins. *J Chromatogr A,* 1566**,** 89-101.

Maschberger, M., Resch, H. M., Duhr, S. & Breitsprecher, D. 2015. Exploring Protein Stability by nanoDSF.

http://www.samedanltd.com/uploads/pdf/white_paper/4f37dc17e15d809788f691ee052b2561.pdf: nanoTemper Technologies.
Mccleary, R. J. & Kini, R. M. 2013. Non-enzymatic proteins from snake venoms: a gold mine of pharmacological tools and drug leads. *Toxicon,* 62**,** 56-74.
Mccomb, S., Chan, P. K., Guinot, A., Hartmannsdottir, H., Jenni, S., Dobay, M. P., Bourquin, J. P. & Bornhauser, B. C. 2019. Efficient apoptosis requires feedback amplification of upstream apoptotic signals by effector caspase-3 or -7. *Sci Adv,* 5**,** eaau9433.
Mccullough, D., Atofanei, C., Knight, E., Trim, S. A. & Trim, C. M. 2020. Kinome scale profiling of venom effects on cancer cells reveals potential new venom activities. *Toxicon,* 185**,** 129-146.
Mcgovern, S. L., Caselli, E., Grigorieff, N. & Shoichet, B. K. 2002. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J Med Chem,* 45**,** 1712-22.
Mclaren, D. G., Shah, V., Wisniewski, T., Ghislain, L., Liu, C., Zhang, H. & Saldanha, S. A. 2021. High-Throughput Mass Spectrometry for Hit Identification: Current Landscape and Future Perspectives. *SLAS Discov,* 26**,** 168-191.
Mcmurry, J. 2011. *Organic Chemistry with Biological Applications*, Brooks/Cole Cengage Learning.
Mcnicholl, B. P., Mcgrath, J. W. & Quinn, J. P. 2007. Development and application of a resazurin-based biomass activity test for activated sludge plant management. *Water Research,* 41**,** 127-133.
Mcwilliams, R. R., Wieben, E. D., Rabe, K. G., Pedersen, K. S., Wu, Y., Sicotte, H. & Petersen, G. M. 2011. Prevalence of CDKN2A mutations in pancreatic cancer patients: implications for genetic counseling. *Eur J Hum Genet,* 19**,** 472-8.
Melcher, R., Al-Taie, O., Kudlich, T., Hartmann, E., Maisch, S., Steinlein, C., Schmid, M., Rosenwald, A., Menzel, T., Scheppach, W. & Luhrs, H. 2007. SNP-Array genotyping and spectral karyotyping reveal uniparental disomy as early mutational event in MSS- and MSI-colorectal cancer cell lines. *Cytogenet Genome Res,* 118**,** 214-21.
Mello, S. S., Valente, L. J., Raj, N., Seoane, J. A., Flowers, B. M., Mcclendon, J., Bieging-Rolett, K. T., Lee, J., Ivanochko, D., Kozak, M. M., Chang, D. T., Longacre, T. A., Koong, A. C., Arrowsmith, C. H., Kim, S. K., Vogel, H., Wood, L. D., Hruban, R. H., Curtis, C. & Attardi, L. D. 2017. A p53 Super-tumor Suppressor Reveals a Tumor Suppressive p53-Ptpn14-Yap Axis in Pancreatic Cancer. *Cancer Cell,* 32**,** 460-473 e6.
Melo, S. A., Luecke, L. B., Kahlert, C., Fernandez, A. F., Gammon, S. T., Kaye, J., Lebleu, V. S., Mittendorf, E. A., Weitz, J., Rahbari, N., Reissfelder, C., Pilarsky, C., Fraga, M. F., Piwnica-Worms, D. & Kalluri, R. 2015. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature,* 523**,** 177-182.
Mendez, I., Gutierrez, J. M., Angulo, Y., Calvete, J. J. & Lomonte, B. 2011. Comparative study of the cytolytic activity of snake venoms from African spitting cobras (Naja spp., Elapidae) and its neutralization by a polyspecific antivenom. *Toxicon,* 58**,** 558-64.
Meyer, V. R. 2013. *Practical High-Performance Liquid Chromatography*, Wiley.
Michaelis, M., Wass, M. N. & Cinatl, J. 2019. Drug-adapted cancer cell lines as preclinical models of acquired resistance. *Cancer Drug Resistance,* 2**,** 447-456.
Miksad, R. A., Schnipper, L. & Goldstein, M. 2007. Does a statistically significant survival benefit of erlotinib plus gemcitabine for advanced pancreatic cancer translate into clinical significance and value? *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 25**,** 4506-7; author reply 4508.
Minea, R., Swenson, S., Costa, F., Chen, T. C. & Markland, F. S. 2005. Development of a novel recombinant disintegrin, contortrostatin, as an effective anti-tumor and anti-angiogenic agent. *Pathophysiology of haemostasis and thrombosis,* 34**,** 177-83.

Mini, E., Nobili, S., Caciagli, B., Landini, I. & Mazzei, T. 2006. Cellular pharmacology of gemcitabine. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO,* 17 Suppl 5**,** v7-12.

Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. & Prunotto, M. 2017. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat Rev Drug Discov,* 16**,** 531-543.

Moffitt, R. A., Marayati, R., Flate, E. L., Volmar, K. E., Loeza, S. G. H., Hoadley, K. A., Rashid, N. U., Williams, L. A., Eaton, S. C., Chung, A. H., Smyla, J. K., Anderson, J. M., Kim, H. J., Bentrem, D. J., Talamonti, M. S., Iacobuzio-Donahue, C. A., Hollingsworth, M. A. & Yeh, J. J. 2015. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nature genetics,* 47**,** 1168-78.

Moore, M. J., Goldstein, D., Hamm, J., Figer, A., Hecht, J. R., Gallinger, S., Au, H. J., Murawa, P., Walde, D., Wolff, R. A., Campos, D., Lim, R., Ding, K., Clark, G., Voskoglou-Nomikos, T., Ptasynski, M. & Parulekar, W. 2007. Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 25**,** 1960-6.

Morrison, K. L. & Weiss, G. A. 2001. Combinatorial alanine-scanning. *Curr Opin Chem Biol,* 5**,** 302-7.

Morton, J. P., Timpson, P., Karim, S. A., Ridgway, R. A., Athineos, D., Doyle, B., Jamieson, N. B., Oien, K. A., Lowy, A. M., Brunton, V. G., Frame, M. C., Evans, T. R. & Sansom, O. J. 2010. Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer. *Proc Natl Acad Sci U S A,* 107**,** 246-51.

Motulsky, H. J. & Brown, R. E. 2006. Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics,* 7.

Moyer, B. D., Murray, J. K., Ligutti, J., Andrews, K., Favreau, P., Jordan, J. B., Lee, J. H., Liu, D., Long, J., Sham, K., Shi, L., Stocklin, R., Wu, B., Yin, R., Yu, V., Zou, A., Biswas, K. & Miranda, L. P. 2018. Pharmacological characterization of potent and selective NaV1.7 inhibitors engineered from Chilobrachys jingzhao tarantula venom peptide JzTx-V. *PLoS One,* 13**,** e0196791.

Mozzi, A., Forcella, M., Riva, A., Difrancesco, C., Molinari, F., Martin, V., Papini, N., Bernasconi, B., Nonnis, S., Tedeschi, G., Mazzucchelli, L., Monti, E., Fusi, P. & Frattini, M. 2015. NEU3 activity enhances EGFR activation without affecting EGFR expression and acts on its sialylation levels. *Glycobiology,* 25**,** 855-68.

Mukherjee, A. K., Saviola, A. J., Burns, P. D. & Mackessy, S. P. 2015. Apoptosis induction in human breast cancer (MCF-7) cells by a novel venom L-amino acid oxidase (Rusvinoxidase) is independent of its enzymatic activity and is accompanied by caspase-7 activation and reactive oxygen species production. *Apoptosis,* 20**,** 1358-72.

Mullin, R. 2004. As high-throughput screening draws fire, researchers leverage science to put automation into perspective. *Chem. Eng. News,* 82**,** 23-32.

Murphy, S. J., Hart, S. N., Lima, J. F., Kipp, B. R., Klebig, M., Winters, J. L., Szabo, C., Zhang, L., Eckloff, B. W., Petersen, G. M., Scherer, S. E., Gibbs, R. A., Mcwilliams, R. R., Vasmatzis, G. & Couch, F. J. 2013. Genetic Alterations Associated With Progression From Pancreatic Intraepithelial Neoplasia to Invasive Pancreatic Tumor. *Gastroenterology,* 145**,** 1098-1109 e1.

Murray, D., Morrin, M. & Mcdonnell, S. 2004. Increased invasion and expression of MMP-9 in human colorectal cell lines by a CD44-dependent mechanism. *Anticancer Res,* 24**,** 489-94.

Myszka, D. G. 1997. Kinetic Analysis of Macromolecular Interactions Using Surface Plasmon Resonance Biosensors. *Curr Opin Biotechnol,* 8**,** 50-7.

Mzurikwao, D., Khan, M. U., Samuel, O. W., Cinatl, J., Jr., Wass, M., Michaelis, M., Marcelli, G. & Ang, C. S. 2020. Towards image-based cancer cell lines authentication using deep neural networks. *Sci Rep,* 10**,** 19857.

Nabel, E. G. 2002. CDKs and CKIs: molecular targets for tissue remodelling. *Nat Rev Drug Discov,* 1**,** 587-98.

Nachtigall, P. G., Rautsaw, R. M., Ellsworth, S. A., Mason, A. J., Rokyta, D. R., Parkinson, C. L. & Junqueira-De-Azevedo, I. L. M. 2021. ToxCodAn: a new toxin annotator and guide to venom gland transcriptomics. *Brief Bioinform*.

Nakae, Y., Naruse, S., Shibata, T., Kitagawa, M., Kondo, T., Hayakawa, T., Kuno, N. & Kurimoto, K. 1994. Early detection of pancreatic cancer by serum markers. *Rinsho byori. The Japanese journal of clinical pathology,* 42**,** 139-42.

Nakai, Y., Isayama, H., Ijichi, H., Sasaki, T., Sasahira, N., Hirano, K., Kogure, H., Kawakubo, K., Yagioka, H., Yashima, Y., Mizuno, S., Yamamoto, K., Arizumi, T., Togawa, O., Matsubara, S., Tsujino, T., Tateishi, K., Tada, M., Omata, M. & Koike, K. 2010. Inhibition of renin-angiotensin system affects prognosis of advanced pancreatic cancer receiving gemcitabine. *British journal of cancer,* 103**,** 1644-8.

Nascimento, D. G., Rates, B., Santos, D. M., Verano-Braga, T., Barbosa-Silva, A., Dutra, A. a. A., Biondi, I., Martin-Eauclaire, M. F., De Lima, M. E. & Pimenta, A. M. C. 2006. Moving pieces in a taxonomic puzzle: Venom 2D-LC/MS and data clustering analyses to infer phylogenetic relationships in some scorpions from the Buthidae family (Scorpiones). *Toxicon,* 47**,** 628-639.

Naumann, G. B., Silva, L. F., Silva, L., Faria, G., Richardson, M., Evangelista, K., Kohlhoff, M., Gontijo, C. M. F., Navdaev, A., De Rezende, F. F., Eble, J. A. & Sanchez, E. F. 2011. Cytotoxicity and inhibition of platelet aggregation caused by an l-amino acid oxidase from Bothrops leucurus venom. *Biochimica et biophysica acta,* 1810**,** 683-94.

Nawarak, J., Sinchaikul, S., Wu, C. Y., Liau, M. Y., Phutrakul, S. & Chen, S. T. 2003. Proteomics of snake venoms from Elapidae and Viperidae families by multidimensional chromatographic methods. *Electrophoresis,* 24**,** 2838-54.

Neff, R. A., Flinspach, M., Gibbs, A., Shih, A. Y., Minassian, N. A., Liu, Y., Fellows, R., Libiger, O., Young, S., Pennington, M. W., Hunter, M. J. & Wickenden, A. D. 2020. Comprehensive engineering of the tarantula venom peptide huwentoxin-IV to inhibit the human voltage-gated sodium channel hNav1.7. *Journal of Biological Chemistry,* 295**,** 1315-1327.

Neoptolemos, J. P., Stocken, D. D., Friess, H., Bassi, C., Dunn, J. A., Hickey, H., Beger, H., Fernandez-Cruz, L., Dervenis, C., Lacaine, F., Falconi, M., Pederzoli, P., Pap, A., Spooner, D., Kerr, D. J. & Büchler, M. W. 2004. A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer. *The New England journal of medicine,* 350**,** 1200-1210.

Newman, D. J. & Cragg, G. M. 2020. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod,* 83**,** 770-803.

Ngoi, N. Y. L., Eu, J. Q., Hirpara, J., Wang, L., Lim, J. S. J., Lee, S. C., Lim, Y. C., Pervaiz, S., Goh, B. C. & Wong, A. L. A. 2020. Targeting Cell Metabolism as Cancer Therapy. *Antioxid Redox Signal,* 32**,** 285-308.

Nolte, S., De Castro Damasio, D., Barea, A. C., Gomes, J., Magalhaes, A., Mello Zischler, L. F., Stuelp-Campelo, P. M., Elifio-Esposito, S. L., Roque-Barreira, M. C., Reis, C. A. & Moreno-Amaral, A. N. 2012. BJcuL, a lectin purified from Bothrops jararacussu venom, induces apoptosis in human gastric carcinoma cells accompanied by inhibition of cell adhesion and actin cytoskeleton disassembly. *Toxicon,* 59**,** 81-5.

Nugent, K. D., Burton, W. G., Slattery, T. K., Johnson, B. F. & Snyder, L. R. 1988. Separation of proteins by reversed-phase high-performance liquid chromatography. II. Optimizing sample pretreatment and mobile phase conditions. *J Chromatogr,* 443**,** 381-97.

O'brien, J., Wilson, I., Orton, T. & Pognan, F. 2000. Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. *Eur J Biochem,* 267**,** 5421-6.

Oettle, H. 2005. *RE: A phase III trial of pemetrexed plus gemcitabine versus gemcitabine in patients with unresectable or metastatic pancreatic cancer.*

Oettle, H., Post, S., Neuhaus, P., Gellert, K., Langrehr, J., Ridwelski, K., Schramm, H., Fahlke, J., Zuelke, C., Burkart, C., Gutberlet, K., Kettner, E., Schmalenberg, H., Weigang-Koehler, K., Bechstein, W.-O., Niedergethmann, M., Schmidt-Wolf, I., Roll, L., Doerken, B. & Riess, H. 2007. Adjuvant chemotherapy with gemcitabine vs observation in patients undergoing curative-intent resection of pancreatic cancer: a randomized controlled trial. *Jama,* 297**,** 267-77.

Oh, D. Y. & Bang, Y. J. 2020. HER2-targeted therapies - a role beyond breast cancer. *Nat Rev Clin Oncol,* 17**,** 33-48.

Oldrati, V., Arrell, M., Violette, A., Perret, F., Sprungli, X., Wolfender, J. L. & Stocklin, R. 2016. Advances in venomics. *Mol Biosyst,* 12**,** 3530-3543.

Olsen, P. A., Solberg, N. T., Lund, K., Vehus, T., Gelazauskaite, M., Wilson, S. R. & Krauss, S. 2014. Implications of targeted genomic disruption of beta-catenin in BxPC-3 pancreatic adenocarcinoma cells. *PLoS One,* 9**,** e115496.

Ottaiano, A., Capozzi, M., De Divitiis, C., De Stefano, A., Botti, G., Avallone, A. & Tafuto, S. 2017. Gemcitabine mono-therapy versus gemcitabine plus targeted therapy in advanced pancreatic cancer: a meta-analysis of randomized phase III trials. *Acta Oncol,* 56**,** 377-383.

Pace, R. T. & Burg, K. J. 2015. Toxic effects of resazurin on cell cultures. *Cytotechnology,* 67**,** 13-7.

Paixao-Cavalcante, D., Van Den Berg, C. W., Goncalves-De-Andrade, R. M., Fernandes-Pedrosa Mde, F., Okamoto, C. K. & Tambourgi, D. V. 2007. Tetracycline protects against dermonecrosis induced by Loxosceles spider venom. *J Invest Dermatol,* 127**,** 1410-8.

Palagi, A., Koh, J. M. S., Leblanc, M., Wilson, D., Dutertre, S., King, G. F., Nicholson, G. M. & Escoubas, P. 2013. Unravelling the complex venom landscapes of lethal Australian funnel-web spiders (Hexathelidae: Atracinae) using LC-MALDI-TOF mass spectrometry. *Journal of proteomics,* 80**,** 292-310.

Panagides, N., Jackson, T. N., Ikonomopoulou, M. P., Arbuckle, K., Pretzler, R., Yang, D. C., Ali, S. A., Koludarov, I., Dobson, J., Sanker, B., Asselin, A., Santana, R. C., Hendrikx, I., Van Der Ploeg, H., Tai, A. P. J., Van Den Bergh, R., Kerkkamp, H. M., Vonk, F. J., Naude, A., Strydom, M. A., Jacobsz, L., Dunstan, N., Jaeger, M., Hodgson, W. C., Miles, J. & Fry, B. G. 2017. How the Cobra Got Its Flesh-Eating Venom: Cytotoxicity as a Defensive Innovation and Its Co-Evolution with Hooding, Aposematic Marking, and Spitting. *Toxins (Basel),* 9**,** 103.

Paniagua, D., Crowns, K., Montonera, M., Wertheimer, A., Alagón, A. & Boyer, L. 2020. Postmortem histopathology and detection of venom by ELISA following suicide by cobra (Naja kaouthia) envenomation. *Forensic Toxicology,* 38**,** 523-528.

Pereira-Fantini, P. M., Rajapaksa, A. E., Oakley, R. & Tingay, D. G. 2016. Selection of Reference Genes for Gene Expression Studies related to lung injury in a preterm lamb model. *Sci Rep,* 6**,** 26476.

Perrot, S., Dutertre-Catella, H., Martin, C., Warnet, J. M. & Rat, P. 2003. A new nondestructive cytometric assay based on resazurin metabolism and an organ culture model for the assessment of corneal viability. *Cytometry A,* 55**,** 7-14.

Petras, D., Sanz, L., Segura, A., Herrera, M., Villalta, M., Solano, D., Vargas, M., Leon, G., Warrell, D. A., Theakston, R. D., Harrison, R. A., Durfa, N., Nasidi, A., Gutierrez, J. M. & Calvete, J. J. 2011. Snake venomics of African spitting cobras: toxin composition and assessment of congeneric cross-reactivity of the pan-African EchiTAb-Plus-ICP

antivenom by antivenomics and neutralization approaches. *J Proteome Res,* 10**,** 1266-80.

Philip, P. A., Benedetti, J., Corless, C. L., Wong, R., O'reilly, E. M., Flynn, P. J., Rowland, K. M., Atkins, J. N., Mirtsching, B. C., Rivkin, S. E., Khorana, A. A., Goldman, B., Fenoglio-Preiser, C. M., Abbruzzese, J. L. & Blanke, C. D. 2010. Phase III study comparing gemcitabine plus cetuximab versus gemcitabine in patients with advanced pancreatic adenocarcinoma: Southwest Oncology Group-directed intergroup trial S0205. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 28**,** 3605-10.

Philip, P. A., Goldman, B., Ramanathan, R. K., Lenz, H.-J., Lowy, A. M., Whitehead, R. P., Wakatsuki, T., Iqbal, S., Gaur, R., Benedetti, J. K. & Blanke, C. D. 2014. Dual blockade of epidermal growth factor receptor and insulin-like growth factor receptor-1 signaling in metastatic pancreatic cancer: phase Ib and randomized phase II trial of gemcitabine, erlotinib, and cixutumumab versus gemcitabine plus erlotinib (SWO. *Cancer,* 120**,** 2980-5.

Phillips, H. L., Williamson, J. C., Van Elburg, K. A., Snijders, A. P., Wright, P. C. & Dickman, M. J. 2010. Shotgun proteome analysis utilising mixed mode (reversed phase-anion exchange chromatography) in conjunction with reversed phase liquid chromatography mass spectrometry analysis. *Proteomics,* 10**,** 2950-60.

Pirok, B. W. J. & Schoenmakers, P. J. 2018. Practical Approaches to Overcome the Challenges of Comprehensive Two-Dimensional Liquid Chromatography. *LCGC Europe,* 31**,** 242-249.

Pirok, B. W. J., Stoll, D. R. & Schoenmakers, P. J. 2019. Recent Developments in Two-Dimensional Liquid Chromatography: Fundamental Improvements for Practical Applications. *Anal Chem,* 91**,** 240-263.

Pollakova, M., Petrilla, V., Andrejcakova, Z., Petrillova, M., Sopkova, D. & Petrovova, E. 2021. Spitting cobras: Experimental assay employing the model of chicken embryo and the chick chorioallantoic membrane for imaging and evaluation of effects of venom from African and Asian species (Naja ashei, Naja nigricollis, Naja siamensis, Naja sumatrana). *Toxicon,* 189**,** 79-90.

Posner, B. A., Xi, H. & Mills, J. E. 2009. Enhanced HTS hit selection via a local hit rate analysis. *J Chem Inf Model,* 49**,** 2202-10.

Post, Y., Puschhof, J., Beumer, J., Kerkkamp, H. M., De Bakker, M. a. G., Slagboom, J., De Barbanson, B., Wevers, N. R., Spijkers, X. M., Olivier, T., Kazandjian, T. D., Ainsworth, S., Iglesias, C. L., Van De Wetering, W. J., Heinz, M. C., Van Ineveld, R. L., Van Kleef, R., Begthel, H., Korving, J., Bar-Ephraim, Y. E., Getreuer, W., Rios, A. C., Westerink, R. H. S., Snippert, H. J. G., Van Oudenaarden, A., Peters, P. J., Vonk, F. J., Kool, J., Richardson, M. K., Casewell, N. R. & Clevers, H. 2020. Snake Venom Gland Organoids. *Cell,* 180**,** 233-247 e21.

Postollec, F., Falentin, H., Pavan, S., Combrisson, J. & Sohier, D. 2011. Recent advances in quantitative PCR (qPCR) applications in food microbiology. *Food Microbiol,* 28**,** 848-61.

Poulin, E. J. & Haigis, K. M. 2017. No back seat for a progression event-K-RAS as a therapeutic target in CRC. *Genes Dev,* 31**,** 333-335.

Povey, J. F., Saintas, E., Aderemi, A. V., Rothweiler, F., Zehner, R., Dirks, W. G., Cinatl, J., Jr., Racher, A. J., Wass, M. N., Smales, C. M. & Michaelis, M. 2019. Intact-Cell MALDI-ToF Mass Spectrometry for the Authentication of Drug-Adapted Cancer Cell Lines. *Cells,* 8**,** 1194.

Pozzi, L., Hodgson, J. A., Burrell, A. S., Sterner, K. N., Raaum, R. L. & Disotell, T. R. 2014. Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. *Mol Phylogenet Evol,* 75**,** 165-83.

Prashanth, J. R., Hasaballah, N. & Vetter, I. 2017. Pharmacological screening technologies for venom peptide discovery. *Neuropharmacology,* 127**,** 4-19.

Prat, S. 2018. First hominin settlements out of Africa. Tempo and dispersal mode: Review and perspectives. *Comptes Rendus Palevol,* 17**,** 6-16.

Promega. 2016. *CellTiter-Blue® Cell Viability Assay* [Online]. Available: http://www.promega.co.uk/~/media/files/resources/protocols/technical%20bulletins/101/celltiter-blue%20cell%20viability%20assay%20protocol.pdf [Accessed].

Pucca, M. B., Ahmadi, S., Cerni, F. A., Ledsgaard, L., Sorensen, C. V., Mcgeoghan, F. T. S., Stewart, T., Schoof, E., Lomonte, B., Auf Dem Keller, U., Arantes, E. C., Caliskan, F. & Laustsen, A. H. 2020. Unity Makes Strength: Exploring Intraspecies and Interspecies Toxin Synergism between Phospholipases A2 and Cytotoxins. *Front Pharmacol,* 11**,** 611.

Rachagani, S., Kumar, S. & Batra, S. K. 2010. MicroRNA in pancreatic cancer: pathological, diagnostic and therapeutic implications. *Cancer Lett,* 292**,** 8-16.

Radis-Baptista, G. 2021. Cell-Penetrating Peptides Derived from Animal Venoms and Toxins. *Toxins (Basel),* 13.

Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M. & Matrisian, L. M. 2014. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer research,* 74**,** 2913-21.

Ramos, O. H. & Selistre-De-Araujo, H. S. 2006. Snake venom metalloproteases--structure and function of catalytic and disintegrin domains. *Comp Biochem Physiol C Toxicol Pharmacol,* 142**,** 328-46.

Rampersad, S. N. 2012. Multiple applications of Alamar Blue as an indicator of metabolic function and cellular health in cell viability bioassays. *Sensors (Basel),* 12**,** 12347-60.

Rang, H. P., Dale, M. M., Ritter, J. M. & Flower, R. J. 2008. *Rang & Dale's Pharmacology*, Churchill Livingstone.

Rao, X., Huang, X., Zhou, Z. & Lin, X. 2013. An improvement of the 2^(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. *Biostat Bioinforma Biomath,* 3**,** 71-85.

Reed, S. E. 2012. *Studies on Cytoskeletal and Signalling Proteins in Cells from Rat Dorsal Root Ganglia using Arachnid Venoms as Probes.* Msc. in Biochemistry Thesis, University of Kent.

Reeks, T. A., Fry, B. G. & Alewood, P. F. 2015. Privileged frameworks from snake venom. *Cell Mol Life Sci,* 72**,** 1939-58.

Rees, B., Bilwes, A., Samama, J. P. & Moras, D. 1990. Cardiotoxin VII4 from Naja mossambica mossambica. The refined crystal structure. *J Mol Biol,* 214**,** 281-97.

Rehfeld, J. F. 2017. Cholecystokinin-From Local Gut Hormone to Ubiquitous Messenger. *Front Endocrinol (Lausanne),* 8**,** 47.

Reni, M., Cereda, S., Balzano, G., Passoni, P., Rognone, A., Fugazza, C., Mazza, E., Zerbi, A., Di Carlo, V. & Villa, E. 2009. Carbohydrate antigen 19-9 change during chemotherapy for advanced pancreatic adenocarcinoma. *Cancer,* 115**,** 2630-9.

Rishton, G. M. 2008. Natural products as a robust source of new drugs and drug leads: past successes and present day issues. *Am J Cardiol,* 101**,** 43D-49D.

Robinson, S. D., Undheim, E. a. B., Ueberheide, B. & King, G. F. 2017. Venom peptides as therapeutics: advances, challenges and the future of venom-peptide discovery. *Expert Rev Proteomics,* 14**,** 931-939.

Rocha Lima, C. M., Green, M. R., Rotche, R., Miller, W. H., Jeffrey, G. M., Cisar, L. A., Morganti, A., Orlando, N., Gruia, G. & Miller, L. L. 2004. Irinotecan plus gemcitabine results in no survival advantage compared with gemcitabine monotherapy in patients with locally advanced or metastatic pancreatic cancer despite increased tumor response rate. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 22**,** 3776-83.

Rodrigues, R. S., Da Silva, J. F., Boldrini França, J., Fonseca, F. P. P., Otaviano, A. R., Henrique Silva, F., Hamaguchi, A., Magro, A. J., Braz, A. S. K., Dos Santos, J. I., Homsi-

Brandeburgo, M. I., Fontes, M. R. M., Fuly, A. L., Soares, A. M. & Rodrigues, V. M. 2009. Structural and functional properties of Bp-LAAO, a new l-amino acid oxidase isolated from Bothrops pauloensis snake venom. *Biochimie,* 91**,** 490-501.

Rozanova, S., Barkovits, K., Nikolov, M., Schmidt, C., Urlaub, H. & Marcus, K. 2021. Quantitative Mass Spectrometry-Based Proteomics: An Overview. *In:* MARCUS, K., EISENACHER, M. & SITEK, B. (eds.) *Quantitative Methods in Proteomics.* New York, NY: Springer US.

Rulyak, S. J., Lowenfels, A. B., Maisonneuve, P. & Brentnall, T. A. 2003. Risk factors for the development of pancreatic cancer in familial pancreatic cancer kindreds. *Gastroenterology,* 124**,** 1292-1299.

Rusnak, D. W., Alligood, K. J., Mullin, R. J., Spehar, G. M., Arenas-Elliott, C., Martin, A. M., Degenhardt, Y., Rudolph, S. K., Haws, T. F., Jr., Hudson-Curtis, B. L. & Gilmer, T. M. 2007. Assessment of epidermal growth factor receptor (EGFR, ErbB1) and HER2 (ErbB2) protein expression levels and response to lapatinib (Tykerb, GW572016) in an expanded panel of human normal and tumour cell lines. *Cell Prolif,* 40**,** 580-94.

Safran, H., Miner, T., Bahary, N., Whiting, S., Lopez, C. D., Sun, W., Charpentier, K., Shipley, J., Anderson, E., Mcnulty, B., Schumacher, A., Clark, A., Vakharia, J., Kennedy, T. & Sio, T. 2011. Lapatinib and gemcitabine for metastatic pancreatic cancer. A phase II study. *American journal of clinical oncology,* 34**,** 50-2.

Said, A. H., Raufman, J. P. & Xie, G. 2014. The role of matrix metalloproteinases in colorectal cancer. *Cancers (Basel),* 6**,** 366-75.

Sakurai, Y., Takatsuka, H., Yoshioka, A., Matsui, T., Suzuki, M., Titani, K. & Fujimura, Y. 2001. Inhibition of human platelet aggregation by L-amino acid oxidase purified from Naja naja kaouthia venom. *Toxicon,* 39**,** 1827-33.

Samel, M., Vija, H., Ronnholm, G., Siigur, J., Kalkkinen, N. & Siigur, E. 2006. Isolation and characterization of an apoptotic and platelet aggregation inhibiting L-amino acid oxidase from Vipera berus berus (common viper) venom. *Biochim Biophys Acta,* 1764**,** 707-14.

Sampaio, S. C., Santos, M. F., Costa, E. P., Rangel-Santos, A. C., Carneiro, S. M., Curi, R. & Cury, Y. 2006. Crotoxin induces actin reorganization and inhibits tyrosine phosphorylation and activity of small GTPases in rat macrophages. *Toxicon,* 47**,** 909-19.

Sanz, L., Gibbs, H. L., Mackessy, S. P. & Calvete, J. J. 2006. Venom Proteomes of Closely Related Sistrurus Rattlesnakes with Divergent Diets. *Journal of Proteome Research,* 5**,** 2098-2112.

Sarker, S. D., Nahar, L. & Kumarasamy, Y. 2007. Microtitre plate-based antibacterial assay incorporating resazurin as an indicator of cell growth, and its application in the in vitro antibacterial screening of phytochemicals. *Methods,* 42**,** 321-4.

Saviola, A. J., Burns, P. D., Mukherjee, A. K. & Mackessy, S. P. 2016. The disintegrin tzabcanin inhibits adhesion and migration in melanoma and lung cancer cells. *Int J Biol Macromol,* 88**,** 457-64.

Scarborough, R. M., Rose, J. W., Hsu, M. A., Phillips, D. R., Fried, V., Campbell, A. M., Nannizzi, L. & Charo, I. F. 1991. Barbourin: A GPIIb-IIIa-Specific integrin antagonist from the venom of Sistrurus M. Barbouri. *The Journal of biological chemistry,* 266**,** 9359-62.

Schatoff, E. M., Leach, B. I. & Dow, L. E. 2017. Wnt Signaling and Colorectal Cancer. *Curr Colorectal Cancer Rep,* 13**,** 101-110.

Seidel, S. A., Wienken, C. J., Geissler, S., Jerabek-Willemsen, M., Duhr, S., Reiter, A., Trauner, D., Braun, D. & Baaske, P. 2012. Label-Free Microscale Thermophoresis Discriminates Sites and Affinity of Protein–Ligand Binding. *Angewandte Chemie,* 51**,** 10656-10659.

Senji Laxme, R. R., Attarde, S., Khochare, S., Suranse, V., Martin, G., Casewell, N. R., Whitaker, R. & Sunagar, K. 2021. Biogeographical venom variation in the Indian spectacled cobra (Naja naja) underscores the pressing need for pan-India efficacious snakebite therapy. *PLoS Negl Trop Dis,* 15**,** e0009150.

Shenoy, N., Stenson, M., Lawson, J., Abeykoon, J., Patnaik, M., Wu, X. & Witzig, T. 2017. Drugs with anti-oxidant properties can interfere with cell viability measurements by assays that rely on the reducing property of viable cells. *Lab Invest,* 97**,** 494-497.

Shiloh, M. U., Ruan, J. & Nathan, C. 1997. Evaluation of bacterial survival and phagocyte function with a fluorescence-based microplate assay. *Infect Immun,* 65**,** 3193-8.

Shima, K., Nosho, K., Baba, Y., Cantor, M., Meyerhardt, J. A., Giovannucci, E. L., Fuchs, C. S. & Ogino, S. 2011. Prognostic significance of CDKN2A (p16) promoter methylation and loss of expression in 902 colorectal cancers: Cohort study and literature review. *Int J Cancer,* 128**,** 1080-94.

Shoichet, B. K. 2004. Virtual screening of chemical libraries. *Nature,* 432**,** 862-5.

Shulepko, M. A., Lyukmanova, E. N., Shenkarev, Z. O., Dubovskii, P. V., Astapova, M. V., Feofanov, A. V., Arseniev, A. S., Utkin, Y. N., Kirpichnikov, M. P. & Dolgikh, D. A. 2017. Towards universal approach for bacterial production of three-finger Ly6/uPAR proteins: Case study of cytotoxin I from cobra N. oxiana. *Protein Expr Purif,* 130**,** 13-20.

Shum, D., Radu, C., Kim, E., Cajuste, M., Shao, Y., Seshan, V. E. & Djaballah, H. 2008. A high density assay format for the detection of novel cytotoxic agents in large chemical libraries. *J Enzyme Inhib Med Chem,* 6366**,** 1-1.

Shun, T. Y., Lazo, J. S., Sharlow, E. R. & Johnston, P. A. 2011. Identifying actives from HTS data sets: practical approaches for the selection of an appropriate HTS data-processing method and quality control review. *J Biomol Screen,* 16**,** 1-14.

Si, L., Xu, L., Yin, L., Qi, Y., Han, X., Xu, Y., Zhao, Y., Liu, K. & Peng, J. 2017. Potent effects of dioscin against pancreatic cancer via miR-149-3P-mediated inhibition of the Akt1 signalling pathway. *Br J Pharmacol,* 174**,** 553-568.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., Mcwilliam, H., Remmert, M., Soding, J., Thompson, J. D. & Higgins, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7**,** 539.

Simon, M. A., Ecsedi, P., Kovacs, G. M., Poti, A. L., Remenyi, A., Kardos, J., Gogl, G. & Nyitray, L. 2020. High-throughput competitive fluorescence polarization assay reveals functional redundancy in the S100 protein family. *FEBS J,* 287**,** 2834-2846.

Sinclair, I., Bachman, M., Addison, D., Rohman, M., Murray, D. C., Davies, G., Mouchet, E., Tonge, M. E., Stearns, R. G., Ghislain, L., Datwani, S. S., Majlof, L., Hall, E., Jones, G. R., Hoyes, E., Olechno, J., Ellson, R. N., Barran, P. E., Pringle, S. D., Morris, M. R. & Wingfield, J. 2019a. Acoustic Mist Ionization Platform for Direct and Contactless Ultrahigh-Throughput Mass Spectrometry Analysis of Liquid Samples. *Anal Chem,* 91**,** 3790-3794.

Sinclair, I., Davies, G. & Semple, H. 2019b. Acoustic mist ionization mass spectrometry (AMI-MS) as a drug discovery platform. *Expert Opin Drug Discov,* 14**,** 609-617.

Sink, R., Gobec, S., Pecar, S. & Zega, A. 2010. False Positives in the Early Stages of Drug Discovery. *Current Medicinal Chemistry,* 17**,** 4231-4255.

Sitprija, V. & Sitprija, S. 2012. Renal effects and injury induced by animal toxins. *Toxicon,* 60**,** 943-53.

Siu, L. L., Awada, A., Takimoto, C. H., Piccart, M., Schwartz, B., Giannaris, T., Lathia, C., Petrenciuc, O. & Moore, M. J. 2006. Phase I trial of sorafenib and gemcitabine in advanced solid tumors with an expanded cohort in advanced pancreatic cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 12**,** 144-51.

Slagboom, J., Mladic, M., Xie, C., Kazandjian, T. D., Vonk, F., Somsen, G. W., Casewell, N. R. & Kool, J. 2020. High throughput screening and identification of coagulopathic snake venom proteins and peptides using nanofractionation and proteomics approaches. *PLoS Negl Trop Dis,* 14**,** e0007802.

Smith, C. G. & Vane, J. R. 2003. The discovery of captopril. *FASEB J,* 17**,** 788-9.

Smyth, M. S. & Martin, J. H. 2000. x ray crystallography. *Mol Pathol,* 53**,** 8-14.
Song, H., Niu, X., Quann, J., Li, Y., Yuan, L., Wang, J. & Ma, C. 2020. Discovery of specific HDAC6 inhibitor with anti-metastatic effects in pancreatic cancer cells through virtual screening and biological evaluation. *Bioorganic Chemistry,* 97.
Stoffel, E. M. & Kastrinos, F. 2014. Familial colorectal cancer, beyond Lynch syndrome. *Clin Gastroenterol Hepatol,* 12**,** 1059-68.
Stojanovic, N., Hassan, Z., Wirth, M., Wenzel, P., Beyer, M., Schafer, C., Brand, P., Kroemer, A., Stauber, R. H., Schmid, R. M., Arlt, A., Sellmer, A., Mahboobi, S., Rad, R., Reichert, M., Saur, D., Kramer, O. H. & Schneider, G. 2017. HDAC1 and HDAC2 integrate the expression of p53 mutants in pancreatic cancer. *Oncogene,* 36**,** 1804-1815.
Stoll, D. R. & Carr, P. W. 2017. Two-Dimensional Liquid Chromatography: A State of the Art Tutorial. *Anal. Chem.,* 89**,** 519-531.
Su, B. H., Tu, Y. S., Lin, O. A., Harn, Y. C., Shen, M. Y. & Tseng, Y. J. 2015. Rule-based classification models of molecular autofluorescence. *J Chem Inf Model,* 55**,** 434-45.
Su, J. C., Lin, K. L., Chien, C. M., Chuang, P. W., Chang, L. S. & Lin, S. R. 2010. Concomitant inactivation of the epidermal growth factor receptor, phosphatidylinositol 3-kinase/Akt and Janus tyrosine kinase 2/signal transducer and activator of transcription 3 signalling pathways in cardiotoxin III-treated A549 cells. *Clin Exp Pharmacol Physiol,* 37**,** 833-40.
Sunagar, K., Jackson, T. N., Undheim, E. A., Ali, S. A., Antunes, A. & Fry, B. G. 2013. Three-fingered RAVERs: Rapid Accumulation of Variations in Exposed Residues of snake venom toxins. *Toxins (Basel),* 5**,** 2172-208.
Sundberg, S. A. 2000. High-throughput and ultra-high-throughput screening: solution- and cell-based approaches. *Curr Opin Biotechnol,* 11**,** 47-53.
Swenson, S., Costa, F., Minea, R., Sherwin, R. P., Ernst, W., Fujii, G., Yang, D. & Markland, F. S., Jr. 2004. Intravenous liposomal delivery of the snake venom disintegrin contortrostatin limits breast cancer progression. *Mol. Cancer Ther.,* 3**,** 499-511.
Swinney, D. C. 2013. Phenotypic vs. target-based drug discovery for first-in-class medicines. *Clin Pharmacol Ther,* 93**,** 299-301.
Tan, C., H., Wong, K. Y., Tan, N. H., Ng, T. S. & Tan, K. Y. 2019. Distinctive Distribution of Secretory Phospholipases A2 in the Venoms of Afro-Asian Cobras (Subgenus: Naja, Afronaja, Boulengerina and Uraeus). *Toxins,* 11**,** 116.
Tan, X., Lambert, P. F., Rapraeger, A. C. & Anderson, R. A. 2016. Stress-Induced EGFR Trafficking: Mechanisms, Functions, and Therapeutic Implications. *Trends Cell Biol,* 26**,** 352-366.
Tang, X., Zhang, N., Si, H. & Calderon-Urrea, A. 2017. Selection and validation of reference genes for RT-qPCR analysis in potato under abiotic stress. *Plant Methods,* 13**,** 85.
Tavares, C., Maciel, T., Burin, S., Ambrosio, L., Ghisla, S., Sampaio, S. & Castro, F. 2016. l-Amino acid oxidase isolated from Calloselasma rhodostoma snake venom induces cytotoxicity and apoptosis in JAK2V617F-positive cell lines. *Rev Bras Hematol Hemoter,* 38**,** 128-34.
Taylor, S. C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M. & Fenrich, J. 2019. The Ultimate qPCR Experiment: Producing Publication Quality, Reproducible Data the First Time. *Trends Biotechnol,* 37**,** 761-774.
Teixeira, T., Silva, V., Da Cunha, D., Polettini, F., Thomaz, C., Pianca, A. & Zambom, F. 2016. Isolation, characterization and screening of thein vitrocytotoxicactivity of a novel L-amino acid oxidase (LAAOcdt) fromCrotalusdurissus terrificusvenom on human cancer cell lines. *Toxicon,* 119**,** 203-217.
Thornburg, C. C., Britt, J. R., Evans, J. R., Akee, R. K., Whitt, J. A., Trinh, S. K., Harris, M. J., Thompson, J. R., Ewing, T. L., Shipley, S. M., Grothaus, P. G., Newman, D. J., Schneider, J. P., Grkovic, T. & O'keefe, B. R. 2018. NCI Program for Natural Product Discovery: A Publicly-Accessible Library of Natural Product Fractions for High-Throughput Screening. *ACS Chem Biol,* 13**,** 2484-2497.

Tian, M., Cui, Y. Z., Song, G. H., Zong, M. J., Zhou, X. Y., Chen, Y. & Han, J. X. 2008. Proteomic analysis identifies MMP-9, DJ-1 and A1BG as overexpressed proteins in pancreatic juice from pancreatic ductal adenocarcinoma patients. *BMC Cancer,* 8**,** 241.

Tibbitts, J., Canter, D., Graff, R., Smith, A. & Khawli, L. A. 2016. Key factors influencing ADME properties of therapeutic proteins: A need for ADME characterization in drug discovery and development. *MAbs,* 8**,** 229-45.

Tingle, S. J., Moir, J. A. & White, S. A. 2015. Role of anti-stromal polypharmacy in increasing survival after pancreaticoduodenectomy for pancreatic ductal adenocarcinoma. *World journal of gastrointestinal pathophysiology,* 6**,** 235-42.

Touchard, A., Koh, J. M., Aili, S. R., Dejean, A., Nicholson, G. M., Orivel, J. & Escoubas, P. 2015. The complexity and structural diversity of ant venom peptidomes is revealed by mass spectrometry profiling. *Rapid Commun Mass Spectrom,* 29**,** 385-96.

Tsai, P.-C., Hsieh, C.-Y., Chiu, C.-C., Wang, C.-K., Chang, L.-S. & Lin, S.-R. 2012. Cardiotoxin III suppresses MDA-MB-231 cell metastasis through the inhibition of EGF/EGFR-mediated signaling pathway. *Toxicon : official journal of the International Society on Toxinology,* 60**,** 734-43.

Turchetto, J., Sequeira, A. F., Ramond, L., Peysson, F., Bras, J. L., Saez, N. J., Duhoo, Y., Blemont, M., Guerreiro, C. I., Quinton, L., De Pauw, E., Gilles, N., Darbon, H., Fontes, C. M. & Vincentelli, R. 2017. High-throughput expression of animal venom toxins in Escherichia coli to generate a large library of oxidized disulphide-reticulated peptides for drug discovery. *Microb Cell Fact,* 16**,** 6.

Turillazzi, S., Mastrobuoni, G., Dani, F. R., Moneti, G., Pieraccini, G., La Marca, G., Bartolucci, G., Perito, B., Lambardi, D., Cavallini, V. & Dapporto, L. 2006. Dominulin A and B: two new antibacterial peptides identified on the cuticle and in the venom of the social paper wasp Polistes dominulus using MALDI-TOF, MALDI-TOF/TOF, and ESI-ion trap. *J Am Soc Mass Spectrom,* 17**,** 376-83.

Tuveson, D. A. & Neoptolemos, J. P. 2012. Understanding metastasis in pancreatic cancer: a call for new clinical approaches. *Cell,* 148**,** 21-3.

Utkin, Y. N. 2015. Animal venom studies: Current benefits and future developments. *World J Biol Chem,* 6**,** 28-33.

Uversky, V. N. & Finkelstein, A. V. 2019. Life in Phases: Intra- and Inter- Molecular Phase Transitions in Protein Solutions. *Biomolecules,* 9**,** 842.

Uzarski, J., Divito, M., Wertheim, J. & Miller, W. 2017. Essential Design Considerations for the Resazurin Reduction Assay to Noninvasively Quantify Cell Expansion within Perfused Extracellular Matrix Scaffolds. *Biomaterials,* 129**,** 163-175.

Vallo, S., Kopp, R., Michaelis, M., Rothweiler, F., Bartsch, G., Brandt, M. P., Gust, K. M., Wezel, F., Blaheta, R. A., Haferkamp, A. & Cinatl, J., Jr. 2017. Resistance to nanoparticle albumin-bound paclitaxel is mediated by ABCB1 in urothelial cancer cells. *Oncol Lett,* 13**,** 4085-4092.

Valsecchi, M. E., Díaz-Cantón, E., De La Vega, M. & Littman, S. J. 2014. Recent treatment advances and novel therapies in pancreas cancer: a review. *Journal of gastrointestinal cancer,* 45**,** 190-201.

Van Cutsem, E., Lenz, H.-J., Furuse, J., Tabernero, J., Heinemann, V., Ioka, T., Bazin, I., Ueno, M., Csoszi, T., Wasan, H., Melichar, B., Karasek, P., Macarulla, T., Guillen Ponce, C., Kalinka-Warzocha, E., Horvath, Z., Prenen, H., Schlichting, M., Mehdi, F. & Bendell, J. C. 2016. Evofosfamide (TH-302) in combination with gemcitabine in previously untreated patients with metastatic or locally advanced unresectable pancreatic ductal adenocarcinoma: Primary analysis of the randomized, double-blind phase III MAESTRO study. *ASCO Meeting Abstracts,* 34**,** 193-193.

Van Cutsem, E., Van De Velde, H., Karasek, P., Oettle, H., Vervenne, W. L., Szawlowski, A., Schoffski, P., Post, S., Verslype, C., Neumann, H., Safran, H., Humblet, Y., Perez Ruixo, J., Ma, Y. & Von Hoff, D. 2004. Phase III trial of gemcitabine plus tipifarnib compared

with gemcitabine plus placebo in advanced pancreatic cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 22**,** 1430-8.

Van Cutsem, E., Vervenne, W. L., Bennouna, J., Humblet, Y., Gill, S., Van Laethem, J.-L., Verslype, C., Scheithauer, W., Shang, A., Cosaert, J. & Moore, M. J. 2009. Phase III trial of bevacizumab in combination with gemcitabine and erlotinib in patients with metastatic pancreatic cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 27**,** 2231-7.

Vander Heiden, M. G. 2011. Targeting cancer metabolism: a therapeutic window opens. *Nat Rev Drug Discov,* 10**,** 671-84.

Vetter, I., Davis, J. L., Rash, L. D., Anangi, R., Mobli, M., Alewood, P. F., Lewis, R. J. & King, G. F. 2011. Venomics: a new paradigm for natural products-based drug discovery. *Amino Acids,* 40**,** 15-28.

Von Hoff, D. D., Ervin, T., Arena, F. P., Chiorean, E. G., Infante, J., Moore, M., Seay, T., Tjulandin, S. A., Ma, W. W., Saleh, M. N., Harris, M., Reni, M., Dowden, S., Laheru, D., Bahary, N., Ramanathan, R. K., Tabernero, J., Hidalgo, M., Goldstein, D., Van Cutsem, E., Wei, X., Iglesias, J. & Renschler, M. F. 2013. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *The New England journal of medicine,* 369**,** 1691-703.

Vonk, F. J., Jackson, K., Doley, R., Madaras, F., Mirtschin, P. J. & Vidal, N. 2011. Snake venom: From fieldwork to the clinic: Recent insights into snake biology, together with new technology allowing high-throughput screening of venom, bring new hope for drug discovery. *BioEssays : news and reviews in molecular, cellular and developmental biology,* 33**,** 269-79.

Vyas, V. K., Brahmbhatt, K., Bhatt, H. & Parmar, U. 2013. Therapeutic potential of snake venom in cancer therapy: current perspectives. *Asian Pacific journal of tropical biomedicine,* 3**,** 156-62.

Walewska, A., Skalicky, J. J., Davis, D. R., Zhang, M. M., Lopez-Vera, E., Watkins, M., Han, T. S., Yoshikami, D., Olivera, B. M. & Bulaj, G. 2008. NMR-based mapping of disulfide bridges in cysteine-rich peptides: application to the mu-conotoxin SxIIIA. *J Am Chem Soc,* 130**,** 14280-6.

Wallach, V. a. N., Wüster, W. & Broadley, D. G. 2009. In praise of subgenera: taxonomic status of cobras of the genus Naja Laurenti (Serpentes: Elapidae). *Zootaxa; Vol 2236, No 1: 21 Sep. 2009DO - 10.11646/zootaxa.2236.1.2*.

Walsh, N., Kennedy, S., Larkin, A., Corkery, B., O'driscoll, L., Clynes, M., Crown, J. & O'donovan, N. 2013. EGFR and HER2 inhibition in pancreatic cancer. *Investigational new drugs,* 31**,** 558-66.

Wang-Gillam, A., Li, C.-P., Bodoky, G., Dean, A., Shan, Y.-S., Jameson, G., Macarulla, T., Lee, K.-H., Cunningham, D., Blanc, J. F., Hubner, R. A., Chiu, C.-F., Schwartsmann, G., Siveke, J. T., Braiteh, F., Moyo, V., Belanger, B., Dhindsa, N., Bayever, E., Von Hoff, D. D. & Chen, L.-T. 2015. Nanoliposomal irinotecan with fluorouracil and folinic acid in metastatic pancreatic cancer after previous gemcitabine-based therapy (NAPOLI-1): a global, randomised, open-label, phase 3 trial. *Lancet (London, England),* 387**,** 545-557.

Waters, A. M. & Der, C. J. 2018. KRAS: The Critical Driver and Therapeutic Target for Pancreatic Cancer. *Cold Spring Harb Perspect Med,* 8.

Wei, W. T., Nian, X. X., Wang, S. Y., Jiao, H. L., Wang, Y. X., Xiao, Z. Y., Yang, R. W., Ding, Y. Q., Ye, Y. P. & Liao, W. T. 2017. miR-422a inhibits cell proliferation in colorectal cancer by targeting AKT1 and MAPK1. *Cancer Cell Int,* 17**,** 91.

Weise, K. H., Carlsson, F. H., Joubert, F. J. & Strydom, D. J. 1973. Snake venom toxins. The purification of toxins VII1 and VII2, two cytotoxin homologues from banded Egyptian cobra (Naja haje annulifera) venom, and the complete amino acid sequence of toxin VII1. *Hoppe Seylers Z Physiol Chem,* 354**,** 1317-26.

Werner, U., Haschke, G., Herling, A. W. & Kramer, W. 2010. Pharmacological profile of lixisenatide: A new GLP-1 receptor agonist for the treatment of type 2 diabetes. *Regul Pept,* 164**,** 58-64.

Wheeler, D. & Bhagwat, M. 2007. BLAST QuickStart: Example-Driven Web-Based BLAST Tutorial. *In:* NH, B. (ed.) *Comparative Genomics: Volumes 1 and 2.* https://www.ncbi.nlm.nih.gov/books/NBK1734/: Humana Press.

Winston, R., Weiss, G., Baldwin, A., Frank, S., Tims, L., Lashawn Brooks, N. & Kundranda, M. 2015. Effect of renin-angiotensin system (RAS) inhibition on survival in advanced ductal adenocarcinoma (PDAC). *Journal of Clinical Oncology,* 33.

Wray, K. P., Margres, M. J., Seavy, M. & Rokyta, D. R. 2015. Early significant ontogenetic changes in snake venoms. *Toxicon,* 96**,** 74-81.

Wurbel, H. 2017. More than 3Rs: the importance of scientific validity for harm-benefit analysis of animal research. *Lab Anim (NY),* 46**,** 164-166.

Wüster, W. & Broadley, D. G. 2004. A review of the southern African 'non-spitting' cobras (Serpentes : Elapidae : Naja) : original article.

Wüster, W. & Broadley, D. G. 2007. Get an eyeful of this: a new species of giant spitting cobra from eastern and north-eastern Africa (Squamata: Serpentes: Elapidae: Naja ). *Zootaxa; Vol 1532, No 1: 26 Jul. 2007DO - 10.11646/zootaxa.1532.1.4*.

Wüster, W., Crookes, S., Ineich, I., Mane, Y., Pook, C. E., Trape, J. F. & Broadley, D. G. 2007. The phylogeny of cobras inferred from mitochondrial DNA sequences: evolution of venom spitting and the phylogeography of the African spitting cobras (Serpentes: Elapidae: Naja nigricollis complex). *Mol Phylogenet Evol,* 45**,** 437-53.

Xie, F., Smith, R. D. & Shen, Y. 2012. Advanced proteomic liquid chromatography. *J Chromatogr A,* 1261**,** 78-90.

Xu, D. C., Arthurton, L. & Baena-Lopez, L. A. 2018a. Learning on the Fly: The Interplay between Caspases and Cancer. *Biomed Res Int,* 2018**,** 5473180.

Xu, R. L., He, W., Tang, J., Guo, W., Zhuang, P., Wang, C. Q., Fu, W. M. & Zhang, J. F. 2018b. Primate-specific miRNA-637 inhibited tumorigenesis in human pancreatic ductal adenocarcinoma cells by suppressing Akt1 expression. *Exp Cell Res,* 363**,** 310-314.

Xu, X., Ehdaie, B., Ohara, N., Yoshino, T. & Deng, C. X. 2010. Synergistic action of Smad4 and Pten in suppressing pancreatic ductal adenocarcinoma formation in mice. *Oncogene,* 29**,** 674-86.

Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R. H., Eshleman, J. R., Nowak, M. A., Velculescu, V. E., Kinzler, K. W., Vogelstein, B. & Iacobuzio-Donahue, C. A. 2010. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature,* 467**,** 1114-7.

Yamanouye, N., Kerchove, C. M., Moura-Da-Silva, A. M., Carneiro, S. M. & Markus, R. P. 2006. Long-term primary culture of secretory cells of Bothrops jararaca venom gland for venom production in vitro. *Nat Protoc,* 1**,** 2763-6.

Yan, H., Xiang, P., Zhang, J., Xie, L. & Shen, M. 2017. Dynamic changes of serum protein in rats with acute intoxication of Chinese cobra snake venom by proteomic analysis. *Forensic Sci Res,* 5**,** 309-321.

Yan, J., Zhang, H., Xiang, J., Zhao, Y., Yuan, X., Sun, B. & Lin, A. 2018a. The BH3-only protein BAD mediates TNFalpha cytotoxicity despite concurrent activation of IKK and NF-kappaB in septic shock. *Cell Res,* 28**,** 701-718.

Yan, W., Liu, Z., Yang, W. & Wu, G. 2018b. miRNA expression profiles in Smad4-positive and Smad4-negative SW620 human colon cancer cells detected by next-generation small RNA sequencing. *Cancer Manag Res,* 10**,** 5479-5490.

Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., Stephens, R. M., Okamoto, A., Yokota, J., Tanaka, T., Calin, G. A., Liu, C. G., Croce, C. M. & Harris, C. C. 2006. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell,* 9**,** 189-98.

Yang, B., Tang, F., Zhang, B., Zhao, Y., Feng, J. & Rao, Z. 2014. Matrix metalloproteinase-9 overexpression is closely related to poor prognosis in patients with colon cancer. *World J Surg Oncol,* 12**,** 24.

Yang, F., Qu, W., Du, M., Mai, Z., Wang, B., Ma, Y., Wang, X. & Chen, T. 2020a. Stoichiometry and regulation network of Bcl-2 family complexes quantified by live-cell FRET assay. *Cell Mol Life Sci,* 77**,** 2387-2406.

Yang, J. L., Qu, X. J., Russell, P. J. & Goldstein, D. 2004. Regulation of epidermal growth factor receptor in human colon cancer cell lines by interferon alpha. *Gut,* 53**,** 123-9.

Yang, Z. Y., He, J. H., Lu, A. P., Hou, T. J. & Cao, D. S. 2020b. Frequent hitters: nuisance artifacts in high-throughput screening. *Drug Discov Today,* 25**,** 657-667.

Yasutome, M., Gunn, J. & Korc, M. 2005. Restoration of Smad4 in BxPC3 pancreatic cancer cells attenuates proliferation without altering angiogenesis. *Clin Exp Metastasis,* 22**,** 461-73.

Yuan, J. S., Reed, A., Chen, F. & Stewart, C. N., Jr. 2006. Statistical analysis of real-time PCR data. *BMC Bioinformatics,* 7**,** 85.

Yurgelun, M. B., Kulke, M. H., Fuchs, C. S., Allen, B. A., Uno, H., Hornick, J. L., Ukaegbu, C. I., Brais, L. K., Mcnamara, P. G., Mayer, R. J., Schrag, D., Meyerhardt, J. A., Ng, K., Kidd, J., Singh, N., Hartman, A. R., Wenstrup, R. J. & Syngal, S. 2017. Cancer Susceptibility Gene Mutations in Individuals With Colorectal Cancer. *J Clin Oncol,* 35**,** 1086-1095.

Zamani, A., Sääksjärvi, I. E. & Prendini, L. 2021. Amateur venom-extraction business may hasten extinction of scorpions. *Arachnologische Mitteilungen: Arachnology Letters,* 61**,** 20-23.

Zare, M., Amin, M. M., Nikaeen, M., Bina, B., Pourzamani, H., Fatehizadeh, A. & Taheri, E. 2015. Resazurin reduction assay, a useful tool for assessment of heavy metal toxicity in acidic conditions. *Environ Monit Assess,* 187**,** 276.

Zhang, B., Li, F., Chen, Z., Shrivastava, I. H., Gasanoff, E. S. & Dagda, R. K. 2019a. Naja mossambica mossambica Cobra Cardiotoxin Targets Mitochondria to Disrupt Mitochondrial Membrane Structure and Function. *Toxins,* 11**,** 152.

Zhang, F. C., Sun, Z. Y., Liao, L. P., Zuo, Y., Zhang, D., Wang, J., Chen, Y. T., Xiao, S. H., Jiang, H., Lu, T., Xu, P., Yue, L. Y., Du, D. H., Zhang, H., Liu, C. P. & Luo, C. 2020. Discovery of novel CBP bromodomain inhibitors through TR-FRET-based high-throughput screening. *Acta Pharmacol Sin,* 41**,** 286-292.

Zhang, H., Cui, W. & Gross, M. L. 2014. Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies. *FEBS Lett,* 588**,** 308-17.

Zhang, H. X., Du, G. H. & Zhang, J. T. 2004. Assay of mitochondrial functions by resazurin in vitro. *Acta Pharmacol Sin,* 25**,** 385-9.

Zhang, J. H., Chung, T. D. & Oldenburg, K. R. 1999. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of biomolecular screening,* 4**,** 67-73.

Zhang, L. & Wei, L.-J. 2007. ACTX-8, a cytotoxic L-amino acid oxidase isolated from Agkistrodon acutus snake venom, induces apoptosis in Hela cervical cancer cells. *Life sciences,* 80**,** 1189-97.

Zhang, Q., Xiao, M., Gu, S., Xu, Y., Liu, T., Li, H., Yu, Y., Qin, L., Zhu, Y., Chen, F., Wang, Y., Ding, C., Wu, H., Ji, H., Chen, Z., Zu, Y., Malkoski, S., Li, Y., Liang, T., Ji, J., Qin, J., Xu, P., Zhao, B., Shen, L., Lin, X. & Feng, X. H. 2019b. ALK phosphorylates SMAD4 on tyrosine to disable TGF-beta tumour suppressor functions. *Nat Cell Biol,* 21**,** 179-189.

Zhang, X., Fang, A., Riley, C. P., Wang, M., Regnier, F. E. & Buck, C. 2010. Multi-dimensional liquid chromatography in proteomics--a review. *Anal Chim Acta,* 664**,** 101-13.

Zhang, X., Hu, F., Li, G., Li, G., Yang, X., Liu, L., Zhang, R., Zhang, B. & Feng, Y. 2018. Human colorectal cancer-derived mesenchymal stem cells promote colorectal cancer progression through IL-6/JAK2/STAT3 signaling. *Cell Death Dis,* 9**,** 25.

Zhao, M., Lei, C., Yang, Y., Bu, X., Ma, H., Gong, H., Liu, J., Fang, X., Hu, Z. & Fang, Q. 2015. Abraxane, the Nanoparticle Formulation of Paclitaxel Can Induce Drug Resistance by Up-Regulation of P-gp. *PLoS One,* 10**,** e0131429.

Zhao, M., Mishra, L. & Deng, C. X. 2018. The role of TGF-β/SMAD4 signaling in cancer. *Int J Biol Sci,* 14**,** 111-23.

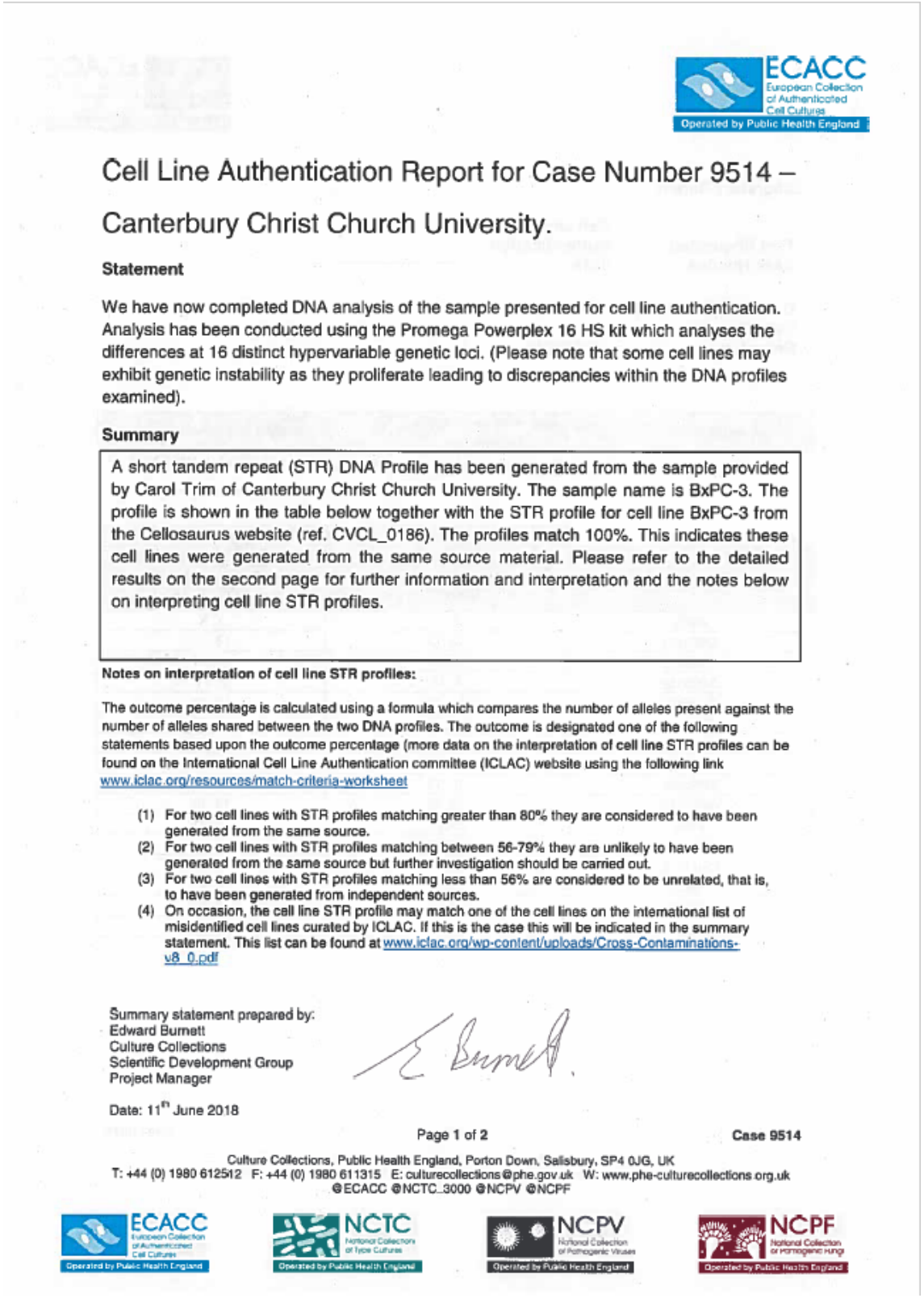
Zhao, R., Choi, B. Y., Lee, M. H., Bode, A. M. & Dong, Z. 2016. Implications of Genetic and Epigenetic Alterations of CDKN2A (p16(INK4a)) in Cancer. *EBioMedicine,* 8**,** 30-39.

Zrimšek, P., Kunc, J., Kosec, M. & Mrkun, J. 2004. Spectrophotometric application of resazurin reduction assay to evaluate boar semen quality. *International Journal of Andrology,* 27**,** 57-62.

# Appendix Contents

# Appendix I: STR Profiling reports

## BxPC-3



ECACC
European Collection of Authenticated Cell Cultures
Operated by Public Health England

Cell Line Authentication Report for Case Number 9514 – Canterbury Christ Church University.

**Statement**

We have now completed DNA analysis of the sample presented for cell line authentication. Analysis has been conducted using the Promega Powerplex 16 HS kit which analyses the differences at 16 distinct hypervariable genetic loci. (Please note that some cell lines may exhibit genetic instability as they proliferate leading to discrepancies within the DNA profiles examined).

**Summary**

A short tandem repeat (STR) DNA Profile has been generated from the sample provided by Carol Trim of Canterbury Christ Church University. The sample name is BxPC-3. The profile is shown in the table below together with the STR profile for cell line BxPC-3 from the Cellosaurus website (ref. CVCL_0186). The profiles match 100%. This indicates these cell lines were generated from the same source material. Please refer to the detailed results on the second page for further information and interpretation and the notes below on interpreting cell line STR profiles.

**Notes on interpretation of cell line STR profiles:**

The outcome percentage is calculated using a formula which compares the number of alleles present against the number of alleles shared between the two DNA profiles. The outcome is designated one of the following statements based upon the outcome percentage (more data on the interpretation of cell line STR profiles can be found on the International Cell Line Authentication committee (ICLAC) website using the following link www.iclac.org/resources/match-criteria-worksheet

(1) For two cell lines with STR profiles matching greater than 80% they are considered to have been generated from the same source.
(2) For two cell lines with STR profiles matching between 56-79% they are unlikely to have been generated from the same source but further investigation should be carried out.
(3) For two cell lines with STR profiles matching less than 56% are considered to be unrelated, that is, to have been generated from independent sources.
(4) On occasion, the cell line STR profile may match one of the cell lines on the international list of misidentified cell lines curated by ICLAC. If this is the case this will be indicated in the summary statement. This list can be found at www.iclac.org/wp-content/uploads/Cross-Contaminations-v8_0.pdf

Summary statement prepared by:
Edward Burnett
Culture Collections
Scientific Development Group
Project Manager

Date: 11th June 2018

Page 1 of 2 Case 9514

Culture Collections, Public Health England, Porton Down, Salisbury, SP4 0JG, UK
T: +44 (0) 1980 612512 F: +44 (0) 1980 611315 E: culturecollections@phe.gov.uk W: www.phe-culturecollections.org.uk
@ECACC @NCTC_3000 @NCPV @NCPF

ECACC Operated by Public Health England
NCTC National Collection of Type Cultures Operated by Public Health England
NCPV National Collection of Pathogenic Viruses Operated by Public Health England
NCPF National Collection of Pathogenic Fungi Operated by Public Health England

ECACC
European Collection of Authenticated Cell Cultures
Operated by Public Health England

**Laboratory Report**

| | |
|---|---|
| **Test Requested** | **Cell Line Authentication** |
| **Case Number** | **9514** |
| **Date Sample Tested** | **15/3/2018** |
| **Date Sample Reported** | **21/3/2018** |

| Sample Name | Sample/Comparison Profile Source | Sample Reference |
|---|---|---|
| BxPC-3 | Cellosaurus BxPC-3 CVCL_0186 | RS_05_D1015669_9514_LW6000087-18 |

| Table of Allelic Data | Genotypes | |
|---|---|---|
| **STR Locus** | **Test Sample – BxPC-3 – Canterbury Christ Church University** | **Database Sample – Cellosaurus BxPC-3 CVCL_0186** |
| AMEL | X, X | X, X |
| CSF1PO | 13, 13 | 13, 13 |
| D13S317 | 11, 11 | 11, 11 |
| D16S539 | 9, 11 | 9, 11 |
| D18S51 | 12, 12 | 12, 12 |
| D21S11 | 29, 29 | 29, 29 |
| D3S1358 | 14, 16 | 14, 16 |
| D5S818 | 11, 11 | 11, 11 |
| D7S820 | 10, 13 | 10, 13 |
| D8S1179 | 13, 13 | 13, 13 |
| FGA | 20, 21 | 20, 21 |
| PENTA D | 14, 14 | 14, 14 |
| PENTA E | 12, 14 | 12, 14 |
| THO1 | 9, 9 | 9, 9 |
| TPOX | 8, 8 | 8, 8 |
| vWA | 14, 18 | 14, 18 |



Culture Collections, Public Health England, Porton Down, Salisbury, SP4 0JG, UK
T: +44 (0) 1980 612512 F: +44 (0) 1980 611315 E: culturecollections@phe.gov.uk W: www.phe-culturecollections.org.uk
@ECACC @NCTC_3000 @NCPV @NCPF

ECACC European Collection of Authenticated Cell Cultures Operated by Public Health England
NCTC National Collection of Type Cultures Operated by Public Health England
NCPV National Collection of Pathogenic Viruses Operated by Public Health England
NCPF National Collection of Pathogenic Fungi Operated by Public Health England

ECACC
European Collection of Authenticated Cell Cultures
Operated by Public Health England

# Cell Line Authentication Report for Case Number 9514 – Canterbury Christ Church University.

**Statement**

We have now completed DNA analysis of the sample presented for cell line authentication. Analysis has been conducted using the Promega Powerplex 16 HS kit which analyses the differences at 16 distinct hypervariable genetic loci. (Please note that some cell lines may exhibit genetic instability as they proliferate leading to discrepancies within the DNA profiles examined).

**Summary**

A short tandem repeat (STR) DNA Profile has been generated from the sample provided by Carol Trim of Canterbury Christ Church University. The sample name is MIA PaCa2. The profile is shown in the table below together with the STR profile for cell line MIA PaCa2 from the Cellosaurus website (ref. CVCL_0428). The profiles match 27% (8/30 alleles). This indicates these cell lines were generated from **different** source material. However, the Canterbury Christ Church University MIA PaCa2 cell lines matches cell line SW620 (Cellosaurus ref CVCL_0547) 100% meaning these cell lines were derived from the same source material. Please refer to the detailed results on the second page for further information and interpretation and the notes below on interpreting cell line STR profiles.

**Notes on interpretation of cell line STR profiles:**

The outcome percentage is calculated using a formula which compares the number of alleles present against the number of alleles shared between the two DNA profiles. The outcome is designated one of the following statements based upon the outcome percentage (more data on the interpretation of cell line STR profiles can be found on the International Cell Line Authentication committee (ICLAC) website using the following link www.iclac.org/resources/match-criteria-worksheet

(1) For two cell lines with STR profiles matching greater than 80% they are considered to have been generated from the same source.
(2) For two cell lines with STR profiles matching between 56-79% they are unlikely to have been generated from the same source but further investigation should be carried out.
(3) For two cell lines with STR profiles matching less than 56% are considered to be unrelated, that is, to have been generated from independent sources.
(4) On occasion, the cell line STR profile may match one of the cell lines on the international list of misidentified cell lines curated by ICLAC. If this is the case this will be indicated in the summary statement. This list can be found at www.iclac.org/wp-content/uploads/Cross-Contaminations-v8_0.pdf

Summary statement prepared by:
Edward Burnett
Culture Collections
Scientific Development Group
Project Manager

E. Burnett.

Date: 11th June 2018



Culture Collections, Public Health England, Porton Down, Salisbury, SP4 0JG, UK
T: +44 (0) 1980 612512 F: +44 (0) 1980 611315 E: culturecollections@phe.gov.uk W: www.phe-culturecollections.org.uk
@ECACC @NCTC_3000 @NCPV @NCPF

ECACC European Collection of Authenticated Cell Cultures Operated by Public Health England
NCTC National Collection of Type Cultures Operated by Public Health England
NCPV National Collection of Pathogenic Viruses Operated by Public Health England
NCPF National Collection of Pathogenic Fungi Operated by Public Health England

ECACC
European Collection of Authenticated Cell Cultures
Operated by Public Health England

**Laboratory Report**

| | |
|---|---|
| **Test Requested** | **Cell Line Authentication** |
| **Case Number** | **9514** |
| **Date Sample Tested** | **15/3/2018** |
| **Date Sample Reported** | **21/3/2018** |

| Sample Name | Sample/Comparison Profile Source | Sample Reference |
|---|---|---|
| MIA PaCa2 | Cellosaurus MIA PaCa2 CVCL_0428 | RS_04_D1015667_9514_LW6000087-18 |

| Table of Allelic Data | Genotypes | |
|---|---|---|
| **STR Locus** | **Test Sample – MIA PaCa2 – Canterbury Christ Church University** | **Database Sample – Cellosaurus MIA PaCa2 CVCL_0428** |
| AMEL | X, X | X, X |
| CSF1PO | 13, 14 | 10, 10 |
| D13S317 | 12, 12 | 12, 13 |
| D16S539 | 9, 13 | 10, 13 |
| D18S51 | n/d | 12, 12 |
| D21S11 | 30, 30.2 | 29, 31 |
| D3S1358 | 16, 16 | 16, 16 |
| D5S818 | 13, 13 | 13, 13 |
| D7S820 | 8, 9 | 12, 13 |
| D8S1179 | 13, 13 | 16, 16 |
| FGA | 24, 24 | 22, 22 |
| PENTA D | 9, 15 | 12, 16 |
| PENTA E | 10, 10 | 13, 18 |
| TH01 | 8, 18 | 9, 10 |
| TPOX | 11, 11 | 9, 9 |
| vWA | 16, 16 | 15, 15 |



Culture Collections, Public Health England, Porton Down, Salisbury, SP4 0JG, UK
T: +44 (0) 1980 612512 F: +44 (0) 1980 611315 E: culturecollections@phe.gov.uk W: www.phe-culturecollections.org.uk
@ECACC @NCTC_3000 @NCPV @NCPF

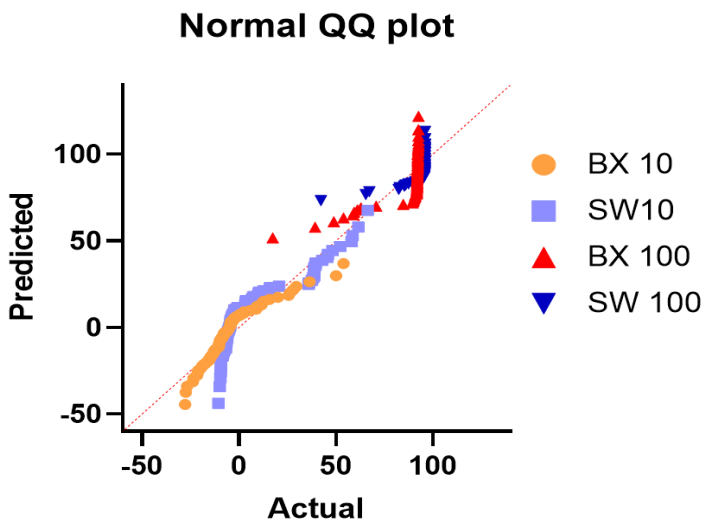
ECACC European Collection of Authenticated Cell Cultures Operated by Public Health England
NCTC National Collection of Type Cultures Operated by Public Health England
NCPV National Collection of Pathogenic Viruses Operated by Public Health England
NCPF National Collection of Pathogenic Fungi Operated by Public Health England

# Appendix II: Statistical analysis

## Cobra Screen

To assess whether the cobra screen data from section 3.2.2 were normally distributed, normality tests and associated quantile-quantile (QQ) plots were performed. Results are as follows:

| Tests for normal distribution | | | | |
|---|---|---|---|---|
| **Anderson-Darling test** | | | | |
| A2* | 2.992 | 5.423 | 17.77 | 18.81 |
| P value | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| Passed normality test (alpha=0.05)? | No | No | No | No |
| P value summary | **** | **** | **** | **** |
| **D'Agostino & Pearson test** | | | | |
| K2 | 27.62 | 11 | 66.42 | 104.5 |
| P value | <0.0001 | 0.0041 | <0.0001 | <0.0001 |
| Passed normality test (alpha=0.05)? | No | No | No | No |
| P value summary | **** | ** | **** | **** |
| **Shapiro-Wilk test** | | | | |
| W | 0.8755 | 0.821 | 0.4465 | 0.3634 |
| P value | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| Passed normality test (alpha=0.05)? | No | No | No | No |
| P value summary | **** | **** | **** | **** |
| **Kolmogorov-Smirnov test** | | | | |
| KS distance | 0.1886 | 0.2128 | 0.4342 | 0.4009 |
| P value | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| Passed normality test (alpha=0.05)? | No | No | No | No |
| P value summary | **** | **** | **** | **** |
| Number of values | 76 | 72 | 71 | 75 |



### *Kruskal-Wallis tests*

As none of the data sets were normally distributed, Kruskal-Wallis with Dunn's multiple comparison tests were selected for assessment of statistical significance:

### *BxPC-3 10 µg/ml*

| Kruskal-Wallis test | |
|---|---|
| P value | 0.0001 |
| Exact or approximate P value? | Gaussian Approximation |
| P value summary | *** |
| Do the medians vary signif. (P < 0.05) | Yes |
| Number of groups | 21 |
| Kruskal-Wallis statistic | 52.2 |

| Dunn's Multiple Comparison Test | Difference in rank sum | Significant? P < 0.05? | Summary |
|---|---|---|---|
| N.aan vs Neg con | -47.06 | No | ns |
| N.anc vs Neg con | -28.06 | No | ns |
| N.haj vs Neg con | -31.31 | No | ns |
| N.hle vs Neg con | -50.31 | Yes | * |
| N.kao vs Neg con | -45.56 | No | ns |
| N.niv vs Neg con | -23.81 | No | ns |
| N.nub vs Neg con | -50.56 | Yes | * |

### *BxPC-3 100 µg/ml*

| Kruskal-Wallis test | |
|---|---|
| P value | P<0.0001 |
| Exact or approximate P value? | Gaussian Approximation |
| P value summary | *** |
| Do the medians vary signif. (P < 0.05) | Yes |
| Number of groups | 21 |
| Kruskal-Wallis statistic | 99.37 |

| Dunn's Multiple Comparison Test | Difference in rank sum | Significant? P < 0.05? | Summary |
|---|---|---|---|
| N.aan vs Neg | 22.93 | No | ns |
| N.anc vs Neg | 22.34 | No | ns |
| N.atr vs Neg | 49.34 | No | ns |
| N.hle vs Neg | 49.97 | No | ns |
| N.kao vs Neg | 46.47 | No | ns |
| N.mel vs Neg | 65.22 | Yes | ** |
| N.mos vs Neg | 75.09 | Yes | *** |
| N.naj vs Neg | 45.59 | No | ns |
| N.nct vs Neg | 66.84 | Yes | ** |
| N.nig vs Neg | 73.34 | Yes | *** |
| N.niv vs Neg | 35.26 | No | ns |
| N.nka vs Neg | 58.59 | Yes | * |
| N.nub vs Neg | 81.76 | Yes | *** |
| N.pal vs Neg | 66.72 | Yes | ** |
| N.sam vs Neg | 38.72 | No | ns |
| N.sia vs Neg | 71.09 | Yes | *** |
| N.spu vs Neg | 67.97 | Yes | ** |

*SW620 10 µg/ml*

| Kruskal-Wallis test | |
|---|---|
| P value | P<0.0001 |
| Exact or approximate P value? | Gaussian Approximation |
| P value summary | *** |
| Do the medians vary signif. (P < 0.05) | Yes |
| Number of groups | 21 |
| Kruskal-Wallis statistic | 76.51 |

| Dunn's Multiple Comparison Test | Difference in rank sum | Significant? P < 0.05? | Summary |
|---|---|---|---|
| N.mel vs Neg | -24.36 | No | ns |
| N.mos vs Neg | 47.22 | Yes | * |
| N.nct vs Neg | 49.72 | Yes | * |
| N.nig vs Neg | 49.09 | Yes | * |
| N.niv vs Neg | -11.36 | No | ns |
| N.nka vs Neg | -42.36 | No | ns |
| N.nub vs Neg | 50.22 | Yes | * |
| N.pal vs Neg | 40.47 | No | ns |

*SW620 100 µg/ml*

| Kruskal-Wallis test | |
|---|---|
| P value | P<0.0001 |
| Exact or approximate P value? | Gaussian Approximation |
| P value summary | *** |
| Do the medians vary signif. (P < 0.05) | Yes |
| Number of groups | 21 |
| Kruskal-Wallis statistic | 99.88 |

| Dunn's Multiple Comparison Test | Difference in rank sum | Significant? P < 0.05? | Summary |
|---|---|---|---|
| N.aan vs Neg | 54 | Yes | * |
| N.anc vs Neg | 25 | No | ns |
| N.atr vs Neg | 52.5 | No | ns |
| N.haj vs Neg | 42.13 | No | ns |
| N.hle vs Neg | 55.75 | Yes | * |
| N.kao vs Neg | 59.63 | Yes | * |
| N.mel vs Neg | 69.13 | Yes | ** |
| N.mos vs Neg | 79.25 | Yes | *** |
| N.naj vs Neg | 37.88 | No | ns |
| N.nct vs Neg | 56.5 | Yes | * |
| N.nig vs Neg | 47.5 | No | ns |
| N.niv vs Neg | 70.67 | Yes | ** |
| N.nka vs Neg | 55.75 | Yes | * |
| N.nub vs Neg | 82.75 | Yes | *** |
| N.pal vs Neg | 73.5 | Yes | *** |
| N.sam vs Neg | 21.25 | No | ns |
| N.sia vs Neg | 60.75 | Yes | * |
| N.spu vs Neg | 57 | Yes | * |
| O.han vs Neg | 20 | No | ns |
| Pos vs Neg | 94.43 | Yes | *** |

## First dimension RP HPLC

### *Normality testing*

### *N.mos normality test*

Shapiro-Wilk test

| | N.mos_r1 | N.mos_r2 | N.mos_r3 | N.mos_r4 | N.mos_r5 | N.mos_r6 | N.mos_r7 | N.mos_r8 | N.mos_r9 | N.mos_r10 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.9787 | 0.8559 | 0.9388 | 0.809 | 0.8286 | 0.9707 | 0.9317 | 0.8988 | 0.9281 | 0.9821 |
| P value | 0.7201 | 0.2565 | 0.5224 | 0.1362 | 0.1849 | 0.6713 | 0.4952 | 0.3816 | 0.4816 | 0.7439 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.mos_r11 | N.mos_r12 | N.mos_r13 | N.mos_r14 | N.mos_r15 | N.mos_r16 | N.mos_r17 | N.mos_r18 | N.mos_r19 | N.mos_r20 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.7662 | 0.8043 | 0.9885 | 0.9829 | 0.9216 | 0.9814 | 0.9922 | 0.9721 | 0.9639 | 0.9811 |
| P value | 0.0362 | 0.1249 | 0.795 | 0.7498 | 0.4581 | 0.7384 | 0.831 | 0.6794 | 0.6349 | 0.7366 |
| Passed normality test (alpha=0.05)? | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | * | ns | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.mos_r21 | N.mos_r22 | N.mos_r23 | N.mos_r24 | N.mos_r25 | N.mos_r26 | N.mos_r27 | N.mos_r28 | N.mos Crude | Neg con | Pos con |
|---|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.87 | 0.9966 | 0.8477 | 0.9913 | 0.9851 | 0.846 | 0.9403 | 0.9998 | 0.907 | 0.8231 | 0.9488 |
| P value | 0.2956 | 0.8885 | 0.2342 | 0.8215 | 0.766 | 0.2299 | 0.5286 | 0.9725 | 0.4082 | 0.0173 | 0.5063 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes |
| P value summary | ns | ns | ns | ns | ns | ns | ns | ns | ns | * | ns |

*N.nct normality test*

Shapiro-Wilk test

| | N.nct_r1 | N.nct_r2 | N.nct_r3 | N.nct_r4 | N.nct_r5 | N.nct_r6 | N.nct_r7 | N.nct_r8 | N.nct_r9 | N.nct_r10 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.9358 | 0.9803 | 0.7546 | 0.9618 | 0.9982 | 0.7662 | 0.7594 | 0.9677 | 0.9276 | 0.8597 |
| P value | 0.5106 | 0.7309 | 0.0102 | 0.6242 | 0.919 | 0.0361 | 0.0208 | 0.6548 | 0.4796 | 0.2667 |
| Passed normality test (alpha=0.05)? | Yes | Yes | No | Yes | Yes | No | No | Yes | Yes | Yes |
| P value summary | ns | ns | * | ns | ns | * | * | ns | ns | ns |

| | N.nct_r11 | N.nct_r12 | N.nct_r13 | N.nct_r14 | N.nct_r15 | N.nct_r16 | N.nct_r17 | N.nct_r18 | N.nct_r19 | N.nct_r20 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.9888 | 0.9458 | 0.9349 | 0.9578 | 0.9112 | 0.9494 | 0.8762 | 0.8216 | 0.7835 | 0.9505 |
| P value | 0.7971 | 0.5514 | 0.5074 | 0.6048 | 0.4221 | 0.5665 | 0.3133 | 0.1673 | 0.0756 | 0.5715 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.nct_r21 | N.nct_r22 | N.nct Crude | Neg con | Pos con |
|---|---|---|---|---|---|
| W | 0.9989 | 0.782 | 0.8263 | 0.8218 | 0.9488 |
| P value | 0.9368 | 0.0723 | 0.179 | 0.0168 | 0.5064 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | No | Yes |
| P value summary | ns | ns | ns | * | ns |

*N.nig normality test*

Shapiro-Wilk test

| | N.nig_r1 | N.nig_r2 | N.nig_r3 | N.nig_r4 | N.nig_r5 | N.nig_r6 | N.nig_r7 | N.nig_r8 | N.nig_r9 | N.nig_r10 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.933 | 0.9292 | 0.8887 | 0.9166 | 0.8042 | 0.9541 | 0.823 | 0.838 | 0.7796 | 0.952 |
| P value | 0.5 | 0.4855 | 0.3504 | 0.4404 | 0.1247 | 0.5878 | 0.1706 | 0.2088 | 0.0666 | 0.5782 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.nig_r11 | N.nig_r12 | N.nig_r13 | N.nig_r14 | N.nig_r15 | N.nig_r16 | N.nig_r17 | N.nig_r18 | N.nig_r19 | N.nig_r20 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.8318 | 0.8767 | 0.8746 | 0.9287 | 0.9898 | 0.9996 | 0.9199 | 0.9354 | 0.9782 | 0.9592 |
| P value | 0.193 | 0.3148 | 0.3086 | 0.4837 | 0.8071 | 0.9631 | 0.452 | 0.509 | 0.7169 | 0.6116 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.nig_r21 | N.nig_r22 | N.nig_r23 | N.nig_r24 | N.nig_r25 | N.nig Crude | Neg con | Pos con |
|---|---|---|---|---|---|---|---|---|
| W | 0.779 | 0.8421 | N too small | N too small | N too small | 0.9494 | 0.8218 | 0.9488 |
| P value | 0.0653 | 0.2196 | | | | 0.5665 | 0.0168 | 0.5064 |
| Passed normality test (alpha=0.05)? | Yes | Yes | | | | Yes | No | Yes |
| P value summary | ns | ns | | | | ns | * | ns |

*N.nub normality test*

Shapiro-Wilk test

| | N.nub_r1 | N.nub_r2 | N.nub_r3 | N.nub_r4 | N.nub_r5 | N.nub_r6 | N.nub_r7 | N.nub_r8 | N.nub_r9 | N.nub_r10 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.8806 | 0.7546 | 0.8338 | 0.837 | 0.8088 | 0.9932 | 0.7961 | 0.8819 | 0.9354 | 0.8818 |
| P value | 0.3262 | 0.0102 | 0.1981 | 0.2062 | 0.1358 | 0.8421 | 0.1053 | 0.3301 | 0.5092 | 0.3296 |
| Passed normality test (alpha=0.05)? | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | ns | * | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.nub_r11 | N.nub_r12 | N.nub_r13 | N.nub_r14 | N.nub_r15 | N.nub_r16 | N.nub_r17 | N.nub_r18 | N.nub_r19 | N.nub_r20 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.8282 | 0.9882 | 0.9275 | 0.9937 | 0.8032 | 0.8092 | 0.8893 | 0.9058 | 0.9959 | 0.9979 |
| P value | 0.1839 | 0.7922 | 0.4792 | 0.8479 | 0.1222 | 0.1367 | 0.3523 | 0.4041 | 0.8769 | 0.9132 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | ns | ns | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.nub_r21 | N.nub Crude | Neg con | Pos con |
|---|---|---|---|---|
| W | 0.9897 | 0.8451 | 0.8521 | 0.9488 |
| P value | 0.8056 | 0.2275 | 0.0786 | 0.5063 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | Yes |
| P value summary | ns | ns | ns | ns |

*N.pal normality test*

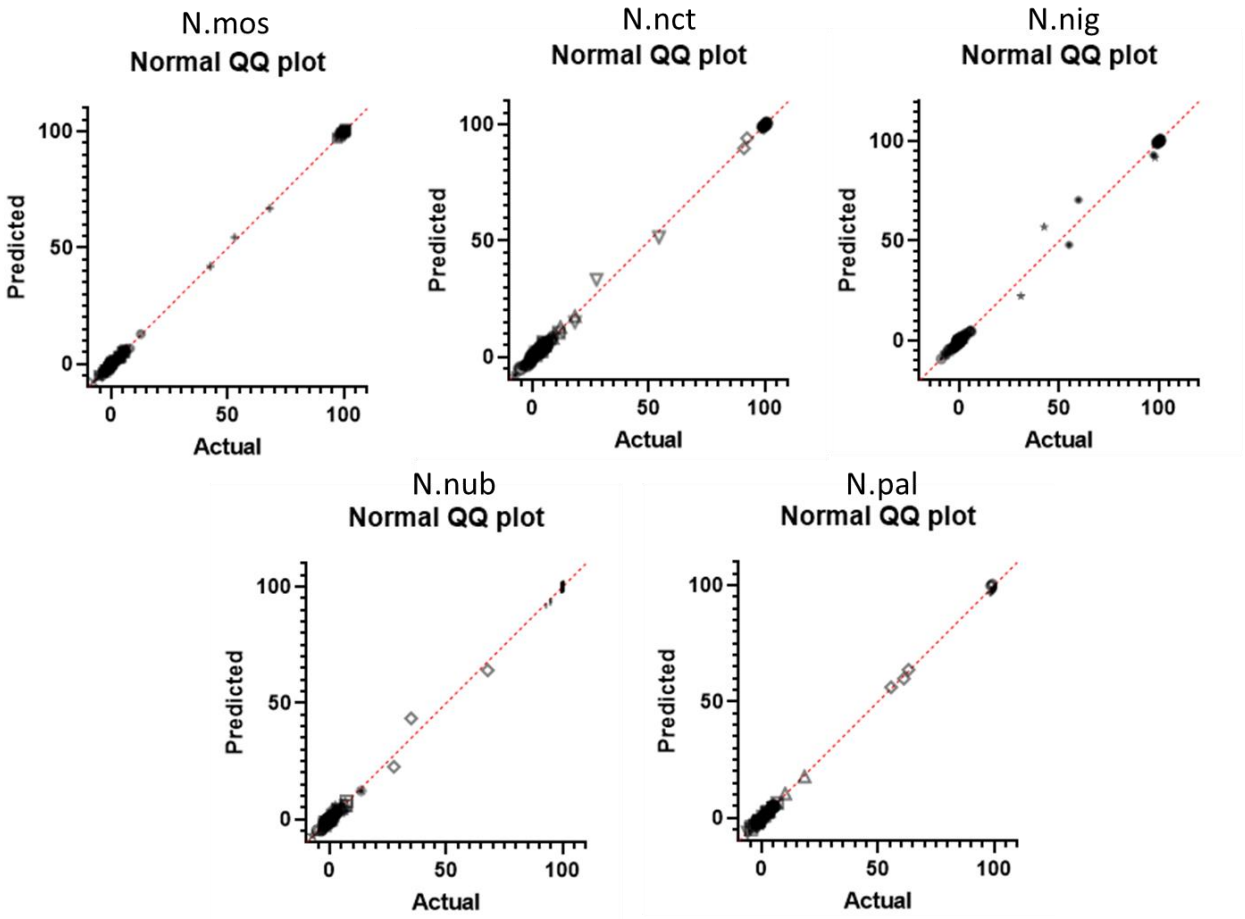
Shapiro-Wilk test

| | N.pal_r1 | N.pal_r2 | N.pal_r3 | N.pal_r4 | N.pal_r5 | N.pal_r6 | N.pal_r7 | N.pal_r8 | N.pal_r9 | N.pal_r10 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.9065 | 0.7548 | 0.9871 | 0.9344 | 0.9952 | 0.9246 | 0.7994 | 0.9972 | 0.9898 | 0.9367 |
| P value | 0.4065 | 0.0106 | 0.7826 | 0.5052 | 0.868 | 0.4687 | 0.113 | 0.8994 | 0.807 | 0.5141 |
| Passed normality test (alpha=0.05)? | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | ns | * | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.pal_r11 | N.pal_r12 | N.pal_r13 | N.pal_r14 | N.pal_r15 | N.pal_r16 | N.pal_r17 | N.pal_r18 | N.pal_r19 | N.pal_r20 |
|---|---|---|---|---|---|---|---|---|---|---|
| W | 0.769 | 0.7595 | 0.8226 | 0.7735 | 0.9999 | 0.9943 | 0.9989 | 0.99 | 0.9947 | 0.9994 |
| P value | 0.0425 | 0.0211 | 0.1696 | 0.0526 | 0.9767 | 0.8562 | 0.9363 | 0.809 | 0.8614 | 0.952 |
| Passed normality test (alpha=0.05)? | No | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | * | * | ns | ns | ns | ns | ns | ns | ns | ns |

| | N.pal Crude | Neg con | Pos con |
|---|---|---|---|
| W | 0.9757 | 0.8231 | 0.9488 |
| P value | 0.7013 | 0.0173 | 0.5063 |
| Passed normality test (alpha=0.05)? | Yes | No | Yes |
| P value summary | ns | * | ns |

*1D RP Fraction QQ plots*



## First Dimension RP Fraction Screen

*Kruskal Wallis with Dunn's multiple comparison tests*

*N.mos KW & Dunn's*

| Kruskal-Wallis test | |
|---|---|
| P value | 0.0001 |
| Exact or approximate P value? | Approximate |
| P value summary | *** |
| Do the medians vary signif. (P < 0.05)? | Yes |
| Number of groups | 28 |
| Kruskal-Wallis statistic | 62.91 |
| | |
| Data summary | |
| Number of treatments (columns) | 28 |
| Number of values (total) | 84 |

| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value |
|---|---|---|---|---|
| 1 vs. 2 | 21 | No | ns | >0.9999 |
| 1 vs. 3 | 12.67 | No | ns | >0.9999 |
| 1 vs. 4 | 20 | No | ns | >0.9999 |
| 1 vs. 5 | 3.333 | No | ns | >0.9999 |
| 1 vs. 6 | 6.333 | No | ns | >0.9999 |
| 1 vs. 7 | 8.667 | No | ns | >0.9999 |
| 1 vs. 8 | 6 | No | ns | >0.9999 |
| 1 vs. 9 | 5 | No | ns | >0.9999 |
| 1 vs. 10 | 0.3333 | No | ns | >0.9999 |
| 1 vs. 11 | 54.67 | No | ns | >0.9999 |
| 1 vs. 12 | 61 | No | ns | 0.8289 |
| 1 vs. 13 | -12.33 | No | ns | >0.9999 |
| 1 vs. 14 | 68 | No | ns | 0.2418 |
| 1 vs. 15 | 30.67 | No | ns | >0.9999 |
| 1 vs. 16 | 50.33 | No | ns | >0.9999 |
| 1 vs. 17 | 39.67 | No | ns | >0.9999 |
| 1 vs. 18 | 20.33 | No | ns | >0.9999 |
| 1 vs. 19 | 60 | No | ns | 0.9792 |
| 1 vs. 20 | 29 | No | ns | >0.9999 |
| 1 vs. 21 | 41 | No | ns | >0.9999 |
| 1 vs. 22 | 25.33 | No | ns | >0.9999 |
| 1 vs. 23 | 55.33 | No | ns | >0.9999 |
| 1 vs. 24 | 33 | No | ns | >0.9999 |
| 1 vs. 25 | 43.67 | No | ns | >0.9999 |
| 1 vs. 26 | 32.33 | No | ns | >0.9999 |
| 1 vs. 27 | 36 | No | ns | >0.9999 |
| 1 vs. 28 | 37.33 | No | ns | >0.9999 |
| 2 vs. 3 | -8.333 | No | ns | >0.9999 |
| 2 vs. 4 | -1 | No | ns | >0.9999 |
| 2 vs. 5 | -17.67 | No | ns | >0.9999 |
| 2 vs. 6 | -14.67 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 2 vs. 7 | -12.33 | No | ns | >0.9999 |
| 2 vs. 8 | -15 | No | ns | >0.9999 |
| 2 vs. 9 | -16 | No | ns | >0.9999 |
| 2 vs. 10 | -20.67 | No | ns | >0.9999 |
| 2 vs. 11 | 33.67 | No | ns | >0.9999 |
| 2 vs. 12 | 40 | No | ns | >0.9999 |
| 2 vs. 13 | -33.33 | No | ns | >0.9999 |
| 2 vs. 14 | 47 | No | ns | >0.9999 |
| 2 vs. 15 | 9.667 | No | ns | >0.9999 |
| 2 vs. 16 | 29.33 | No | ns | >0.9999 |
| 2 vs. 17 | 18.67 | No | ns | >0.9999 |
| 2 vs. 18 | -0.6667 | No | ns | >0.9999 |
| 2 vs. 19 | 39 | No | ns | >0.9999 |
| 2 vs. 20 | 8 | No | ns | >0.9999 |
| 2 vs. 21 | 20 | No | ns | >0.9999 |
| 2 vs. 22 | 4.333 | No | ns | >0.9999 |
| 2 vs. 23 | 34.33 | No | ns | >0.9999 |
| 2 vs. 24 | 12 | No | ns | >0.9999 |
| 2 vs. 25 | 22.67 | No | ns | >0.9999 |
| 2 vs. 26 | 11.33 | No | ns | >0.9999 |
| 2 vs. 27 | 15 | No | ns | >0.9999 |
| 2 vs. 28 | 16.33 | No | ns | >0.9999 |
| 3 vs. 4 | 7.333 | No | ns | >0.9999 |
| 3 vs. 5 | -9.333 | No | ns | >0.9999 |
| 3 vs. 6 | -6.333 | No | ns | >0.9999 |
| 3 vs. 7 | -4 | No | ns | >0.9999 |
| 3 vs. 8 | -6.667 | No | ns | >0.9999 |
| 3 vs. 9 | -7.667 | No | ns | >0.9999 |
| 3 vs. 10 | -12.33 | No | ns | >0.9999 |
| 3 vs. 11 | 42 | No | ns | >0.9999 |
| 3 vs. 12 | 48.33 | No | ns | >0.9999 |
| 3 vs. 13 | -25 | No | ns | >0.9999 |
| 3 vs. 14 | 55.33 | No | ns | >0.9999 |
| 3 vs. 15 | 18 | No | ns | >0.9999 |
| 3 vs. 16 | 37.67 | No | ns | >0.9999 |
| 3 vs. 17 | 27 | No | ns | >0.9999 |
| 3 vs. 18 | 7.667 | No | ns | >0.9999 |
| 3 vs. 19 | 47.33 | No | ns | >0.9999 |
| 3 vs. 20 | 16.33 | No | ns | >0.9999 |
| 3 vs. 21 | 28.33 | No | ns | >0.9999 |
| 3 vs. 22 | 12.67 | No | ns | >0.9999 |
| 3 vs. 23 | 42.67 | No | ns | >0.9999 |
| 3 vs. 24 | 20.33 | No | ns | >0.9999 |
| 3 vs. 25 | 31 | No | ns | >0.9999 |
| 3 vs. 26 | 19.67 | No | ns | >0.9999 |
| 3 vs. 27 | 23.33 | No | ns | >0.9999 |
| 3 vs. 28 | 24.67 | No | ns | >0.9999 |

| | | | | | |
|---|---|---|---|---|---|
| 4 vs. 5 | -16.67 | No | ns | >0.9999 | |
| 4 vs. 6 | -13.67 | No | ns | >0.9999 | |
| 4 vs. 7 | -11.33 | No | ns | >0.9999 | |
| 4 vs. 8 | -14 | No | ns | >0.9999 | |
| 4 vs. 9 | -15 | No | ns | >0.9999 | |
| 4 vs. 10 | -19.67 | No | ns | >0.9999 | |
| 4 vs. 11 | 34.67 | No | ns | >0.9999 | |
| 4 vs. 12 | 41 | No | ns | >0.9999 | |
| 4 vs. 13 | -32.33 | No | ns | >0.9999 | |
| 4 vs. 14 | 48 | No | ns | >0.9999 | |
| 4 vs. 15 | 10.67 | No | ns | >0.9999 | |
| 4 vs. 16 | 30.33 | No | ns | >0.9999 | |
| 4 vs. 17 | 19.67 | No | ns | >0.9999 | |
| 4 vs. 18 | 0.3333 | No | ns | >0.9999 | |
| 4 vs. 19 | 40 | No | ns | >0.9999 | |
| 4 vs. 20 | 9 | No | ns | >0.9999 | |
| 4 vs. 21 | 21 | No | ns | >0.9999 | |
| 4 vs. 22 | 5.333 | No | ns | >0.9999 | |
| 4 vs. 23 | 35.33 | No | ns | >0.9999 | |
| 4 vs. 24 | 13 | No | ns | >0.9999 | |
| 4 vs. 25 | 23.67 | No | ns | >0.9999 | |
| 4 vs. 26 | 12.33 | No | ns | >0.9999 | |
| 4 vs. 27 | 16 | No | ns | >0.9999 | |
| 4 vs. 28 | 17.33 | No | ns | >0.9999 | |
| 5 vs. 6 | 3 | No | ns | >0.9999 | |
| 5 vs. 7 | 5.333 | No | ns | >0.9999 | |
| 5 vs. 8 | 2.667 | No | ns | >0.9999 | |
| 5 vs. 9 | 1.667 | No | ns | >0.9999 | |
| 5 vs. 10 | -3 | No | ns | >0.9999 | |
| 5 vs. 11 | 51.33 | No | ns | >0.9999 | |
| 5 vs. 12 | 57.67 | No | ns | >0.9999 | |
| 5 vs. 13 | -15.67 | No | ns | >0.9999 | |
| 5 vs. 14 | 64.67 | No | ns | | 0.441 |
| 5 vs. 15 | 27.33 | No | ns | >0.9999 | |
| 5 vs. 16 | 47 | No | ns | >0.9999 | |
| 5 vs. 17 | 36.33 | No | ns | >0.9999 | |
| 5 vs. 18 | 17 | No | ns | >0.9999 | |
| 5 vs. 19 | 56.67 | No | ns | >0.9999 | |
| 5 vs. 20 | 25.67 | No | ns | >0.9999 | |
| 5 vs. 21 | 37.67 | No | ns | >0.9999 | |
| 5 vs. 22 | 22 | No | ns | >0.9999 | |
| 5 vs. 23 | 52 | No | ns | >0.9999 | |
| 5 vs. 24 | 29.67 | No | ns | >0.9999 | |
| 5 vs. 25 | 40.33 | No | ns | >0.9999 | |
| 5 vs. 26 | 29 | No | ns | >0.9999 | |
| 5 vs. 27 | 32.67 | No | ns | >0.9999 | |
| 5 vs. 28 | 34 | No | ns | >0.9999 | |

| | | | | | |
|---|---|---|---|---|---|
| 6 vs. 7 | 2.333 | No | ns | >0.9999 | |
| 6 vs. 8 | -0.3333 | No | ns | >0.9999 | |
| 6 vs. 9 | -1.333 | No | ns | >0.9999 | |
| 6 vs. 10 | -6 | No | ns | >0.9999 | |
| 6 vs. 11 | 48.33 | No | ns | >0.9999 | |
| 6 vs. 12 | 54.67 | No | ns | >0.9999 | |
| 6 vs. 13 | -18.67 | No | ns | >0.9999 | |
| 6 vs. 14 | 61.67 | No | ns | | 0.7408 |
| 6 vs. 15 | 24.33 | No | ns | >0.9999 | |
| 6 vs. 16 | 44 | No | ns | >0.9999 | |
| 6 vs. 17 | 33.33 | No | ns | >0.9999 | |
| 6 vs. 18 | 14 | No | ns | >0.9999 | |
| 6 vs. 19 | 53.67 | No | ns | >0.9999 | |
| 6 vs. 20 | 22.67 | No | ns | >0.9999 | |
| 6 vs. 21 | 34.67 | No | ns | >0.9999 | |
| 6 vs. 22 | 19 | No | ns | >0.9999 | |
| 6 vs. 23 | 49 | No | ns | >0.9999 | |
| 6 vs. 24 | 26.67 | No | ns | >0.9999 | |
| 6 vs. 25 | 37.33 | No | ns | >0.9999 | |
| 6 vs. 26 | 26 | No | ns | >0.9999 | |
| 6 vs. 27 | 29.67 | No | ns | >0.9999 | |
| 6 vs. 28 | 31 | No | ns | >0.9999 | |
| 7 vs. 8 | -2.667 | No | ns | >0.9999 | |
| 7 vs. 9 | -3.667 | No | ns | >0.9999 | |
| 7 vs. 10 | -8.333 | No | ns | >0.9999 | |
| 7 vs. 11 | 46 | No | ns | >0.9999 | |
| 7 vs. 12 | 52.33 | No | ns | >0.9999 | |
| 7 vs. 13 | -21 | No | ns | >0.9999 | |
| 7 vs. 14 | 59.33 | No | ns | >0.9999 | |
| 7 vs. 15 | 22 | No | ns | >0.9999 | |
| 7 vs. 16 | 41.67 | No | ns | >0.9999 | |
| 7 vs. 17 | 31 | No | ns | >0.9999 | |
| 7 vs. 18 | 11.67 | No | ns | >0.9999 | |
| 7 vs. 19 | 51.33 | No | ns | >0.9999 | |
| 7 vs. 20 | 20.33 | No | ns | >0.9999 | |
| 7 vs. 21 | 32.33 | No | ns | >0.9999 | |
| 7 vs. 22 | 16.67 | No | ns | >0.9999 | |
| 7 vs. 23 | 46.67 | No | ns | >0.9999 | |
| 7 vs. 24 | 24.33 | No | ns | >0.9999 | |
| 7 vs. 25 | 35 | No | ns | >0.9999 | |
| 7 vs. 26 | 23.67 | No | ns | >0.9999 | |
| 7 vs. 27 | 27.33 | No | ns | >0.9999 | |
| 7 vs. 28 | 28.67 | No | ns | >0.9999 | |
| 8 vs. 9 | -1 | No | ns | >0.9999 | |
| 8 vs. 10 | -5.667 | No | ns | >0.9999 | |
| 8 vs. 11 | 48.67 | No | ns | >0.9999 | |
| 8 vs. 12 | 55 | No | ns | >0.9999 | |

| | | | | | |
|---|---|---|---|---|---|
| 8 vs. 13 | -18.33 | No | ns | >0.9999 | |
| 8 vs. 14 | 62 | No | ns | | 0.7 |
| 8 vs. 15 | 24.67 | No | ns | >0.9999 | |
| 8 vs. 16 | 44.33 | No | ns | >0.9999 | |
| 8 vs. 17 | 33.67 | No | ns | >0.9999 | |
| 8 vs. 18 | 14.33 | No | ns | >0.9999 | |
| 8 vs. 19 | 54 | No | ns | >0.9999 | |
| 8 vs. 20 | 23 | No | ns | >0.9999 | |
| 8 vs. 21 | 35 | No | ns | >0.9999 | |
| 8 vs. 22 | 19.33 | No | ns | >0.9999 | |
| 8 vs. 23 | 49.33 | No | ns | >0.9999 | |
| 8 vs. 24 | 27 | No | ns | >0.9999 | |
| 8 vs. 25 | 37.67 | No | ns | >0.9999 | |
| 8 vs. 26 | 26.33 | No | ns | >0.9999 | |
| 8 vs. 27 | 30 | No | ns | >0.9999 | |
| 8 vs. 28 | 31.33 | No | ns | >0.9999 | |
| 9 vs. 10 | -4.667 | No | ns | >0.9999 | |
| 9 vs. 11 | 49.67 | No | ns | >0.9999 | |
| 9 vs. 12 | 56 | No | ns | >0.9999 | |
| 9 vs. 13 | -17.33 | No | ns | >0.9999 | |
| 9 vs. 14 | 63 | No | ns | | 0.5898 |
| 9 vs. 15 | 25.67 | No | ns | >0.9999 | |
| 9 vs. 16 | 45.33 | No | ns | >0.9999 | |
| 9 vs. 17 | 34.67 | No | ns | >0.9999 | |
| 9 vs. 18 | 15.33 | No | ns | >0.9999 | |
| 9 vs. 19 | 55 | No | ns | >0.9999 | |
| 9 vs. 20 | 24 | No | ns | >0.9999 | |
| 9 vs. 21 | 36 | No | ns | >0.9999 | |
| 9 vs. 22 | 20.33 | No | ns | >0.9999 | |
| 9 vs. 23 | 50.33 | No | ns | >0.9999 | |
| 9 vs. 24 | 28 | No | ns | >0.9999 | |
| 9 vs. 25 | 38.67 | No | ns | >0.9999 | |
| 9 vs. 26 | 27.33 | No | ns | >0.9999 | |
| 9 vs. 27 | 31 | No | ns | >0.9999 | |
| 9 vs. 28 | 32.33 | No | ns | >0.9999 | |
| 10 vs. 11 | 54.33 | No | ns | >0.9999 | |
| 10 vs. 12 | 60.67 | No | ns | | 0.8765 |
| 10 vs. 13 | -12.67 | No | ns | >0.9999 | |
| 10 vs. 14 | 67.67 | No | ns | | 0.257 |
| 10 vs. 15 | 30.33 | No | ns | >0.9999 | |
| 10 vs. 16 | 50 | No | ns | >0.9999 | |
| 10 vs. 17 | 39.33 | No | ns | >0.9999 | |
| 10 vs. 18 | 20 | No | ns | >0.9999 | |
| 10 vs. 19 | 59.67 | No | ns | >0.9999 | |
| 10 vs. 20 | 28.67 | No | ns | >0.9999 | |
| 10 vs. 21 | 40.67 | No | ns | >0.9999 | |
| 10 vs. 22 | 25 | No | ns | >0.9999 | |

| | | | | | |
|---|---|---|---|---|---|
| 10 vs. 23 | 55 | No | ns | >0.9999 | |
| 10 vs. 24 | 32.67 | No | ns | >0.9999 | |
| 10 vs. 25 | 43.33 | No | ns | >0.9999 | |
| 10 vs. 26 | 32 | No | ns | >0.9999 | |
| 10 vs. 27 | 35.67 | No | ns | >0.9999 | |
| 10 vs. 28 | 37 | No | ns | >0.9999 | |
| 11 vs. 12 | 6.333 | No | ns | >0.9999 | |
| 11 vs. 13 | -67 | No | ns | | 0.2903 |
| 11 vs. 14 | 13.33 | No | ns | >0.9999 | |
| 11 vs. 15 | -24 | No | ns | >0.9999 | |
| 11 vs. 16 | -4.333 | No | ns | >0.9999 | |
| 11 vs. 17 | -15 | No | ns | >0.9999 | |
| 11 vs. 18 | -34.33 | No | ns | >0.9999 | |
| 11 vs. 19 | 5.333 | No | ns | >0.9999 | |
| 11 vs. 20 | -25.67 | No | ns | >0.9999 | |
| 11 vs. 21 | -13.67 | No | ns | >0.9999 | |
| 11 vs. 22 | -29.33 | No | ns | >0.9999 | |
| 11 vs. 23 | 0.6667 | No | ns | >0.9999 | |
| 11 vs. 24 | -21.67 | No | ns | >0.9999 | |
| 11 vs. 25 | -11 | No | ns | >0.9999 | |
| 11 vs. 26 | -22.33 | No | ns | >0.9999 | |
| 11 vs. 27 | -18.67 | No | ns | >0.9999 | |
| 11 vs. 28 | -17.33 | No | ns | >0.9999 | |
| 12 vs. 13 | -73.33 | No | ns | | 0.0875 |
| 12 vs. 14 | 7 | No | ns | >0.9999 | |
| 12 vs. 15 | -30.33 | No | ns | >0.9999 | |
| 12 vs. 16 | -10.67 | No | ns | >0.9999 | |
| 12 vs. 17 | -21.33 | No | ns | >0.9999 | |
| 12 vs. 18 | -40.67 | No | ns | >0.9999 | |
| 12 vs. 19 | -1 | No | ns | >0.9999 | |
| 12 vs. 20 | -32 | No | ns | >0.9999 | |
| 12 vs. 21 | -20 | No | ns | >0.9999 | |
| 12 vs. 22 | -35.67 | No | ns | >0.9999 | |
| 12 vs. 23 | -5.667 | No | ns | >0.9999 | |
| 12 vs. 24 | -28 | No | ns | >0.9999 | |
| 12 vs. 25 | -17.33 | No | ns | >0.9999 | |
| 12 vs. 26 | -28.67 | No | ns | >0.9999 | |
| 12 vs. 27 | -25 | No | ns | >0.9999 | |
| 12 vs. 28 | -23.67 | No | ns | >0.9999 | |
| 13 vs. 14 | 80.33 | Yes | * | | 0.0208 |
| 13 vs. 15 | 43 | No | ns | >0.9999 | |
| 13 vs. 16 | 62.67 | No | ns | | 0.6247 |
| 13 vs. 17 | 52 | No | ns | >0.9999 | |
| 13 vs. 18 | 32.67 | No | ns | >0.9999 | |
| 13 vs. 19 | 72.33 | No | ns | | 0.1064 |
| 13 vs. 20 | 41.33 | No | ns | >0.9999 | |
| 13 vs. 21 | 53.33 | No | ns | >0.9999 | |

| | | | | | |
|---|---|---|---|---|---|
| 13 vs. 22 | 37.67 | No | ns | >0.9999 | |
| 13 vs. 23 | 67.67 | No | ns | | 0.257 |
| 13 vs. 24 | 45.33 | No | ns | >0.9999 | |
| 13 vs. 25 | 56 | No | ns | >0.9999 | |
| 13 vs. 26 | 44.67 | No | ns | >0.9999 | |
| 13 vs. 27 | 48.33 | No | ns | >0.9999 | |
| 13 vs. 28 | 49.67 | No | ns | >0.9999 | |
| 14 vs. 15 | -37.33 | No | ns | >0.9999 | |
| 14 vs. 16 | -17.67 | No | ns | >0.9999 | |
| 14 vs. 17 | -28.33 | No | ns | >0.9999 | |
| 14 vs. 18 | -47.67 | No | ns | >0.9999 | |
| 14 vs. 19 | -8 | No | ns | >0.9999 | |
| 14 vs. 20 | -39 | No | ns | >0.9999 | |
| 14 vs. 21 | -27 | No | ns | >0.9999 | |
| 14 vs. 22 | -42.67 | No | ns | >0.9999 | |
| 14 vs. 23 | -12.67 | No | ns | >0.9999 | |
| 14 vs. 24 | -35 | No | ns | >0.9999 | |
| 14 vs. 25 | -24.33 | No | ns | >0.9999 | |
| 14 vs. 26 | -35.67 | No | ns | >0.9999 | |
| 14 vs. 27 | -32 | No | ns | >0.9999 | |
| 14 vs. 28 | -30.67 | No | ns | >0.9999 | |
| 15 vs. 16 | 19.67 | No | ns | >0.9999 | |
| 15 vs. 17 | 9 | No | ns | >0.9999 | |
| 15 vs. 18 | -10.33 | No | ns | >0.9999 | |
| 15 vs. 19 | 29.33 | No | ns | >0.9999 | |
| 15 vs. 20 | -1.667 | No | ns | >0.9999 | |
| 15 vs. 21 | 10.33 | No | ns | >0.9999 | |
| 15 vs. 22 | -5.333 | No | ns | >0.9999 | |
| 15 vs. 23 | 24.67 | No | ns | >0.9999 | |
| 15 vs. 24 | 2.333 | No | ns | >0.9999 | |
| 15 vs. 25 | 13 | No | ns | >0.9999 | |
| 15 vs. 26 | 1.667 | No | ns | >0.9999 | |
| 15 vs. 27 | 5.333 | No | ns | >0.9999 | |
| 15 vs. 28 | 6.667 | No | ns | >0.9999 | |
| 16 vs. 17 | -10.67 | No | ns | >0.9999 | |
| 16 vs. 18 | -30 | No | ns | >0.9999 | |
| 16 vs. 19 | 9.667 | No | ns | >0.9999 | |
| 16 vs. 20 | -21.33 | No | ns | >0.9999 | |
| 16 vs. 21 | -9.333 | No | ns | >0.9999 | |
| 16 vs. 22 | -25 | No | ns | >0.9999 | |
| 16 vs. 23 | 5 | No | ns | >0.9999 | |
| 16 vs. 24 | -17.33 | No | ns | >0.9999 | |
| 16 vs. 25 | -6.667 | No | ns | >0.9999 | |
| 16 vs. 26 | -18 | No | ns | >0.9999 | |
| 16 vs. 27 | -14.33 | No | ns | >0.9999 | |
| 16 vs. 28 | -13 | No | ns | >0.9999 | |
| 17 vs. 18 | -19.33 | No | ns | >0.9999 | |

| | | | | |
|---|---|---|---|---|
| 17 vs. 19 | 20.33 | No | ns | >0.9999 |
| 17 vs. 20 | -10.67 | No | ns | >0.9999 |
| 17 vs. 21 | 1.333 | No | ns | >0.9999 |
| 17 vs. 22 | -14.33 | No | ns | >0.9999 |
| 17 vs. 23 | 15.67 | No | ns | >0.9999 |
| 17 vs. 24 | -6.667 | No | ns | >0.9999 |
| 17 vs. 25 | 4 | No | ns | >0.9999 |
| 17 vs. 26 | -7.333 | No | ns | >0.9999 |
| 17 vs. 27 | -3.667 | No | ns | >0.9999 |
| 17 vs. 28 | -2.333 | No | ns | >0.9999 |
| 18 vs. 19 | 39.67 | No | ns | >0.9999 |
| 18 vs. 20 | 8.667 | No | ns | >0.9999 |
| 18 vs. 21 | 20.67 | No | ns | >0.9999 |
| 18 vs. 22 | 5 | No | ns | >0.9999 |
| 18 vs. 23 | 35 | No | ns | >0.9999 |
| 18 vs. 24 | 12.67 | No | ns | >0.9999 |
| 18 vs. 25 | 23.33 | No | ns | >0.9999 |
| 18 vs. 26 | 12 | No | ns | >0.9999 |
| 18 vs. 27 | 15.67 | No | ns | >0.9999 |
| 18 vs. 28 | 17 | No | ns | >0.9999 |
| 19 vs. 20 | -31 | No | ns | >0.9999 |
| 19 vs. 21 | -19 | No | ns | >0.9999 |
| 19 vs. 22 | -34.67 | No | ns | >0.9999 |
| 19 vs. 23 | -4.667 | No | ns | >0.9999 |
| 19 vs. 24 | -27 | No | ns | >0.9999 |
| 19 vs. 25 | -16.33 | No | ns | >0.9999 |
| 19 vs. 26 | -27.67 | No | ns | >0.9999 |
| 19 vs. 27 | -24 | No | ns | >0.9999 |
| 19 vs. 28 | -22.67 | No | ns | >0.9999 |
| 20 vs. 21 | 12 | No | ns | >0.9999 |
| 20 vs. 22 | -3.667 | No | ns | >0.9999 |
| 20 vs. 23 | 26.33 | No | ns | >0.9999 |
| 20 vs. 24 | 4 | No | ns | >0.9999 |
| 20 vs. 25 | 14.67 | No | ns | >0.9999 |
| 20 vs. 26 | 3.333 | No | ns | >0.9999 |
| 20 vs. 27 | 7 | No | ns | >0.9999 |
| 20 vs. 28 | 8.333 | No | ns | >0.9999 |
| 21 vs. 22 | -15.67 | No | ns | >0.9999 |
| 21 vs. 23 | 14.33 | No | ns | >0.9999 |
| 21 vs. 24 | -8 | No | ns | >0.9999 |
| 21 vs. 25 | 2.667 | No | ns | >0.9999 |
| 21 vs. 26 | -8.667 | No | ns | >0.9999 |
| 21 vs. 27 | -5 | No | ns | >0.9999 |
| 21 vs. 28 | -3.667 | No | ns | >0.9999 |
| 22 vs. 23 | 30 | No | ns | >0.9999 |
| 22 vs. 24 | 7.667 | No | ns | >0.9999 |
| 22 vs. 25 | 18.33 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 22 vs. 26 | 7 | No | ns | >0.9999 |
| 22 vs. 27 | 10.67 | No | ns | >0.9999 |
| 22 vs. 28 | 12 | No | ns | >0.9999 |
| 23 vs. 24 | -22.33 | No | ns | >0.9999 |
| 23 vs. 25 | -11.67 | No | ns | >0.9999 |
| 23 vs. 26 | -23 | No | ns | >0.9999 |
| 23 vs. 27 | -19.33 | No | ns | >0.9999 |
| 23 vs. 28 | -18 | No | ns | >0.9999 |
| 24 vs. 25 | 10.67 | No | ns | >0.9999 |
| 24 vs. 26 | -0.6667 | No | ns | >0.9999 |
| 24 vs. 27 | 3 | No | ns | >0.9999 |
| 24 vs. 28 | 4.333 | No | ns | >0.9999 |
| 25 vs. 26 | -11.33 | No | ns | >0.9999 |
| 25 vs. 27 | -7.667 | No | ns | >0.9999 |
| 25 vs. 28 | -6.333 | No | ns | >0.9999 |
| 26 vs. 27 | 3.667 | No | ns | >0.9999 |
| 26 vs. 28 | 5 | No | ns | >0.9999 |
| 27 vs. 28 | 1.333 | No | ns | >0.9999 |

*N.nct KW & Dunn's*

| Kruskal-Wallis test | |
|---|---|
| P value | <0.0001 |
| Exact or approximate P value? | Approximate |
| P value summary | **** |
| Do the medians vary signif. (P < 0.05)? | Yes |
| Number of groups | 22 |
| Kruskal-Wallis statistic | 54.76 |
| | |
| Data summary | |
| Number of treatments (columns) | 22 |
| Number of values (total) | 66 |

| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value |
|---|---|---|---|---|
| 1 vs. 2 | 16 | No | ns | >0.9999 |
| 1 vs. 3 | 26.67 | No | ns | >0.9999 |
| 1 vs. 4 | 30.33 | No | ns | >0.9999 |
| 1 vs. 5 | 26 | No | ns | >0.9999 |
| 1 vs. 6 | 29.67 | No | ns | >0.9999 |
| 1 vs. 7 | 33.33 | No | ns | >0.9999 |
| 1 vs. 8 | -4 | No | ns | >0.9999 |
| 1 vs. 9 | -7 | No | ns | >0.9999 |
| 1 vs. 10 | -10.33 | No | ns | >0.9999 |
| 1 vs. 11 | 50.33 | No | ns | 0.3052 |
| 1 vs. 12 | 50 | No | ns | 0.3286 |
| 1 vs. 13 | 47.33 | No | ns | 0.5841 |

| | | | | |
|---|---|---|---|---|
| 1 vs. 14 | 22.33 | No | ns | >0.9999 |
| 1 vs. 15 | 9.667 | No | ns | >0.9999 |
| 1 vs. 16 | 6.667 | No | ns | >0.9999 |
| 1 vs. 17 | 16.67 | No | ns | >0.9999 |
| 1 vs. 18 | 9.667 | No | ns | >0.9999 |
| 1 vs. 19 | 39.33 | No | ns | >0.9999 |
| 1 vs. 20 | 20.33 | No | ns | >0.9999 |
| 1 vs. 21 | 15.67 | No | ns | >0.9999 |
| 1 vs. 22 | 37 | No | ns | >0.9999 |
| 2 vs. 3 | 10.67 | No | ns | >0.9999 |
| 2 vs. 4 | 14.33 | No | ns | >0.9999 |
| 2 vs. 5 | 10 | No | ns | >0.9999 |
| 2 vs. 6 | 13.67 | No | ns | >0.9999 |
| 2 vs. 7 | 17.33 | No | ns | >0.9999 |
| 2 vs. 8 | -20 | No | ns | >0.9999 |
| 2 vs. 9 | -23 | No | ns | >0.9999 |
| 2 vs. 10 | -26.33 | No | ns | >0.9999 |
| 2 vs. 11 | 34.33 | No | ns | >0.9999 |
| 2 vs. 12 | 34 | No | ns | >0.9999 |
| 2 vs. 13 | 31.33 | No | ns | >0.9999 |
| 2 vs. 14 | 6.333 | No | ns | >0.9999 |
| 2 vs. 15 | -6.333 | No | ns | >0.9999 |
| 2 vs. 16 | -9.333 | No | ns | >0.9999 |
| 2 vs. 17 | 0.6667 | No | ns | >0.9999 |
| 2 vs. 18 | -6.333 | No | ns | >0.9999 |
| 2 vs. 19 | 23.33 | No | ns | >0.9999 |
| 2 vs. 20 | 4.333 | No | ns | >0.9999 |
| 2 vs. 21 | -0.3333 | No | ns | >0.9999 |
| 2 vs. 22 | 21 | No | ns | >0.9999 |
| 3 vs. 4 | 3.667 | No | ns | >0.9999 |
| 3 vs. 5 | -0.6667 | No | ns | >0.9999 |
| 3 vs. 6 | 3 | No | ns | >0.9999 |
| 3 vs. 7 | 6.667 | No | ns | >0.9999 |
| 3 vs. 8 | -30.67 | No | ns | >0.9999 |
| 3 vs. 9 | -33.67 | No | ns | >0.9999 |
| 3 vs. 10 | -37 | No | ns | >0.9999 |
| 3 vs. 11 | 23.67 | No | ns | >0.9999 |
| 3 vs. 12 | 23.33 | No | ns | >0.9999 |
| 3 vs. 13 | 20.67 | No | ns | >0.9999 |
| 3 vs. 14 | -4.333 | No | ns | >0.9999 |
| 3 vs. 15 | -17 | No | ns | >0.9999 |
| 3 vs. 16 | -20 | No | ns | >0.9999 |
| 3 vs. 17 | -10 | No | ns | >0.9999 |
| 3 vs. 18 | -17 | No | ns | >0.9999 |
| 3 vs. 19 | 12.67 | No | ns | >0.9999 |
| 3 vs. 20 | -6.333 | No | ns | >0.9999 |
| 3 vs. 21 | -11 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 3 vs. 22 | 10.33 | No | ns | >0.9999 |
| 4 vs. 5 | -4.333 | No | ns | >0.9999 |
| 4 vs. 6 | -0.6667 | No | ns | >0.9999 |
| 4 vs. 7 | 3 | No | ns | >0.9999 |
| 4 vs. 8 | -34.33 | No | ns | >0.9999 |
| 4 vs. 9 | -37.33 | No | ns | >0.9999 |
| 4 vs. 10 | -40.67 | No | ns | >0.9999 |
| 4 vs. 11 | 20 | No | ns | >0.9999 |
| 4 vs. 12 | 19.67 | No | ns | >0.9999 |
| 4 vs. 13 | 17 | No | ns | >0.9999 |
| 4 vs. 14 | -8 | No | ns | >0.9999 |
| 4 vs. 15 | -20.67 | No | ns | >0.9999 |
| 4 vs. 16 | -23.67 | No | ns | >0.9999 |
| 4 vs. 17 | -13.67 | No | ns | >0.9999 |
| 4 vs. 18 | -20.67 | No | ns | >0.9999 |
| 4 vs. 19 | 9 | No | ns | >0.9999 |
| 4 vs. 20 | -10 | No | ns | >0.9999 |
| 4 vs. 21 | -14.67 | No | ns | >0.9999 |
| 4 vs. 22 | 6.667 | No | ns | >0.9999 |
| 5 vs. 6 | 3.667 | No | ns | >0.9999 |
| 5 vs. 7 | 7.333 | No | ns | >0.9999 |
| 5 vs. 8 | -30 | No | ns | >0.9999 |
| 5 vs. 9 | -33 | No | ns | >0.9999 |
| 5 vs. 10 | -36.33 | No | ns | >0.9999 |
| 5 vs. 11 | 24.33 | No | ns | >0.9999 |
| 5 vs. 12 | 24 | No | ns | >0.9999 |
| 5 vs. 13 | 21.33 | No | ns | >0.9999 |
| 5 vs. 14 | -3.667 | No | ns | >0.9999 |
| 5 vs. 15 | -16.33 | No | ns | >0.9999 |
| 5 vs. 16 | -19.33 | No | ns | >0.9999 |
| 5 vs. 17 | -9.333 | No | ns | >0.9999 |
| 5 vs. 18 | -16.33 | No | ns | >0.9999 |
| 5 vs. 19 | 13.33 | No | ns | >0.9999 |
| 5 vs. 20 | -5.667 | No | ns | >0.9999 |
| 5 vs. 21 | -10.33 | No | ns | >0.9999 |
| 5 vs. 22 | 11 | No | ns | >0.9999 |
| 6 vs. 7 | 3.667 | No | ns | >0.9999 |
| 6 vs. 8 | -33.67 | No | ns | >0.9999 |
| 6 vs. 9 | -36.67 | No | ns | >0.9999 |
| 6 vs. 10 | -40 | No | ns | >0.9999 |
| 6 vs. 11 | 20.67 | No | ns | >0.9999 |
| 6 vs. 12 | 20.33 | No | ns | >0.9999 |
| 6 vs. 13 | 17.67 | No | ns | >0.9999 |
| 6 vs. 14 | -7.333 | No | ns | >0.9999 |
| 6 vs. 15 | -20 | No | ns | >0.9999 |
| 6 vs. 16 | -23 | No | ns | >0.9999 |
| 6 vs. 17 | -13 | No | ns | >0.9999 |

| | | | | | |
|---|---|---|---|---|---|
| 6 vs. 18 | -20 | No | ns | >0.9999 | |
| 6 vs. 19 | 9.667 | No | ns | >0.9999 | |
| 6 vs. 20 | -9.333 | No | ns | >0.9999 | |
| 6 vs. 21 | -14 | No | ns | >0.9999 | |
| 6 vs. 22 | 7.333 | No | ns | >0.9999 | |
| 7 vs. 8 | -37.33 | No | ns | >0.9999 | |
| 7 vs. 9 | -40.33 | No | ns | >0.9999 | |
| 7 vs. 10 | -43.67 | No | ns | >0.9999 | |
| 7 vs. 11 | 17 | No | ns | >0.9999 | |
| 7 vs. 12 | 16.67 | No | ns | >0.9999 | |
| 7 vs. 13 | 14 | No | ns | >0.9999 | |
| 7 vs. 14 | -11 | No | ns | >0.9999 | |
| 7 vs. 15 | -23.67 | No | ns | >0.9999 | |
| 7 vs. 16 | -26.67 | No | ns | >0.9999 | |
| 7 vs. 17 | -16.67 | No | ns | >0.9999 | |
| 7 vs. 18 | -23.67 | No | ns | >0.9999 | |
| 7 vs. 19 | 6 | No | ns | >0.9999 | |
| 7 vs. 20 | -13 | No | ns | >0.9999 | |
| 7 vs. 21 | -17.67 | No | ns | >0.9999 | |
| 7 vs. 22 | 3.667 | No | ns | >0.9999 | |
| 8 vs. 9 | -3 | No | ns | >0.9999 | |
| 8 vs. 10 | -6.333 | No | ns | >0.9999 | |
| 8 vs. 11 | 54.33 | No | ns | | 0.1218 |
| 8 vs. 12 | 54 | No | ns | | 0.1318 |
| 8 vs. 13 | 51.33 | No | ns | | 0.244 |
| 8 vs. 14 | 26.33 | No | ns | >0.9999 | |
| 8 vs. 15 | 13.67 | No | ns | >0.9999 | |
| 8 vs. 16 | 10.67 | No | ns | >0.9999 | |
| 8 vs. 17 | 20.67 | No | ns | >0.9999 | |
| 8 vs. 18 | 13.67 | No | ns | >0.9999 | |
| 8 vs. 19 | 43.33 | No | ns | >0.9999 | |
| 8 vs. 20 | 24.33 | No | ns | >0.9999 | |
| 8 vs. 21 | 19.67 | No | ns | >0.9999 | |
| 8 vs. 22 | 41 | No | ns | >0.9999 | |
| 9 vs. 10 | -3.333 | No | ns | >0.9999 | |
| 9 vs. 11 | 57.33 | No | ns | | 0.0587 |
| 9 vs. 12 | 57 | No | ns | | 0.0638 |
| 9 vs. 13 | 54.33 | No | ns | | 0.1218 |
| 9 vs. 14 | 29.33 | No | ns | >0.9999 | |
| 9 vs. 15 | 16.67 | No | ns | >0.9999 | |
| 9 vs. 16 | 13.67 | No | ns | >0.9999 | |
| 9 vs. 17 | 23.67 | No | ns | >0.9999 | |
| 9 vs. 18 | 16.67 | No | ns | >0.9999 | |
| 9 vs. 19 | 46.33 | No | ns | | 0.7197 |
| 9 vs. 20 | 27.33 | No | ns | >0.9999 | |
| 9 vs. 21 | 22.67 | No | ns | >0.9999 | |
| 9 vs. 22 | 44 | No | ns | >0.9999 | |

| | | | | |
|---|---|---|---|---|
| 10 vs. 11 | 60.67 | Yes | * | 0.0251 |
| 10 vs. 12 | 60.33 | Yes | * | 0.0274 |
| 10 vs. 13 | 57.67 | No | ns | 0.054 |
| 10 vs. 14 | 32.67 | No | ns | >0.9999 |
| 10 vs. 15 | 20 | No | ns | >0.9999 |
| 10 vs. 16 | 17 | No | ns | >0.9999 |
| 10 vs. 17 | 27 | No | ns | >0.9999 |
| 10 vs. 18 | 20 | No | ns | >0.9999 |
| 10 vs. 19 | 49.67 | No | ns | 0.3536 |
| 10 vs. 20 | 30.67 | No | ns | >0.9999 |
| 10 vs. 21 | 26 | No | ns | >0.9999 |
| 10 vs. 22 | 47.33 | No | ns | 0.5841 |
| 11 vs. 12 | -0.3333 | No | ns | >0.9999 |
| 11 vs. 13 | -3 | No | ns | >0.9999 |
| 11 vs. 14 | -28 | No | ns | >0.9999 |
| 11 vs. 15 | -40.67 | No | ns | >0.9999 |
| 11 vs. 16 | -43.67 | No | ns | >0.9999 |
| 11 vs. 17 | -33.67 | No | ns | >0.9999 |
| 11 vs. 18 | -40.67 | No | ns | >0.9999 |
| 11 vs. 19 | -11 | No | ns | >0.9999 |
| 11 vs. 20 | -30 | No | ns | >0.9999 |
| 11 vs. 21 | -34.67 | No | ns | >0.9999 |
| 11 vs. 22 | -13.33 | No | ns | >0.9999 |
| 12 vs. 13 | -2.667 | No | ns | >0.9999 |
| 12 vs. 14 | -27.67 | No | ns | >0.9999 |
| 12 vs. 15 | -40.33 | No | ns | >0.9999 |
| 12 vs. 16 | -43.33 | No | ns | >0.9999 |
| 12 vs. 17 | -33.33 | No | ns | >0.9999 |
| 12 vs. 18 | -40.33 | No | ns | >0.9999 |
| 12 vs. 19 | -10.67 | No | ns | >0.9999 |
| 12 vs. 20 | -29.67 | No | ns | >0.9999 |
| 12 vs. 21 | -34.33 | No | ns | >0.9999 |
| 12 vs. 22 | -13 | No | ns | >0.9999 |
| 13 vs. 14 | -25 | No | ns | >0.9999 |
| 13 vs. 15 | -37.67 | No | ns | >0.9999 |
| 13 vs. 16 | -40.67 | No | ns | >0.9999 |
| 13 vs. 17 | -30.67 | No | ns | >0.9999 |
| 13 vs. 18 | -37.67 | No | ns | >0.9999 |
| 13 vs. 19 | -8 | No | ns | >0.9999 |
| 13 vs. 20 | -27 | No | ns | >0.9999 |
| 13 vs. 21 | -31.67 | No | ns | >0.9999 |
| 13 vs. 22 | -10.33 | No | ns | >0.9999 |
| 14 vs. 15 | -12.67 | No | ns | >0.9999 |
| 14 vs. 16 | -15.67 | No | ns | >0.9999 |
| 14 vs. 17 | -5.667 | No | ns | >0.9999 |
| 14 vs. 18 | -12.67 | No | ns | >0.9999 |
| 14 vs. 19 | 17 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 14 vs. 20 | -2 | No | ns | >0.9999 |
| 14 vs. 21 | -6.667 | No | ns | >0.9999 |
| 14 vs. 22 | 14.67 | No | ns | >0.9999 |
| 15 vs. 16 | -3 | No | ns | >0.9999 |
| 15 vs. 17 | 7 | No | ns | >0.9999 |
| 15 vs. 18 | 0 | No | ns | >0.9999 |
| 15 vs. 19 | 29.67 | No | ns | >0.9999 |
| 15 vs. 20 | 10.67 | No | ns | >0.9999 |
| 15 vs. 21 | 6 | No | ns | >0.9999 |
| 15 vs. 22 | 27.33 | No | ns | >0.9999 |
| 16 vs. 17 | 10 | No | ns | >0.9999 |
| 16 vs. 18 | 3 | No | ns | >0.9999 |
| 16 vs. 19 | 32.67 | No | ns | >0.9999 |
| 16 vs. 20 | 13.67 | No | ns | >0.9999 |
| 16 vs. 21 | 9 | No | ns | >0.9999 |
| 16 vs. 22 | 30.33 | No | ns | >0.9999 |
| 17 vs. 18 | -7 | No | ns | >0.9999 |
| 17 vs. 19 | 22.67 | No | ns | >0.9999 |
| 17 vs. 20 | 3.667 | No | ns | >0.9999 |
| 17 vs. 21 | -1 | No | ns | >0.9999 |
| 17 vs. 22 | 20.33 | No | ns | >0.9999 |
| 18 vs. 19 | 29.67 | No | ns | >0.9999 |
| 18 vs. 20 | 10.67 | No | ns | >0.9999 |
| 18 vs. 21 | 6 | No | ns | >0.9999 |
| 18 vs. 22 | 27.33 | No | ns | >0.9999 |
| 19 vs. 20 | -19 | No | ns | >0.9999 |
| 19 vs. 21 | -23.67 | No | ns | >0.9999 |
| 19 vs. 22 | -2.333 | No | ns | >0.9999 |
| 20 vs. 21 | -4.667 | No | ns | >0.9999 |
| 20 vs. 22 | 16.67 | No | ns | >0.9999 |
| 21 vs. 22 | 21.33 | No | ns | >0.9999 |

*N.nig KW & Dunn's*

| Kruskal-Wallis test | |
|---|---|
| P value | 0.0002 |
| Exact or approximate P value? | Approximate |
| P value summary | *** |
| Do the medians vary signif. (P < 0.05)? | Yes |
| Number of groups | 25 |
| Kruskal-Wallis statistic | 56.53 |
| | |
| Data summary | |
| Number of treatments (columns) | 25 |
| Number of values (total) | 72 |

| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value |
|---|---|---|---|---|
| 1 vs. 2 | 3.333 | No | ns | >0.9999 |
| 1 vs. 3 | 12.67 | No | ns | >0.9999 |
| 1 vs. 4 | 20 | No | ns | >0.9999 |
| 1 vs. 5 | 0.3333 | No | ns | >0.9999 |
| 1 vs. 6 | -5 | No | ns | >0.9999 |
| 1 vs. 7 | 14.33 | No | ns | >0.9999 |
| 1 vs. 8 | -9 | No | ns | >0.9999 |
| 1 vs. 9 | 3.333 | No | ns | >0.9999 |
| 1 vs. 10 | 8.667 | No | ns | >0.9999 |
| 1 vs. 11 | -20.67 | No | ns | >0.9999 |
| 1 vs. 12 | -19.67 | No | ns | >0.9999 |
| 1 vs. 13 | 44.33 | No | ns | >0.9999 |
| 1 vs. 14 | 43.67 | No | ns | >0.9999 |
| 1 vs. 15 | 45 | No | ns | >0.9999 |
| 1 vs. 16 | 12.67 | No | ns | >0.9999 |
| 1 vs. 17 | 37 | No | ns | >0.9999 |
| 1 vs. 18 | 11.67 | No | ns | >0.9999 |
| 1 vs. 19 | 19.67 | No | ns | >0.9999 |
| 1 vs. 20 | 36 | No | ns | >0.9999 |
| 1 vs. 21 | 10.33 | No | ns | >0.9999 |
| 1 vs. 22 | 3.667 | No | ns | >0.9999 |
| 1 vs. 23 | 27.33 | No | ns | >0.9999 |
| 1 vs. 24 | 28.83 | No | ns | >0.9999 |
| 1 vs. 25 | -2.667 | No | ns | >0.9999 |
| 2 vs. 3 | 9.333 | No | ns | >0.9999 |
| 2 vs. 4 | 16.67 | No | ns | >0.9999 |
| 2 vs. 5 | -3 | No | ns | >0.9999 |
| 2 vs. 6 | -8.333 | No | ns | >0.9999 |
| 2 vs. 7 | 11 | No | ns | >0.9999 |
| 2 vs. 8 | -12.33 | No | ns | >0.9999 |
| 2 vs. 9 | 0 | No | ns | >0.9999 |
| 2 vs. 10 | 5.333 | No | ns | >0.9999 |
| 2 vs. 11 | -24 | No | ns | >0.9999 |
| 2 vs. 12 | -23 | No | ns | >0.9999 |
| 2 vs. 13 | 41 | No | ns | >0.9999 |
| 2 vs. 14 | 40.33 | No | ns | >0.9999 |
| 2 vs. 15 | 41.67 | No | ns | >0.9999 |
| 2 vs. 16 | 9.333 | No | ns | >0.9999 |
| 2 vs. 17 | 33.67 | No | ns | >0.9999 |
| 2 vs. 18 | 8.333 | No | ns | >0.9999 |
| 2 vs. 19 | 16.33 | No | ns | >0.9999 |
| 2 vs. 20 | 32.67 | No | ns | >0.9999 |
| 2 vs. 21 | 7 | No | ns | >0.9999 |
| 2 vs. 22 | 0.3333 | No | ns | >0.9999 |
| 2 vs. 23 | 24 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 2 vs. 24 | 25.5 | No | ns | >0.9999 |
| 2 vs. 25 | -6 | No | ns | >0.9999 |
| 3 vs. 4 | 7.333 | No | ns | >0.9999 |
| 3 vs. 5 | -12.33 | No | ns | >0.9999 |
| 3 vs. 6 | -17.67 | No | ns | >0.9999 |
| 3 vs. 7 | 1.667 | No | ns | >0.9999 |
| 3 vs. 8 | -21.67 | No | ns | >0.9999 |
| 3 vs. 9 | -9.333 | No | ns | >0.9999 |
| 3 vs. 10 | -4 | No | ns | >0.9999 |
| 3 vs. 11 | -33.33 | No | ns | >0.9999 |
| 3 vs. 12 | -32.33 | No | ns | >0.9999 |
| 3 vs. 13 | 31.67 | No | ns | >0.9999 |
| 3 vs. 14 | 31 | No | ns | >0.9999 |
| 3 vs. 15 | 32.33 | No | ns | >0.9999 |
| 3 vs. 16 | 0 | No | ns | >0.9999 |
| 3 vs. 17 | 24.33 | No | ns | >0.9999 |
| 3 vs. 18 | -1 | No | ns | >0.9999 |
| 3 vs. 19 | 7 | No | ns | >0.9999 |
| 3 vs. 20 | 23.33 | No | ns | >0.9999 |
| 3 vs. 21 | -2.333 | No | ns | >0.9999 |
| 3 vs. 22 | -9 | No | ns | >0.9999 |
| 3 vs. 23 | 14.67 | No | ns | >0.9999 |
| 3 vs. 24 | 16.17 | No | ns | >0.9999 |
| 3 vs. 25 | -15.33 | No | ns | >0.9999 |
| 4 vs. 5 | -19.67 | No | ns | >0.9999 |
| 4 vs. 6 | -25 | No | ns | >0.9999 |
| 4 vs. 7 | -5.667 | No | ns | >0.9999 |
| 4 vs. 8 | -29 | No | ns | >0.9999 |
| 4 vs. 9 | -16.67 | No | ns | >0.9999 |
| 4 vs. 10 | -11.33 | No | ns | >0.9999 |
| 4 vs. 11 | -40.67 | No | ns | >0.9999 |
| 4 vs. 12 | -39.67 | No | ns | >0.9999 |
| 4 vs. 13 | 24.33 | No | ns | >0.9999 |
| 4 vs. 14 | 23.67 | No | ns | >0.9999 |
| 4 vs. 15 | 25 | No | ns | >0.9999 |
| 4 vs. 16 | -7.333 | No | ns | >0.9999 |
| 4 vs. 17 | 17 | No | ns | >0.9999 |
| 4 vs. 18 | -8.333 | No | ns | >0.9999 |
| 4 vs. 19 | -0.3333 | No | ns | >0.9999 |
| 4 vs. 20 | 16 | No | ns | >0.9999 |
| 4 vs. 21 | -9.667 | No | ns | >0.9999 |
| 4 vs. 22 | -16.33 | No | ns | >0.9999 |
| 4 vs. 23 | 7.333 | No | ns | >0.9999 |
| 4 vs. 24 | 8.833 | No | ns | >0.9999 |
| 4 vs. 25 | -22.67 | No | ns | >0.9999 |
| 5 vs. 6 | -5.333 | No | ns | >0.9999 |
| 5 vs. 7 | 14 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 5 vs. 8 | -9.333 | No | ns | >0.9999 |
| 5 vs. 9 | 3 | No | ns | >0.9999 |
| 5 vs. 10 | 8.333 | No | ns | >0.9999 |
| 5 vs. 11 | -21 | No | ns | >0.9999 |
| 5 vs. 12 | -20 | No | ns | >0.9999 |
| 5 vs. 13 | 44 | No | ns | >0.9999 |
| 5 vs. 14 | 43.33 | No | ns | >0.9999 |
| 5 vs. 15 | 44.67 | No | ns | >0.9999 |
| 5 vs. 16 | 12.33 | No | ns | >0.9999 |
| 5 vs. 17 | 36.67 | No | ns | >0.9999 |
| 5 vs. 18 | 11.33 | No | ns | >0.9999 |
| 5 vs. 19 | 19.33 | No | ns | >0.9999 |
| 5 vs. 20 | 35.67 | No | ns | >0.9999 |
| 5 vs. 21 | 10 | No | ns | >0.9999 |
| 5 vs. 22 | 3.333 | No | ns | >0.9999 |
| 5 vs. 23 | 27 | No | ns | >0.9999 |
| 5 vs. 24 | 28.5 | No | ns | >0.9999 |
| 5 vs. 25 | -3 | No | ns | >0.9999 |
| 6 vs. 7 | 19.33 | No | ns | >0.9999 |
| 6 vs. 8 | -4 | No | ns | >0.9999 |
| 6 vs. 9 | 8.333 | No | ns | >0.9999 |
| 6 vs. 10 | 13.67 | No | ns | >0.9999 |
| 6 vs. 11 | -15.67 | No | ns | >0.9999 |
| 6 vs. 12 | -14.67 | No | ns | >0.9999 |
| 6 vs. 13 | 49.33 | No | ns | >0.9999 |
| 6 vs. 14 | 48.67 | No | ns | >0.9999 |
| 6 vs. 15 | 50 | No | ns | >0.9999 |
| 6 vs. 16 | 17.67 | No | ns | >0.9999 |
| 6 vs. 17 | 42 | No | ns | >0.9999 |
| 6 vs. 18 | 16.67 | No | ns | >0.9999 |
| 6 vs. 19 | 24.67 | No | ns | >0.9999 |
| 6 vs. 20 | 41 | No | ns | >0.9999 |
| 6 vs. 21 | 15.33 | No | ns | >0.9999 |
| 6 vs. 22 | 8.667 | No | ns | >0.9999 |
| 6 vs. 23 | 32.33 | No | ns | >0.9999 |
| 6 vs. 24 | 33.83 | No | ns | >0.9999 |
| 6 vs. 25 | 2.333 | No | ns | >0.9999 |
| 7 vs. 8 | -23.33 | No | ns | >0.9999 |
| 7 vs. 9 | -11 | No | ns | >0.9999 |
| 7 vs. 10 | -5.667 | No | ns | >0.9999 |
| 7 vs. 11 | -35 | No | ns | >0.9999 |
| 7 vs. 12 | -34 | No | ns | >0.9999 |
| 7 vs. 13 | 30 | No | ns | >0.9999 |
| 7 vs. 14 | 29.33 | No | ns | >0.9999 |
| 7 vs. 15 | 30.67 | No | ns | >0.9999 |
| 7 vs. 16 | -1.667 | No | ns | >0.9999 |
| 7 vs. 17 | 22.67 | No | ns | >0.9999 |

| | | | | | |
|---|---|---|---|---|---|
| 7 vs. 18 | -2.667 | No | ns | >0.9999 | |
| 7 vs. 19 | 5.333 | No | ns | >0.9999 | |
| 7 vs. 20 | 21.67 | No | ns | >0.9999 | |
| 7 vs. 21 | -4 | No | ns | >0.9999 | |
| 7 vs. 22 | -10.67 | No | ns | >0.9999 | |
| 7 vs. 23 | 13 | No | ns | >0.9999 | |
| 7 vs. 24 | 14.5 | No | ns | >0.9999 | |
| 7 vs. 25 | -17 | No | ns | >0.9999 | |
| 8 vs. 9 | 12.33 | No | ns | >0.9999 | |
| 8 vs. 10 | 17.67 | No | ns | >0.9999 | |
| 8 vs. 11 | -11.67 | No | ns | >0.9999 | |
| 8 vs. 12 | -10.67 | No | ns | >0.9999 | |
| 8 vs. 13 | 53.33 | No | ns | | 0.5405 |
| 8 vs. 14 | 52.67 | No | ns | | 0.6167 |
| 8 vs. 15 | 54 | No | ns | | 0.4731 |
| 8 vs. 16 | 21.67 | No | ns | >0.9999 | |
| 8 vs. 17 | 46 | No | ns | >0.9999 | |
| 8 vs. 18 | 20.67 | No | ns | >0.9999 | |
| 8 vs. 19 | 28.67 | No | ns | >0.9999 | |
| 8 vs. 20 | 45 | No | ns | >0.9999 | |
| 8 vs. 21 | 19.33 | No | ns | >0.9999 | |
| 8 vs. 22 | 12.67 | No | ns | >0.9999 | |
| 8 vs. 23 | 36.33 | No | ns | >0.9999 | |
| 8 vs. 24 | 37.83 | No | ns | >0.9999 | |
| 8 vs. 25 | 6.333 | No | ns | >0.9999 | |
| 9 vs. 10 | 5.333 | No | ns | >0.9999 | |
| 9 vs. 11 | -24 | No | ns | >0.9999 | |
| 9 vs. 12 | -23 | No | ns | >0.9999 | |
| 9 vs. 13 | 41 | No | ns | >0.9999 | |
| 9 vs. 14 | 40.33 | No | ns | >0.9999 | |
| 9 vs. 15 | 41.67 | No | ns | >0.9999 | |
| 9 vs. 16 | 9.333 | No | ns | >0.9999 | |
| 9 vs. 17 | 33.67 | No | ns | >0.9999 | |
| 9 vs. 18 | 8.333 | No | ns | >0.9999 | |
| 9 vs. 19 | 16.33 | No | ns | >0.9999 | |
| 9 vs. 20 | 32.67 | No | ns | >0.9999 | |
| 9 vs. 21 | 7 | No | ns | >0.9999 | |
| 9 vs. 22 | 0.3333 | No | ns | >0.9999 | |
| 9 vs. 23 | 24 | No | ns | >0.9999 | |
| 9 vs. 24 | 25.5 | No | ns | >0.9999 | |
| 9 vs. 25 | -6 | No | ns | >0.9999 | |
| 10 vs. 11 | -29.33 | No | ns | >0.9999 | |
| 10 vs. 12 | -28.33 | No | ns | >0.9999 | |
| 10 vs. 13 | 35.67 | No | ns | >0.9999 | |
| 10 vs. 14 | 35 | No | ns | >0.9999 | |
| 10 vs. 15 | 36.33 | No | ns | >0.9999 | |
| 10 vs. 16 | 4 | No | ns | >0.9999 | |

| | | | | |
|---|---|---|---|---|
| 10 vs. 17 | 28.33 | No | ns | >0.9999 |
| 10 vs. 18 | 3 | No | ns | >0.9999 |
| 10 vs. 19 | 11 | No | ns | >0.9999 |
| 10 vs. 20 | 27.33 | No | ns | >0.9999 |
| 10 vs. 21 | 1.667 | No | ns | >0.9999 |
| 10 vs. 22 | -5 | No | ns | >0.9999 |
| 10 vs. 23 | 18.67 | No | ns | >0.9999 |
| 10 vs. 24 | 20.17 | No | ns | >0.9999 |
| 10 vs. 25 | -11.33 | No | ns | >0.9999 |
| 11 vs. 12 | 1 | No | ns | >0.9999 |
| 11 vs. 13 | 65 | Yes | * | 0.0427 |
| 11 vs. 14 | 64.33 | No | ns | 0.05 |
| 11 vs. 15 | 65.67 | Yes | * | 0.0365 |
| 11 vs. 16 | 33.33 | No | ns | >0.9999 |
| 11 vs. 17 | 57.67 | No | ns | 0.2217 |
| 11 vs. 18 | 32.33 | No | ns | >0.9999 |
| 11 vs. 19 | 40.33 | No | ns | >0.9999 |
| 11 vs. 20 | 56.67 | No | ns | 0.2738 |
| 11 vs. 21 | 31 | No | ns | >0.9999 |
| 11 vs. 22 | 24.33 | No | ns | >0.9999 |
| 11 vs. 23 | 48 | No | ns | >0.9999 |
| 11 vs. 24 | 49.5 | No | ns | >0.9999 |
| 11 vs. 25 | 18 | No | ns | >0.9999 |
| 12 vs. 13 | 64 | No | ns | 0.054 |
| 12 vs. 14 | 63.33 | No | ns | 0.0631 |
| 12 vs. 15 | 64.67 | Yes | * | 0.0462 |
| 12 vs. 16 | 32.33 | No | ns | >0.9999 |
| 12 vs. 17 | 56.67 | No | ns | 0.2738 |
| 12 vs. 18 | 31.33 | No | ns | >0.9999 |
| 12 vs. 19 | 39.33 | No | ns | >0.9999 |
| 12 vs. 20 | 55.67 | No | ns | 0.337 |
| 12 vs. 21 | 30 | No | ns | >0.9999 |
| 12 vs. 22 | 23.33 | No | ns | >0.9999 |
| 12 vs. 23 | 47 | No | ns | >0.9999 |
| 12 vs. 24 | 48.5 | No | ns | >0.9999 |
| 12 vs. 25 | 17 | No | ns | >0.9999 |
| 13 vs. 14 | -0.6667 | No | ns | >0.9999 |
| 13 vs. 15 | 0.6667 | No | ns | >0.9999 |
| 13 vs. 16 | -31.67 | No | ns | >0.9999 |
| 13 vs. 17 | -7.333 | No | ns | >0.9999 |
| 13 vs. 18 | -32.67 | No | ns | >0.9999 |
| 13 vs. 19 | -24.67 | No | ns | >0.9999 |
| 13 vs. 20 | -8.333 | No | ns | >0.9999 |
| 13 vs. 21 | -34 | No | ns | >0.9999 |
| 13 vs. 22 | -40.67 | No | ns | >0.9999 |
| 13 vs. 23 | -17 | No | ns | >0.9999 |
| 13 vs. 24 | -15.5 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 13 vs. 25 | -47 | No | ns | >0.9999 |
| 14 vs. 15 | 1.333 | No | ns | >0.9999 |
| 14 vs. 16 | -31 | No | ns | >0.9999 |
| 14 vs. 17 | -6.667 | No | ns | >0.9999 |
| 14 vs. 18 | -32 | No | ns | >0.9999 |
| 14 vs. 19 | -24 | No | ns | >0.9999 |
| 14 vs. 20 | -7.667 | No | ns | >0.9999 |
| 14 vs. 21 | -33.33 | No | ns | >0.9999 |
| 14 vs. 22 | -40 | No | ns | >0.9999 |
| 14 vs. 23 | -16.33 | No | ns | >0.9999 |
| 14 vs. 24 | -14.83 | No | ns | >0.9999 |
| 14 vs. 25 | -46.33 | No | ns | >0.9999 |
| 15 vs. 16 | -32.33 | No | ns | >0.9999 |
| 15 vs. 17 | -8 | No | ns | >0.9999 |
| 15 vs. 18 | -33.33 | No | ns | >0.9999 |
| 15 vs. 19 | -25.33 | No | ns | >0.9999 |
| 15 vs. 20 | -9 | No | ns | >0.9999 |
| 15 vs. 21 | -34.67 | No | ns | >0.9999 |
| 15 vs. 22 | -41.33 | No | ns | >0.9999 |
| 15 vs. 23 | -17.67 | No | ns | >0.9999 |
| 15 vs. 24 | -16.17 | No | ns | >0.9999 |
| 15 vs. 25 | -47.67 | No | ns | >0.9999 |
| 16 vs. 17 | 24.33 | No | ns | >0.9999 |
| 16 vs. 18 | -1 | No | ns | >0.9999 |
| 16 vs. 19 | 7 | No | ns | >0.9999 |
| 16 vs. 20 | 23.33 | No | ns | >0.9999 |
| 16 vs. 21 | -2.333 | No | ns | >0.9999 |
| 16 vs. 22 | -9 | No | ns | >0.9999 |
| 16 vs. 23 | 14.67 | No | ns | >0.9999 |
| 16 vs. 24 | 16.17 | No | ns | >0.9999 |
| 16 vs. 25 | -15.33 | No | ns | >0.9999 |
| 17 vs. 18 | -25.33 | No | ns | >0.9999 |
| 17 vs. 19 | -17.33 | No | ns | >0.9999 |
| 17 vs. 20 | -1 | No | ns | >0.9999 |
| 17 vs. 21 | -26.67 | No | ns | >0.9999 |
| 17 vs. 22 | -33.33 | No | ns | >0.9999 |
| 17 vs. 23 | -9.667 | No | ns | >0.9999 |
| 17 vs. 24 | -8.167 | No | ns | >0.9999 |
| 17 vs. 25 | -39.67 | No | ns | >0.9999 |
| 18 vs. 19 | 8 | No | ns | >0.9999 |
| 18 vs. 20 | 24.33 | No | ns | >0.9999 |
| 18 vs. 21 | -1.333 | No | ns | >0.9999 |
| 18 vs. 22 | -8 | No | ns | >0.9999 |
| 18 vs. 23 | 15.67 | No | ns | >0.9999 |
| 18 vs. 24 | 17.17 | No | ns | >0.9999 |
| 18 vs. 25 | -14.33 | No | ns | >0.9999 |
| 19 vs. 20 | 16.33 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 19 vs. 21 | -9.333 | No | ns | >0.9999 |
| 19 vs. 22 | -16 | No | ns | >0.9999 |
| 19 vs. 23 | 7.667 | No | ns | >0.9999 |
| 19 vs. 24 | 9.167 | No | ns | >0.9999 |
| 19 vs. 25 | -22.33 | No | ns | >0.9999 |
| 20 vs. 21 | -25.67 | No | ns | >0.9999 |
| 20 vs. 22 | -32.33 | No | ns | >0.9999 |
| 20 vs. 23 | -8.667 | No | ns | >0.9999 |
| 20 vs. 24 | -7.167 | No | ns | >0.9999 |
| 20 vs. 25 | -38.67 | No | ns | >0.9999 |
| 21 vs. 22 | -6.667 | No | ns | >0.9999 |
| 21 vs. 23 | 17 | No | ns | >0.9999 |
| 21 vs. 24 | 18.5 | No | ns | >0.9999 |
| 21 vs. 25 | -13 | No | ns | >0.9999 |
| 22 vs. 23 | 23.67 | No | ns | >0.9999 |
| 22 vs. 24 | 25.17 | No | ns | >0.9999 |
| 22 vs. 25 | -6.333 | No | ns | >0.9999 |
| 23 vs. 24 | 1.5 | No | ns | >0.9999 |
| 23 vs. 25 | -30 | No | ns | >0.9999 |
| 24 vs. 25 | -31.5 | No | ns | >0.9999 |

| Kruskal-Wallis test | |
|---|---|
| P value | 0.0009 |
| Exact or approximate P value? | Approximate |
| P value summary | *** |
| Do the medians vary signif. (P < 0.05)? | Yes |
| Number of groups | 21 |
| Kruskal-Wallis statistic | 45.81 |
| | |
| Data summary | |
| Number of treatments (columns) | 21 |
| Number of values (total) | 63 |

| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value |
|---|---|---|---|---|
| 1 vs. 2 | 12.33 | No | ns | >0.9999 |
| 1 vs. 3 | 13.33 | No | ns | >0.9999 |
| 1 vs. 4 | 20 | No | ns | >0.9999 |
| 1 vs. 5 | 10 | No | ns | >0.9999 |
| 1 vs. 6 | 16 | No | ns | >0.9999 |
| 1 vs. 7 | -9.833 | No | ns | >0.9999 |
| 1 vs. 8 | -9.5 | No | ns | >0.9999 |
| 1 vs. 9 | 1.5 | No | ns | >0.9999 |
| 1 vs. 10 | -15.5 | No | ns | >0.9999 |
| 1 vs. 11 | -0.8333 | No | ns | >0.9999 |
| 1 vs. 12 | 20.5 | No | ns | >0.9999 |
| 1 vs. 13 | 14.83 | No | ns | >0.9999 |
| 1 vs. 14 | 15.17 | No | ns | >0.9999 |
| 1 vs. 15 | 31.83 | No | ns | >0.9999 |
| 1 vs. 16 | 31.5 | No | ns | >0.9999 |
| 1 vs. 17 | 25.83 | No | ns | >0.9999 |
| 1 vs. 18 | 16.5 | No | ns | >0.9999 |
| 1 vs. 19 | 28.5 | No | ns | >0.9999 |
| 1 vs. 20 | 43.17 | No | ns | 0.8209 |
| 1 vs. 21 | 39.17 | No | ns | >0.9999 |
| 2 vs. 3 | 1 | No | ns | >0.9999 |
| 2 vs. 4 | 7.667 | No | ns | >0.9999 |
| 2 vs. 5 | -2.333 | No | ns | >0.9999 |
| 2 vs. 6 | 3.667 | No | ns | >0.9999 |
| 2 vs. 7 | -22.17 | No | ns | >0.9999 |
| 2 vs. 8 | -21.83 | No | ns | >0.9999 |
| 2 vs. 9 | -10.83 | No | ns | >0.9999 |
| 2 vs. 10 | -27.83 | No | ns | >0.9999 |
| 2 vs. 11 | -13.17 | No | ns | >0.9999 |
| 2 vs. 12 | 8.167 | No | ns | >0.9999 |
| 2 vs. 13 | 2.5 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 2 vs. 14 | 2.833 | No | ns | >0.9999 |
| 2 vs. 15 | 19.5 | No | ns | >0.9999 |
| 2 vs. 16 | 19.17 | No | ns | >0.9999 |
| 2 vs. 17 | 13.5 | No | ns | >0.9999 |
| 2 vs. 18 | 4.167 | No | ns | >0.9999 |
| 2 vs. 19 | 16.17 | No | ns | >0.9999 |
| 2 vs. 20 | 30.83 | No | ns | >0.9999 |
| 2 vs. 21 | 26.83 | No | ns | >0.9999 |
| 3 vs. 4 | 6.667 | No | ns | >0.9999 |
| 3 vs. 5 | -3.333 | No | ns | >0.9999 |
| 3 vs. 6 | 2.667 | No | ns | >0.9999 |
| 3 vs. 7 | -23.17 | No | ns | >0.9999 |
| 3 vs. 8 | -22.83 | No | ns | >0.9999 |
| 3 vs. 9 | -11.83 | No | ns | >0.9999 |
| 3 vs. 10 | -28.83 | No | ns | >0.9999 |
| 3 vs. 11 | -14.17 | No | ns | >0.9999 |
| 3 vs. 12 | 7.167 | No | ns | >0.9999 |
| 3 vs. 13 | 1.5 | No | ns | >0.9999 |
| 3 vs. 14 | 1.833 | No | ns | >0.9999 |
| 3 vs. 15 | 18.5 | No | ns | >0.9999 |
| 3 vs. 16 | 18.17 | No | ns | >0.9999 |
| 3 vs. 17 | 12.5 | No | ns | >0.9999 |
| 3 vs. 18 | 3.167 | No | ns | >0.9999 |
| 3 vs. 19 | 15.17 | No | ns | >0.9999 |
| 3 vs. 20 | 29.83 | No | ns | >0.9999 |
| 3 vs. 21 | 25.83 | No | ns | >0.9999 |
| 4 vs. 5 | -10 | No | ns | >0.9999 |
| 4 vs. 6 | -4 | No | ns | >0.9999 |
| 4 vs. 7 | -29.83 | No | ns | >0.9999 |
| 4 vs. 8 | -29.5 | No | ns | >0.9999 |
| 4 vs. 9 | -18.5 | No | ns | >0.9999 |
| 4 vs. 10 | -35.5 | No | ns | >0.9999 |
| 4 vs. 11 | -20.83 | No | ns | >0.9999 |
| 4 vs. 12 | 0.5 | No | ns | >0.9999 |
| 4 vs. 13 | -5.167 | No | ns | >0.9999 |
| 4 vs. 14 | -4.833 | No | ns | >0.9999 |
| 4 vs. 15 | 11.83 | No | ns | >0.9999 |
| 4 vs. 16 | 11.5 | No | ns | >0.9999 |
| 4 vs. 17 | 5.833 | No | ns | >0.9999 |
| 4 vs. 18 | -3.5 | No | ns | >0.9999 |
| 4 vs. 19 | 8.5 | No | ns | >0.9999 |
| 4 vs. 20 | 23.17 | No | ns | >0.9999 |
| 4 vs. 21 | 19.17 | No | ns | >0.9999 |
| 5 vs. 6 | 6 | No | ns | >0.9999 |
| 5 vs. 7 | -19.83 | No | ns | >0.9999 |
| 5 vs. 8 | -19.5 | No | ns | >0.9999 |
| 5 vs. 9 | -8.5 | No | ns | >0.9999 |

| | | | | |
|---|---|---|---|---|
| 5 vs. 10 | -25.5 | No | ns | >0.9999 |
| 5 vs. 11 | -10.83 | No | ns | >0.9999 |
| 5 vs. 12 | 10.5 | No | ns | >0.9999 |
| 5 vs. 13 | 4.833 | No | ns | >0.9999 |
| 5 vs. 14 | 5.167 | No | ns | >0.9999 |
| 5 vs. 15 | 21.83 | No | ns | >0.9999 |
| 5 vs. 16 | 21.5 | No | ns | >0.9999 |
| 5 vs. 17 | 15.83 | No | ns | >0.9999 |
| 5 vs. 18 | 6.5 | No | ns | >0.9999 |
| 5 vs. 19 | 18.5 | No | ns | >0.9999 |
| 5 vs. 20 | 33.17 | No | ns | >0.9999 |
| 5 vs. 21 | 29.17 | No | ns | >0.9999 |
| 6 vs. 7 | -25.83 | No | ns | >0.9999 |
| 6 vs. 8 | -25.5 | No | ns | >0.9999 |
| 6 vs. 9 | -14.5 | No | ns | >0.9999 |
| 6 vs. 10 | -31.5 | No | ns | >0.9999 |
| 6 vs. 11 | -16.83 | No | ns | >0.9999 |
| 6 vs. 12 | 4.5 | No | ns | >0.9999 |
| 6 vs. 13 | -1.167 | No | ns | >0.9999 |
| 6 vs. 14 | -0.8333 | No | ns | >0.9999 |
| 6 vs. 15 | 15.83 | No | ns | >0.9999 |
| 6 vs. 16 | 15.5 | No | ns | >0.9999 |
| 6 vs. 17 | 9.833 | No | ns | >0.9999 |
| 6 vs. 18 | 0.5 | No | ns | >0.9999 |
| 6 vs. 19 | 12.5 | No | ns | >0.9999 |
| 6 vs. 20 | 27.17 | No | ns | >0.9999 |
| 6 vs. 21 | 23.17 | No | ns | >0.9999 |
| 7 vs. 8 | 0.3333 | No | ns | >0.9999 |
| 7 vs. 9 | 11.33 | No | ns | >0.9999 |
| 7 vs. 10 | -5.667 | No | ns | >0.9999 |
| 7 vs. 11 | 9 | No | ns | >0.9999 |
| 7 vs. 12 | 30.33 | No | ns | >0.9999 |
| 7 vs. 13 | 24.67 | No | ns | >0.9999 |
| 7 vs. 14 | 25 | No | ns | >0.9999 |
| 7 vs. 15 | 41.67 | No | ns | >0.9999 |
| 7 vs. 16 | 41.33 | No | ns | >0.9999 |
| 7 vs. 17 | 35.67 | No | ns | >0.9999 |
| 7 vs. 18 | 26.33 | No | ns | >0.9999 |
| 7 vs. 19 | 38.33 | No | ns | >0.9999 |
| 7 vs. 20 | 53 | No | ns | 0.0832 |
| 7 vs. 21 | 49 | No | ns | 0.2216 |
| 8 vs. 9 | 11 | No | ns | >0.9999 |
| 8 vs. 10 | -6 | No | ns | >0.9999 |
| 8 vs. 11 | 8.667 | No | ns | >0.9999 |
| 8 vs. 12 | 30 | No | ns | >0.9999 |
| 8 vs. 13 | 24.33 | No | ns | >0.9999 |
| 8 vs. 14 | 24.67 | No | ns | >0.9999 |

| | | | | | |
|---|---|---|---|---|---|
| 8 vs. 15 | 41.33 | No | ns | >0.9999 | |
| 8 vs. 16 | 41 | No | ns | >0.9999 | |
| 8 vs. 17 | 35.33 | No | ns | >0.9999 | |
| 8 vs. 18 | 26 | No | ns | >0.9999 | |
| 8 vs. 19 | 38 | No | ns | >0.9999 | |
| 8 vs. 20 | 52.67 | No | ns | | 0.0905 |
| 8 vs. 21 | 48.67 | No | ns | | 0.2398 |
| 9 vs. 10 | -17 | No | ns | >0.9999 | |
| 9 vs. 11 | -2.333 | No | ns | >0.9999 | |
| 9 vs. 12 | 19 | No | ns | >0.9999 | |
| 9 vs. 13 | 13.33 | No | ns | >0.9999 | |
| 9 vs. 14 | 13.67 | No | ns | >0.9999 | |
| 9 vs. 15 | 30.33 | No | ns | >0.9999 | |
| 9 vs. 16 | 30 | No | ns | >0.9999 | |
| 9 vs. 17 | 24.33 | No | ns | >0.9999 | |
| 9 vs. 18 | 15 | No | ns | >0.9999 | |
| 9 vs. 19 | 27 | No | ns | >0.9999 | |
| 9 vs. 20 | 41.67 | No | ns | >0.9999 | |
| 9 vs. 21 | 37.67 | No | ns | >0.9999 | |
| 10 vs. 11 | 14.67 | No | ns | >0.9999 | |
| 10 vs. 12 | 36 | No | ns | >0.9999 | |
| 10 vs. 13 | 30.33 | No | ns | >0.9999 | |
| 10 vs. 14 | 30.67 | No | ns | >0.9999 | |
| 10 vs. 15 | 47.33 | No | ns | | 0.3269 |
| 10 vs. 16 | 47 | No | ns | | 0.3528 |
| 10 vs. 17 | 41.33 | No | ns | >0.9999 | |
| 10 vs. 18 | 32 | No | ns | >0.9999 | |
| 10 vs. 19 | 44 | No | ns | | 0.6868 |
| 10 vs. 20 | 58.67 | Yes | * | | 0.0185 |
| 10 vs. 21 | 54.67 | No | ns | | 0.0542 |
| 11 vs. 12 | 21.33 | No | ns | >0.9999 | |
| 11 vs. 13 | 15.67 | No | ns | >0.9999 | |
| 11 vs. 14 | 16 | No | ns | >0.9999 | |
| 11 vs. 15 | 32.67 | No | ns | >0.9999 | |
| 11 vs. 16 | 32.33 | No | ns | >0.9999 | |
| 11 vs. 17 | 26.67 | No | ns | >0.9999 | |
| 11 vs. 18 | 17.33 | No | ns | >0.9999 | |
| 11 vs. 19 | 29.33 | No | ns | >0.9999 | |
| 11 vs. 20 | 44 | No | ns | | 0.6868 |
| 11 vs. 21 | 40 | No | ns | >0.9999 | |
| 12 vs. 13 | -5.667 | No | ns | >0.9999 | |
| 12 vs. 14 | -5.333 | No | ns | >0.9999 | |
| 12 vs. 15 | 11.33 | No | ns | >0.9999 | |
| 12 vs. 16 | 11 | No | ns | >0.9999 | |
| 12 vs. 17 | 5.333 | No | ns | >0.9999 | |
| 12 vs. 18 | -4 | No | ns | >0.9999 | |
| 12 vs. 19 | 8 | No | ns | >0.9999 | |

| | | | | |
|---|---|---|---|---|
| 12 vs. 20 | 22.67 | No | ns | >0.9999 |
| 12 vs. 21 | 18.67 | No | ns | >0.9999 |
| 13 vs. 14 | 0.3333 | No | ns | >0.9999 |
| 13 vs. 15 | 17 | No | ns | >0.9999 |
| 13 vs. 16 | 16.67 | No | ns | >0.9999 |
| 13 vs. 17 | 11 | No | ns | >0.9999 |
| 13 vs. 18 | 1.667 | No | ns | >0.9999 |
| 13 vs. 19 | 13.67 | No | ns | >0.9999 |
| 13 vs. 20 | 28.33 | No | ns | >0.9999 |
| 13 vs. 21 | 24.33 | No | ns | >0.9999 |
| 14 vs. 15 | 16.67 | No | ns | >0.9999 |
| 14 vs. 16 | 16.33 | No | ns | >0.9999 |
| 14 vs. 17 | 10.67 | No | ns | >0.9999 |
| 14 vs. 18 | 1.333 | No | ns | >0.9999 |
| 14 vs. 19 | 13.33 | No | ns | >0.9999 |
| 14 vs. 20 | 28 | No | ns | >0.9999 |
| 14 vs. 21 | 24 | No | ns | >0.9999 |
| 15 vs. 16 | -0.3333 | No | ns | >0.9999 |
| 15 vs. 17 | -6 | No | ns | >0.9999 |
| 15 vs. 18 | -15.33 | No | ns | >0.9999 |
| 15 vs. 19 | -3.333 | No | ns | >0.9999 |
| 15 vs. 20 | 11.33 | No | ns | >0.9999 |
| 15 vs. 21 | 7.333 | No | ns | >0.9999 |
| 16 vs. 17 | -5.667 | No | ns | >0.9999 |
| 16 vs. 18 | -15 | No | ns | >0.9999 |
| 16 vs. 19 | -3 | No | ns | >0.9999 |
| 16 vs. 20 | 11.67 | No | ns | >0.9999 |
| 16 vs. 21 | 7.667 | No | ns | >0.9999 |
| 17 vs. 18 | -9.333 | No | ns | >0.9999 |
| 17 vs. 19 | 2.667 | No | ns | >0.9999 |
| 17 vs. 20 | 17.33 | No | ns | >0.9999 |
| 17 vs. 21 | 13.33 | No | ns | >0.9999 |
| 18 vs. 19 | 12 | No | ns | >0.9999 |
| 18 vs. 20 | 26.67 | No | ns | >0.9999 |
| 18 vs. 21 | 22.67 | No | ns | >0.9999 |
| 19 vs. 20 | 14.67 | No | ns | >0.9999 |
| 19 vs. 21 | 10.67 | No | ns | >0.9999 |
| 20 vs. 21 | -4 | No | ns | >0.9999 |

| Kruskal-Wallis test | |
|---|---|
| P value | 0.0005 |
| Exact or approximate P value? | Approximate |
| P value summary | *** |
| Do the medians vary signif. (P < 0.05)? | Yes |
| Number of groups | 20 |
| Kruskal-Wallis statistic | 45.87 |
| | |
| Data summary | |
| Number of treatments (columns) | 20 |
| Number of values (total) | 60 |

| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value |
|---|---|---|---|---|
| 1 vs. 2 | 27.33 | No | ns | >0.9999 |
| 1 vs. 3 | 20 | No | ns | >0.9999 |
| 1 vs. 4 | -9.667 | No | ns | >0.9999 |
| 1 vs. 5 | -10.67 | No | ns | >0.9999 |
| 1 vs. 6 | -12 | No | ns | >0.9999 |
| 1 vs. 7 | 31.67 | No | ns | >0.9999 |
| 1 vs. 8 | -13 | No | ns | >0.9999 |
| 1 vs. 9 | 35 | No | ns | >0.9999 |
| 1 vs. 10 | -21 | No | ns | >0.9999 |
| 1 vs. 11 | 22.33 | No | ns | >0.9999 |
| 1 vs. 12 | 2 | No | ns | >0.9999 |
| 1 vs. 13 | 19.67 | No | ns | >0.9999 |
| 1 vs. 14 | 11 | No | ns | >0.9999 |
| 1 vs. 15 | 10.67 | No | ns | >0.9999 |
| 1 vs. 16 | 11 | No | ns | >0.9999 |
| 1 vs. 17 | 7.333 | No | ns | >0.9999 |
| 1 vs. 18 | 1 | No | ns | >0.9999 |
| 1 vs. 19 | 9.667 | No | ns | >0.9999 |
| 1 vs. 20 | 7.667 | No | ns | >0.9999 |
| 2 vs. 3 | -7.333 | No | ns | >0.9999 |
| 2 vs. 4 | -37 | No | ns | >0.9999 |
| 2 vs. 5 | -38 | No | ns | >0.9999 |
| 2 vs. 6 | -39.33 | No | ns | >0.9999 |
| 2 vs. 7 | 4.333 | No | ns | >0.9999 |
| 2 vs. 8 | -40.33 | No | ns | 0.8885 |
| 2 vs. 9 | 7.667 | No | ns | >0.9999 |
| 2 vs. 10 | -48.33 | No | ns | 0.133 |
| 2 vs. 11 | -5 | No | ns | >0.9999 |
| 2 vs. 12 | -25.33 | No | ns | >0.9999 |
| 2 vs. 13 | -7.667 | No | ns | >0.9999 |

| | | | | | |
|---|---|---|---|---|---|
| 2 vs. 14 | -16.33 | No | ns | >0.9999 | |
| 2 vs. 15 | -16.67 | No | ns | >0.9999 | |
| 2 vs. 16 | -16.33 | No | ns | >0.9999 | |
| 2 vs. 17 | -20 | No | ns | >0.9999 | |
| 2 vs. 18 | -26.33 | No | ns | >0.9999 | |
| 2 vs. 19 | -17.67 | No | ns | >0.9999 | |
| 2 vs. 20 | -19.67 | No | ns | >0.9999 | |
| 3 vs. 4 | -29.67 | No | ns | >0.9999 | |
| 3 vs. 5 | -30.67 | No | ns | >0.9999 | |
| 3 vs. 6 | -32 | No | ns | >0.9999 | |
| 3 vs. 7 | 11.67 | No | ns | >0.9999 | |
| 3 vs. 8 | -33 | No | ns | >0.9999 | |
| 3 vs. 9 | 15 | No | ns | >0.9999 | |
| 3 vs. 10 | -41 | No | ns | | 0.767 |
| 3 vs. 11 | 2.333 | No | ns | >0.9999 | |
| 3 vs. 12 | -18 | No | ns | >0.9999 | |
| 3 vs. 13 | -0.3333 | No | ns | >0.9999 | |
| 3 vs. 14 | -9 | No | ns | >0.9999 | |
| 3 vs. 15 | -9.333 | No | ns | >0.9999 | |
| 3 vs. 16 | -9 | No | ns | >0.9999 | |
| 3 vs. 17 | -12.67 | No | ns | >0.9999 | |
| 3 vs. 18 | -19 | No | ns | >0.9999 | |
| 3 vs. 19 | -10.33 | No | ns | >0.9999 | |
| 3 vs. 20 | -12.33 | No | ns | >0.9999 | |
| 4 vs. 5 | -1 | No | ns | >0.9999 | |
| 4 vs. 6 | -2.333 | No | ns | >0.9999 | |
| 4 vs. 7 | 41.33 | No | ns | | 0.7121 |
| 4 vs. 8 | -3.333 | No | ns | >0.9999 | |
| 4 vs. 9 | 44.67 | No | ns | | 0.3294 |
| 4 vs. 10 | -11.33 | No | ns | >0.9999 | |
| 4 vs. 11 | 32 | No | ns | >0.9999 | |
| 4 vs. 12 | 11.67 | No | ns | >0.9999 | |
| 4 vs. 13 | 29.33 | No | ns | >0.9999 | |
| 4 vs. 14 | 20.67 | No | ns | >0.9999 | |
| 4 vs. 15 | 20.33 | No | ns | >0.9999 | |
| 4 vs. 16 | 20.67 | No | ns | >0.9999 | |
| 4 vs. 17 | 17 | No | ns | >0.9999 | |
| 4 vs. 18 | 10.67 | No | ns | >0.9999 | |
| 4 vs. 19 | 19.33 | No | ns | >0.9999 | |
| 4 vs. 20 | 17.33 | No | ns | >0.9999 | |
| 5 vs. 6 | -1.333 | No | ns | >0.9999 | |
| 5 vs. 7 | 42.33 | No | ns | | 0.5681 |
| 5 vs. 8 | -2.333 | No | ns | >0.9999 | |
| 5 vs. 9 | 45.67 | No | ns | | 0.2588 |
| 5 vs. 10 | -10.33 | No | ns | >0.9999 | |
| 5 vs. 11 | 33 | No | ns | >0.9999 | |
| 5 vs. 12 | 12.67 | No | ns | >0.9999 | |

| | | | | | |
|---|---|---|---|---|---|
| 5 vs. 13 | 30.33 | No | ns | >0.9999 | |
| 5 vs. 14 | 21.67 | No | ns | >0.9999 | |
| 5 vs. 15 | 21.33 | No | ns | >0.9999 | |
| 5 vs. 16 | 21.67 | No | ns | >0.9999 | |
| 5 vs. 17 | 18 | No | ns | >0.9999 | |
| 5 vs. 18 | 11.67 | No | ns | >0.9999 | |
| 5 vs. 19 | 20.33 | No | ns | >0.9999 | |
| 5 vs. 20 | 18.33 | No | ns | >0.9999 | |
| 6 vs. 7 | 43.67 | No | ns | | 0.4173 |
| 6 vs. 8 | -1 | No | ns | >0.9999 | |
| 6 vs. 9 | 47 | No | ns | | 0.1863 |
| 6 vs. 10 | -9 | No | ns | >0.9999 | |
| 6 vs. 11 | 34.33 | No | ns | >0.9999 | |
| 6 vs. 12 | 14 | No | ns | >0.9999 | |
| 6 vs. 13 | 31.67 | No | ns | >0.9999 | |
| 6 vs. 14 | 23 | No | ns | >0.9999 | |
| 6 vs. 15 | 22.67 | No | ns | >0.9999 | |
| 6 vs. 16 | 23 | No | ns | >0.9999 | |
| 6 vs. 17 | 19.33 | No | ns | >0.9999 | |
| 6 vs. 18 | 13 | No | ns | >0.9999 | |
| 6 vs. 19 | 21.67 | No | ns | >0.9999 | |
| 6 vs. 20 | 19.67 | No | ns | >0.9999 | |
| 7 vs. 8 | -44.67 | No | ns | | 0.3294 |
| 7 vs. 9 | 3.333 | No | ns | >0.9999 | |
| 7 vs. 10 | -52.67 | Yes | * | | 0.042 |
| 7 vs. 11 | -9.333 | No | ns | >0.9999 | |
| 7 vs. 12 | -29.67 | No | ns | >0.9999 | |
| 7 vs. 13 | -12 | No | ns | >0.9999 | |
| 7 vs. 14 | -20.67 | No | ns | >0.9999 | |
| 7 vs. 15 | -21 | No | ns | >0.9999 | |
| 7 vs. 16 | -20.67 | No | ns | >0.9999 | |
| 7 vs. 17 | -24.33 | No | ns | >0.9999 | |
| 7 vs. 18 | -30.67 | No | ns | >0.9999 | |
| 7 vs. 19 | -22 | No | ns | >0.9999 | |
| 7 vs. 20 | -24 | No | ns | >0.9999 | |
| 8 vs. 9 | 48 | No | ns | | 0.1448 |
| 8 vs. 10 | -8 | No | ns | >0.9999 | |
| 8 vs. 11 | 35.33 | No | ns | >0.9999 | |
| 8 vs. 12 | 15 | No | ns | >0.9999 | |
| 8 vs. 13 | 32.67 | No | ns | >0.9999 | |
| 8 vs. 14 | 24 | No | ns | >0.9999 | |
| 8 vs. 15 | 23.67 | No | ns | >0.9999 | |
| 8 vs. 16 | 24 | No | ns | >0.9999 | |
| 8 vs. 17 | 20.33 | No | ns | >0.9999 | |
| 8 vs. 18 | 14 | No | ns | >0.9999 | |
| 8 vs. 19 | 22.67 | No | ns | >0.9999 | |
| 8 vs. 20 | 20.67 | No | ns | >0.9999 | |

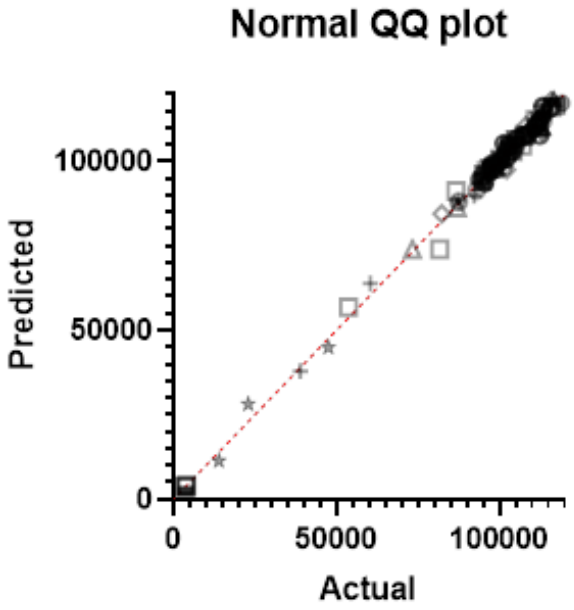
| | | | | | |
|---|---|---|---|---|---|
| 9 vs. 10 | -56 | Yes | * | | 0.0163 |
| 9 vs. 11 | -12.67 | No | ns | >0.9999 | |
| 9 vs. 12 | -33 | No | ns | >0.9999 | |
| 9 vs. 13 | -15.33 | No | ns | >0.9999 | |
| 9 vs. 14 | -24 | No | ns | >0.9999 | |
| 9 vs. 15 | -24.33 | No | ns | >0.9999 | |
| 9 vs. 16 | -24 | No | ns | >0.9999 | |
| 9 vs. 17 | -27.67 | No | ns | >0.9999 | |
| 9 vs. 18 | -34 | No | ns | >0.9999 | |
| 9 vs. 19 | -25.33 | No | ns | >0.9999 | |
| 9 vs. 20 | -27.33 | No | ns | >0.9999 | |
| 10 vs. 11 | 43.33 | No | ns | | 0.4511 |
| 10 vs. 12 | 23 | No | ns | >0.9999 | |
| 10 vs. 13 | 40.67 | No | ns | | 0.8257 |
| 10 vs. 14 | 32 | No | ns | >0.9999 | |
| 10 vs. 15 | 31.67 | No | ns | >0.9999 | |
| 10 vs. 16 | 32 | No | ns | >0.9999 | |
| 10 vs. 17 | 28.33 | No | ns | >0.9999 | |
| 10 vs. 18 | 22 | No | ns | >0.9999 | |
| 10 vs. 19 | 30.67 | No | ns | >0.9999 | |
| 10 vs. 20 | 28.67 | No | ns | >0.9999 | |
| 11 vs. 12 | -20.33 | No | ns | >0.9999 | |
| 11 vs. 13 | -2.667 | No | ns | >0.9999 | |
| 11 vs. 14 | -11.33 | No | ns | >0.9999 | |
| 11 vs. 15 | -11.67 | No | ns | >0.9999 | |
| 11 vs. 16 | -11.33 | No | ns | >0.9999 | |
| 11 vs. 17 | -15 | No | ns | >0.9999 | |
| 11 vs. 18 | -21.33 | No | ns | >0.9999 | |
| 11 vs. 19 | -12.67 | No | ns | >0.9999 | |
| 11 vs. 20 | -14.67 | No | ns | >0.9999 | |
| 12 vs. 13 | 17.67 | No | ns | >0.9999 | |
| 12 vs. 14 | 9 | No | ns | >0.9999 | |
| 12 vs. 15 | 8.667 | No | ns | >0.9999 | |
| 12 vs. 16 | 9 | No | ns | >0.9999 | |
| 12 vs. 17 | 5.333 | No | ns | >0.9999 | |
| 12 vs. 18 | -1 | No | ns | >0.9999 | |
| 12 vs. 19 | 7.667 | No | ns | >0.9999 | |
| 12 vs. 20 | 5.667 | No | ns | >0.9999 | |
| 13 vs. 14 | -8.667 | No | ns | >0.9999 | |
| 13 vs. 15 | -9 | No | ns | >0.9999 | |
| 13 vs. 16 | -8.667 | No | ns | >0.9999 | |
| 13 vs. 17 | -12.33 | No | ns | >0.9999 | |
| 13 vs. 18 | -18.67 | No | ns | >0.9999 | |
| 13 vs. 19 | -10 | No | ns | >0.9999 | |
| 13 vs. 20 | -12 | No | ns | >0.9999 | |
| 14 vs. 15 | -0.3333 | No | ns | >0.9999 | |
| 14 vs. 16 | 0 | No | ns | >0.9999 | |

| | | | | |
|---|---|---|---|---|
| 14 vs. 17 | -3.667 | No | ns | >0.9999 |
| 14 vs. 18 | -10 | No | ns | >0.9999 |
| 14 vs. 19 | -1.333 | No | ns | >0.9999 |
| 14 vs. 20 | -3.333 | No | ns | >0.9999 |
| 15 vs. 16 | 0.3333 | No | ns | >0.9999 |
| 15 vs. 17 | -3.333 | No | ns | >0.9999 |
| 15 vs. 18 | -9.667 | No | ns | >0.9999 |
| 15 vs. 19 | -1 | No | ns | >0.9999 |
| 15 vs. 20 | -3 | No | ns | >0.9999 |
| 16 vs. 17 | -3.667 | No | ns | >0.9999 |
| 16 vs. 18 | -10 | No | ns | >0.9999 |
| 16 vs. 19 | -1.333 | No | ns | >0.9999 |
| 16 vs. 20 | -3.333 | No | ns | >0.9999 |
| 17 vs. 18 | -6.333 | No | ns | >0.9999 |
| 17 vs. 19 | 2.333 | No | ns | >0.9999 |
| 17 vs. 20 | 0.3333 | No | ns | >0.9999 |
| 18 vs. 19 | 8.667 | No | ns | >0.9999 |
| 18 vs. 20 | 6.667 | No | ns | >0.9999 |
| 19 vs. 20 | -2 | No | ns | >0.9999 |

## Second Dimension SEC HPLC

*Normality testing*

| Sample | P value |
|---|---|
| N.nub_r10_s1 | 0.4249 |
| N.nub_r10_s2 | 0.6908 |
| N.nub_r10_s3 | 0.3513 |
| N.nub_r10_s4 | 0.4593 |
| N.nub_r10_s1 | 0.2665 |
| N.nub_r10_s2 | 0.2769 |
| N.nct_r10_s1 | 0.2672 |
| N.nct_r10_s2 | 0.9199 |
| N.nct_r10_s3 | 0.2859 |
| N.nct_r10_s4 | 0.4177 |
| N.nct_r10_s5 | 0.8103 |
| N.nct_r10_s8 | 0.4985 |
| N.nct_r8_s1 | 0.7847 |
| N.nct_r8_s2 | 0.5629 |
| N.nig_r11_s1 | 0.0264 |
| N.nig_r11_s2 | 0.48 |
| N.nig_r11_s3 | 0.8943 |
| N.nig_r11_s4 | 0.2304 |
| N.nig_r11_s5 | 0.6309 |
| N.nig_r12_s1 | 0.5024 |
| N.nig_r12_s2 | 0.8159 |
| N.nig_r12_s3 | 0.6034 |
| N.nct_r9_s1 | 0.5407 |
| N.pal_r10_s1 | 0.1781 |
| N.pal_r10_s2 | 0.0916 |
| N.pal_r10_s3 | 0.3917 |
| N.mos_r13_s1 | 0.789 |
| N.mos_r13_s2 | 0.1242 |
| N.mos_r13_s3 | 0.9315 |
| Pos con | 0.2413 |
| Neg con | 0.1135 |



*Kruskal-Wallis and Dunn's multiple comparisons test on 2nd dimension SEC fraction screen data*

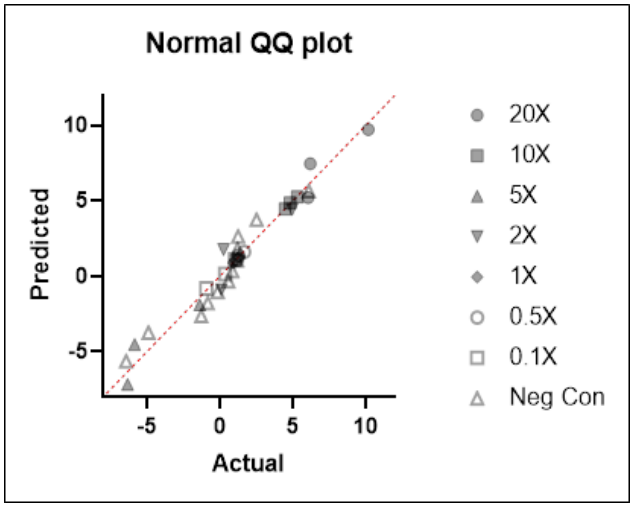
| Kruskal-Wallis test | | |
|---|---|---|
| P value | | 0.0004 |
| Exact or approximate P value? | Approximate | |
| P value summary | *** | |
| Do the medians vary signif. (P < 0.05)? | Yes | |
| Number of groups | | 31 |
| Kruskal-Wallis statistic | | 63.16 |
| | | |
| Data summary | | |
| Number of treatments (columns) | | 31 |
| Number of values (total) | | 93 |

| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value |
|---|---|---|---|---|
| Neg con vs. N.nub_r10_s1 | -28 | No | ns | >0.9999 |
| Neg con vs. N.nub_r10_s2 | -29.33 | No | ns | >0.9999 |
| Neg con vs. N.nub_r10_s3 | -45.67 | No | ns | >0.9999 |
| Neg con vs. N.nub_r10_s4 | -39.33 | No | ns | >0.9999 |
| Neg con vs. N.nub_r10_s1 | -38.33 | No | ns | >0.9999 |
| Neg con vs. N.nub_r10_s2 | 10.33 | No | ns | >0.9999 |
| Neg con vs. N.nct_r10_s1 | 17.67 | No | ns | >0.9999 |
| Neg con vs. N.nct_r10_s2 | 6 | No | ns | >0.9999 |
| Neg con vs. N.nct_r10_s3 | -7 | No | ns | >0.9999 |
| Neg con vs. N.nct_r10_s4 | -17.67 | No | ns | >0.9999 |
| Neg con vs. N.nct_r10_s5 | -8 | No | ns | >0.9999 |
| Neg con vs. N.nct_r10_s8 | 23 | No | ns | >0.9999 |
| Neg con vs. N.nct_r8_s1 | 17.67 | No | ns | >0.9999 |
| Neg con vs. N.nct_r8_s2 | -10.33 | No | ns | >0.9999 |
| Neg con vs. N.nig_r11_s1 | -35 | No | ns | >0.9999 |
| Neg con vs. N.nig_r11_s2 | -27.67 | No | ns | >0.9999 |
| Neg con vs. N.nig_r11_s3 | -12.33 | No | ns | >0.9999 |
| Neg con vs. N.nig_r11_s4 | -29 | No | ns | >0.9999 |
| Neg con vs. N.nig_r11_s5 | -27 | No | ns | >0.9999 |
| Neg con vs. N.nig_r12_s1 | -20 | No | ns | >0.9999 |
| Neg con vs. N.nig_r12_s2 | -4 | No | ns | >0.9999 |
| Neg con vs. N.nig_r12_s3 | -12 | No | ns | >0.9999 |
| Neg con vs. N.nct_r9_s1 | -24.67 | No | ns | >0.9999 |
| Neg con vs. N.pal_r10_s1 | -54 | No | ns | 0.4282 |
| Neg con vs. N.pal_r10_s2 | -37.67 | No | ns | >0.9999 |
| Neg con vs. N.pal_r10_s3 | -39 | No | ns | >0.9999 |
| Neg con vs. N.mos_r13_s1 | -49.67 | No | ns | 0.7265 |
| Neg con vs. N.mos_r13_s2 | -33.67 | No | ns | >0.9999 |
| Neg con vs. N.mos_r13_s3 | -50.33 | No | ns | 0.6712 |

## Salt Buffer Assay Statistics

### *Normality testing*

| Shapiro-Wilk test | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| W | 0.7772 | 0.9875 | 0.8221 | 0.7732 | 0.9304 | 0.8823 | 0.9657 | 0.9224 |
| P value | 0.0612 | 0.7857 | 0.1685 | 0.0519 | 0.4902 | 0.3312 | 0.6441 | 0.3067 |
| Passed normality test (alpha=0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P value summary | ns | ns | ns | ns | ns | ns | ns | ns |




### *Welch's ANOVA with Dunnett's post-hoc test*

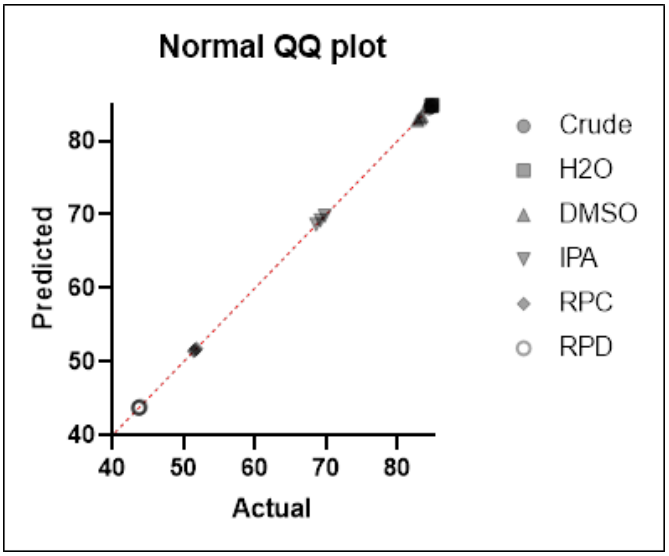
| Welch's ANOVA test | |
|---|---|
| W (DFn, DFd) | 23.71 (7.000, 7.514) |
| P value | 0.0001 |
| P value summary | *** |
| Significant diff. among means (P < 0.05)? | Yes |

| Dunnett's T3 multiple comparisons test | Mean Diff. | 95.00% CI of diff. | Below threshold? | Summary | Adjusted P Value | H-? |
|---|---|---|---|---|---|---|
| Neg Con vs. 20X | -7.467 | -16.55 to 1.614 | Yes | * | 0.0489 | A |
| Neg Con vs. 10X | -4.863 | -8.318 to -1.409 | Yes | ** | 0.002 | B |
| Neg Con vs. 5X | 4.533 | -5.568 to 14.63 | No | ns | 0.279 | C |
| Neg Con vs. 2X | -1.757 | -11.95 to 8.437 | No | ns | 0.9171 | D |
| Neg Con vs. 1X | -1.08 | -4.518 to 2.358 | No | ns | 0.8564 | E |
| Neg Con vs. 0.5X | -1.363 | -4.755 to 2.028 | No | ns | 0.6895 | F |
| Neg Con vs. 0.1X | -0.1467 | -4.073 to 3.780 | No | ns | >0.9999 | G |

## NanoTemper Statistics

### *Normality testing*

| Shapiro-Wilk test | | | | | | |
|---|---|---|---|---|---|---|
| W | 0.9231 | 0.75 | 0.9643 | 1 | 1 | 0.75 |
| P value | 0.4633 | | 0.6369 | >0.9999 | >0.9999 | |
| Passed normality test (alpha=0.05)? | Yes | | Yes | Yes | Yes | |
| P value summary | ns | | ns | ns | ns | |




### *Welch's ANOVA with Dunnett's post-hoc test*

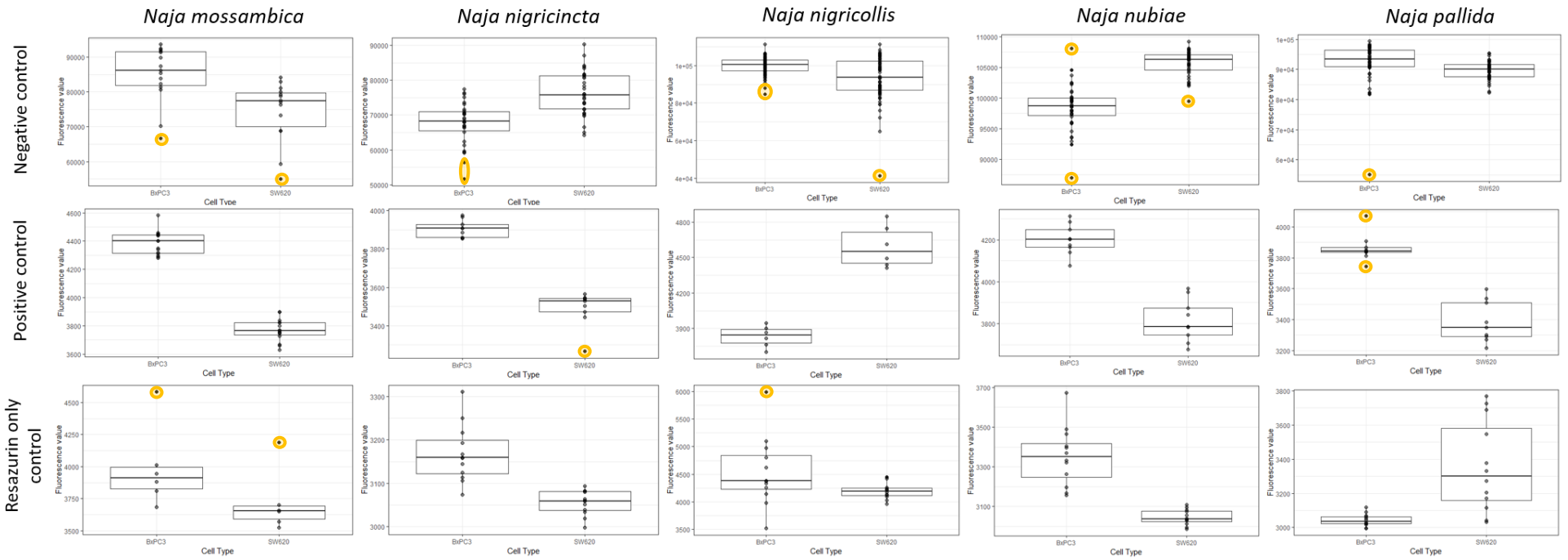
| Welch's ANOVA test | |
|---|---|
| W (DFn, DFd) | 46953 (5.000, 5.175) |
| P value | <0.0001 |
| P value summary | **** |
| Significant diff. among means (P < 0.05)? | Yes |

| Dunnett's T3 multiple comparisons test | Mean Diff. | 95.00% CI of diff. | Below threshold? | Summary | Adjusted P Value | A-? |
|---|---|---|---|---|---|---|
| Crude vs. H2O | -0.2 | -1.501 to 1.101 | No | ns | 0.5801 | B |
| Crude vs. DMSO | 1.5 | 0.3992 to 2.601 | Yes | ** | 0.0082 | C |
| Crude vs. IPA | 15.43 | 11.61 to 19.26 | Yes | ** | 0.0016 | D |
| Crude vs. RPC | 33.03 | 32.17 to 33.89 | Yes | **** | <0.0001 | E |
| Crude vs. RPD | 40.87 | 39.98 to 41.76 | Yes | **** | <0.0001 | F |

## Second Dimension Cobra Screening Statistics

Boxplots to show exclusion of outliers for each of the venom screens for the negative control, positive control and resazurin only control.

Any points which lie outside the “box and whiskers” plot (highlighted in gold) were excluded from the main data prior to data analysis.



*P-values for Normality Tests*

| N.mos | BxPC-3 | SW620 | N.nct | BxPC-3 | SW620 | N.nub | BxPC-3 | SW620 |
|---|---|---|---|---|---|---|---|---|
| i6r1 | 0.7082 | 0.6836 | i1r1 | 0.6558 | 0.5699 | i1r1 | 0.4132 | 0.8194 |
| i6r2 | 0.9082 | 0.7895 | i1r2 | 0.5725 | 0.4167 | i1r2 | 0.7752 | 0.4835 |
| i6r3 | 0.6331 | 0.1623 | i1r3 | 0.0396 | 0.3141 | i4r1 | 0.7441 | 0.25 |
| i6r4 | 0.4295 | 0.0226 | i1r4 | 0.5243 | 0.4227 | i4r2 | 0.2236 | 0.166 |
| i6r5 | 0.9301 | 0.7164 | i4r1 | 0.4898 | 0.9374 | i5r1 | 0.3881 | 0.009 |
| i6r6 | 0.5605 | 0.8445 | i4r2 | 0.0656 | 0.1287 | i5r2 | 0.0109 | 0.2546 |
| i6r7 | 0.0836 | 0.4636 | i4r3 | 0.4808 | 0.7638 | i7r1 | 0.6557 | 0.6233 |
| i11r1 | 0.1773 | 0.2835 | i4r4 | 0.2433 | 0.412 | i7r2 | 0.5164 | 0.6992 |
| i11r2 | 0.6778 | 0.4078 | i9r1 | 0.8647 | 0.0337 | i9r1 | 0.219 | 0.1744 |
| i11r3 | 0.2524 | 0.4014 | i9r2 | 0.2474 | 0.8868 | i9r2 | 0.1165 | 0.0893 |
| i11r4 | 0.045 | 0.9424 | i9r3 | 0.4545 | 0.0833 | i9r3 | 0.5817 | 0.3168 |
| i11r5 | 0.3404 | 0.6059 | i10r1 | 0.753 | 0.1711 | i12r1 | 0.2437 | 0.3041 |
| i13r1 | 0.7114 | 0.4573 | i10r2 | 0.4621 | 0.076 | i13r1 | 0.1916 | 0.0428 |
| i13r2 | 0.1333 | 0.9185 | i10r3 | 0.0584 | 0.9778 | i13r2 | 0.1332 | 0.8638 |
| i13r3 | 0.1907 | 0.2095 | i10r4 | 0.8213 | 0.8533 | i13r3 | 0.4282 | 0.8108 |
| i13r4 | 0.6782 | 0.5819 | i16r1 | 0.9494 | 0.35 | i14r1 | 0.1965 | 0.8179 |
| i16r1 | 0.3434 | 0.3959 | i16r2 | 0.8876 | 0.4061 | i14r2 | 0.7937 | 0.2237 |
| i16r2 | 0.3672 | 0.1099 | i16r3 | 0.9754 | 0.9949 | i14r3 | 0.1906 | 0.2811 |
| i16r3 | 0.2781 | 0.1475 | i18r1 | 0.2786 | 0.8475 | i14r4 | 0.5775 | 0.4633 |
| i16r4 | 0.1867 | 0.863 | i18r2 | 0.0967 | 0.0752 | i14r5 | 0.4474 | 0.5494 |
| i16r5 | 0.6066 | 0.4782 | i19r1 | 0.3685 | 0.5533 | i14r6 | 0.7861 | 0.0345 |
| i16r6 | 0.3647 | 0.7134 | i19r2 | 0.5105 | 0.6791 | i14r7 | 0.6773 | 0.5025 |
| i16r7 | 0.553 | 0.5581 | i19r3 | 0.8212 | 0.2197 | i14r8 | 0.0573 | 0.0287 |
| i16r8 | 0.9915 | 0.9009 | i19r4 | 0.3081 | 0.3207 | i17r1 | 0.6919 | 0.6088 |
| i20r1 | 0.2123 | 0.8752 | i20r1 | 0.3728 | 0.6835 | i17r2 | 0.4053 | 0.6806 |
| i20r2 | 0.1971 | 0.7112 | i20r2 | 0.8593 | 0.1884 | i17r3 | 0.6057 | 0.9427 |
| i20r3 | 0.2589 | 0.7076 | i20r3 | 0.0443 | 0.0028 | i17r4 | 0.4624 | 0.6701 |
| i20r4 | 0.7489 | 0.0927 | i20r4 | 0.176 | 0.4223 | i17r5 | 0.0101 | 0.5185 |
| i20r5 | 0.0142 | 0.1809 | i20r5 | 0.0222 | 0.111 | i18r1 | 0.5352 | 0.2255 |
| i20r6 | 0.8533 | 0.687 | i20r6 | 0.3261 | 0.5933 | i18r2 | 0.4144 | 0.9849 |
| i20r7 | 0.0023 | 0.9775 | i22r1 | 0.3608 | 0.254 | i18r3 | 0.1714 | 0.6302 |
| i20r8 | 0.7433 | 0.0028 | i22r2 | 0.5985 | 0.1206 | i18r4 | 0.0096 | 0.45 |
| i21r1 | 0.7183 | 0.4102 | i22r3 | 0.6243 | 0.5502 | Neg con | 0.8295 | 0.0522 |
| i21r2 | 0.0686 | 0.9571 | i23r1 | 0.9625 | 0.477 | | | |
| i21r3 | 0.3479 | 0.4064 | i23r2 | 0.0542 | 0.8264 | | | |
| i21r4 | 0.5651 | 0.761 | i23r3 | 0.4221 | 0.2358 | | | |
| i21r5 | 0.3685 | 0.3545 | Neg con | 0.3304 | 0.9367 | | | |
| i21r6 | 0.9704 | 0.5926 | | | | | | |
| i21r7 | 0.3035 | 0.8748 | | | | | | |
| i21r8 | 0.354 | 0.865 | | | | | | |
| i21r9 | 0.497 | 0.8968 | | | | | | |
| i25r1 | 0.6361 | 0.7638 | | | | | | |
| i25r2 | 0.6848 | 0.7178 | | | | | | |
| i26r1 | 0.3208 | 0.1847 | | | | | | |
| i26r2 | 0.4196 | 0.5687 | | | | | | |
| i26r3 | 0.9485 | 0.0709 | | | | | | |
| i27r1 | 0.8756 | 0.0446 | | | | | | |
| Neg con | 0.2097 | 0.2097 | | | | | | |

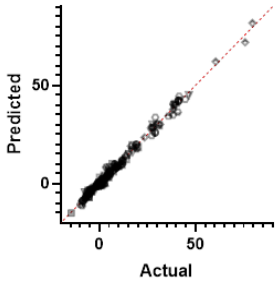
*P-values for Normality Tests continued*

| N.nig | BxPC-3 | SW620 | N.nig | BxPC-3 | SW620 | N.pal | BxPC-3 | SW620 |
|---|---|---|---|---|---|---|---|---|
| i1r1 | 0.3631 | 0.5555 | i12r10 | 0.4361 | 0.333 | i1r1 | 0.1214 | 0.6179 |
| i1r2 | 0.1036 | 0.5151 | i12r11 | 0.5692 | 0.3421 | i1r2 | 0.9543 | 0.0263 |
| i1r3 | 0.4194 | 0.2824 | i12r9 | 0.2252 | 0.9783 | i5r1 | 0.8914 | 0.4601 |
| i1r4 | 0.7023 | 0.2464 | i17r1 | 0.8807 | 0.3942 | i9r1 | 0.9508 | 0.8667 |
| i1r5 | 0.3867 | 0.7595 | i17r2 | 0.5815 | 0.6432 | i9r2 | 0.1531 | 0.2379 |
| i1r6 | 0.2842 | 0.7602 | i17r3 | 0.5915 | 0.4946 | i9r3 | 0.8013 | 0.1597 |
| i1r7 | 0.2054 | 0.5697 | i17r4 | 0.7163 | 0.7843 | i11r1 | 0.9394 | 0.2839 |
| i1r8 | 0.6825 | 0.9727 | i17r5 | 0.9829 | 0.2079 | i11r2 | 0.8311 | 0.6246 |
| i1r9 | 0.3281 | 0.3702 | i18r1 | 0.6767 | 0.2988 | i12r1 | 0.3379 | 0.1902 |
| i1r10 | 0.7937 | 0.2511 | i18r2 | 0.8379 | 0.0143 | i12r2 | 0.1513 | 0.669 |
| i1r11 | 0.7327 | 0.1213 | i18r3 | 0.1642 | 0.4577 | i13r1 | 0.3266 | 0.9103 |
| i5r1 | 0.2842 | 0.459 | i18r4 | 0.9578 | 0.3316 | i14r1 | 0.0434 | 0.4321 |
| i5r2 | 0.2922 | 0.8448 | i18r5 | 0.5214 | 0.4722 | i15r1 | 0.5112 | 0.1387 |
| i5r3 | 0.8144 | 0.9352 | i18r6 | 0.5753 | 0.0014 | i15r2 | 0.4666 | 0.5954 |
| i5r4 | 0.1066 | 0.4525 | i18r7 | 0.3682 | 0.5131 | i15r3 | 0.114 | 0.4918 |
| i5r5 | 0.0557 | 0.4325 | i20r2 | 0.2416 | 0.6142 | i15r4 | 0.4066 | 0.1082 |
| i5r6 | 0.3224 | 0.8832 | i20r3 | 0.8401 | 0.6535 | i15r5 | 0.7816 | 0.7825 |
| i5r7 | 0.8705 | 0.485 | i20r4 | 0.645 | 0.4568 | i17r1 | 0.4336 | 0.0668 |
| i5r8 | 0.8978 | 0.8122 | i20r5 | 0.1131 | 0.6371 | i17r2 | 0.0929 | 0.9504 |
| i5r9 | 0.4741 | 0.8691 | i20r6 | 0.2776 | 0.5581 | i17r3 | 0.0843 | 0.1001 |
| i5r10 | 0.9048 | 0.5823 | i20r7 | 0.2321 | 0.1115 | i17r4 | 0.6986 | 0.3177 |
| i6r1 | 0.4234 | 0.7868 | i20r8 | 0.7876 | 0.8747 | i18r1 | 0.0873 | 0.2389 |
| i6r2 | 0.3148 | 0.5175 | i20r9 | 0.2564 | 0.7793 | i18r2 | 0.6224 | 0.45 |
| i6r3 | 0.2129 | 0.758 | i20r10 | 0.9588 | 0.9231 | i18r3 | 0.5448 | 0.2781 |
| i6r4 | 0.8065 | 0.6921 | i20r11 | 0.8513 | 0.9652 | i18r4 | 0.9194 | 0.3137 |
| i6r5 | 0.7449 | 0.8698 | i21r1 | 0.2877 | 0.713 | i20r1 | 0.7871 | 0.4588 |
| i6r6 | 0.0558 | 0.0628 | i21r2 | 0.7948 | 0.0678 | i20r2 | 0.8603 | 0.6536 |
| i6r7 | 0.8168 | 0.1107 | i21r3 | 0.6951 | 0.7045 | Neg con | 0.0063 | 0.5404 |
| i6r8 | 0.7615 | 0.0237 | i21r4 | 0.0542 | 0.4037 | | | |
| i6r9 | 0.9034 | 0.2174 | i21r5 | 0.8691 | 0.3442 | | | |
| i6r10 | 0.8736 | 0.5663 | i21r6 | 0.4036 | 0.0118 | | | |
| i6r11 | 0.2851 | 0.1655 | i21r7 | 0.5171 | 0.2424 | | | |
| i7r1 | 0.8858 | 0.2839 | i21r8 | 0.6478 | 0.1316 | | | |
| i7r2 | 0.4251 | 0.2182 | i21r9 | 0.8733 | 0.2548 | | | |
| i7r3 | 0.0727 | 0.5849 | i21r10 | 0.9253 | 0.323 | | | |
| i7r4 | 0.3081 | 0.3806 | i21r11 | 0.5657 | 0.961 | | | |
| i7r5 | 0.6011 | 0.9829 | i22r1 | 0.3073 | 0.4215 | | | |
| i7r6 | 0.1389 | 0.6443 | i22r2 | 0.487 | 0.8072 | | | |
| i7r7 | 0.5457 | 0.9328 | i22r3 | 0.9954 | 0.2418 | | | |
| i12r1 | 0.8188 | 0.0236 | i22r4 | 0.7917 | 0.711 | | | |
| i12r2 | 0.2887 | 0.0446 | i22r5 | 0.5227 | 0.2301 | | | |
| i12r3 | 0.1218 | 0.1535 | i22r6 | 0.998 | 0.8624 | | | |
| i12r4 | 0.0207 | 0.8079 | i24r1 | 0.1372 | 0.8868 | | | |
| i12r5 | 0.1139 | 0.7066 | i24r2 | 0.0743 | 0.6797 | | | |
| i12r6 | 0.5367 | 0.5286 | i24r3 | 0.3746 | 0.9729 | | | |
| i12r7 | 0.0893 | 0.6916 | i24r4 | 0.776 | 0.3896 | | | |
| i12r8 | 0.3631 | 0.8003 | i24r5 | 0.7788 | 0.9167 | | | |

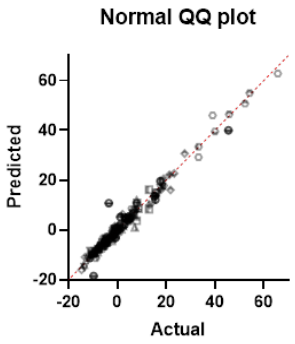
*QQ Plots*

## N.mos

### BxPC-3

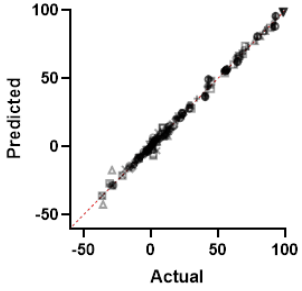



### SW620

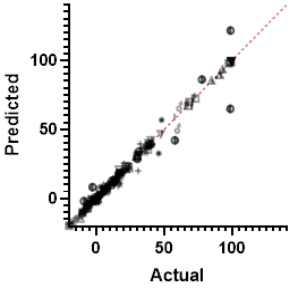



## N.nct

### BxPC-3

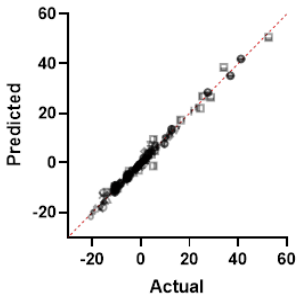



### SW620

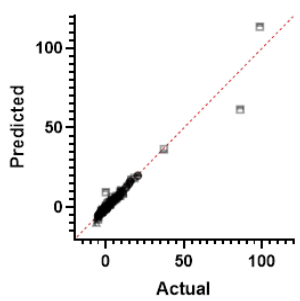



## N.nub

### BxPC-3

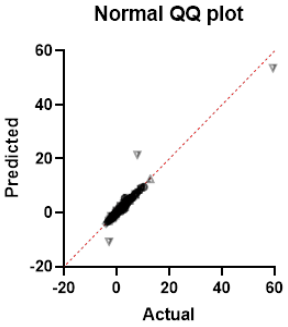


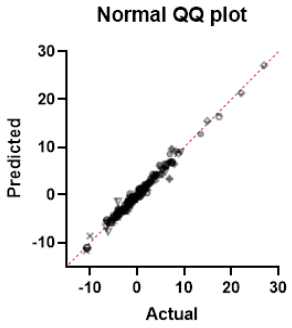

### SW620



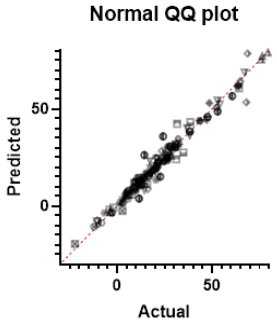

## N.nig

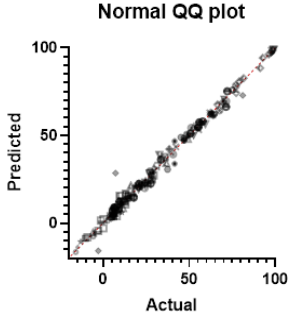
### BxPC-3



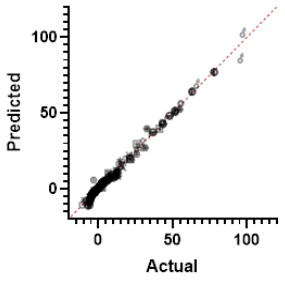

### SW620



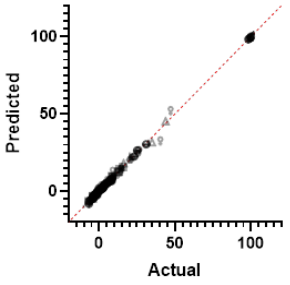




## N.pal

### BxPC-3




### SW620



*Kruskal-Wallis testing of 2D venom fraction assay*

| | N.mos | | N.nct | | N.nub | | N.pal | |
|---|---|---|---|---|---|---|---|---|
| | BxPC-3 | SW620 | BxPC-3 | SW620 | BxPC-3 | SW620 | BxPC-3 | SW620 |
| P value | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| P value summary | **** | **** | **** | **** | **** | **** | **** | **** |
| Do the medians vary signif. (P < 0.05)? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of groups | 48 | 48 | 37 | 37 | 33 | 33 | 28 | 28 |
| Kruskal-Wallis statistic | 93.75 | 110.9 | 112.3 | 102.2 | 97.86 | 76.92 | 94.43 | 94.3 |
| Data summary<br>Number of treatments (columns) | 48 | 48 | 37 | 37 | 33 | 33 | 28 | 28 |
| Number of values (total) | 154 | 154 | 136 | 138 | 121 | 135 | 121 | 123 |

*Dunn's Multiple Comparisons Test*

| | N.mos | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | BxPC-3 | | | | SW620 | | | |
| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value | Mean rank diff. | Significant? | Summary | Adjusted P Value |
| i16r5 vs. Neg con | 40.78 | No | ns | >0.9999 | -30.18 | No | ns | >0.9999 |
| i16r6 vs. Neg con | 48.45 | No | ns | 0.9888 | 4.154 | No | ns | >0.9999 |
| i20r1 vs. Neg con | 31.62 | No | ns | >0.9999 | 5.154 | No | ns | >0.9999 |
| i20r2 vs. Neg con | 5.782 | No | ns | >0.9999 | 48.15 | No | ns | >0.9999 |
| i20r7 vs. Neg con | 40.45 | No | ns | >0.9999 | 42.82 | No | ns | >0.9999 |
| i20r8 vs. Neg con | 27.12 | No | ns | >0.9999 | -11.18 | No | ns | >0.9999 |
| i21r2 vs. Neg con | 13.62 | No | ns | >0.9999 | 9.987 | No | ns | >0.9999 |
| i21r9 vs. Neg con | 42.12 | No | ns | >0.9999 | -3.846 | No | ns | >0.9999 |
| i25r1 vs. Neg con | 57.12 | No | ns | 0.5013 | 40.49 | No | ns | >0.9999 |
| i26r1 vs. Neg con | 80.45 | No | ns | 0.0535 | 90.49 | Yes | * | 0.0169 |
| i26r2 vs. Neg con | 83.12 | Yes | * | 0.0398 | 91.15 | Yes | * | 0.0156 |

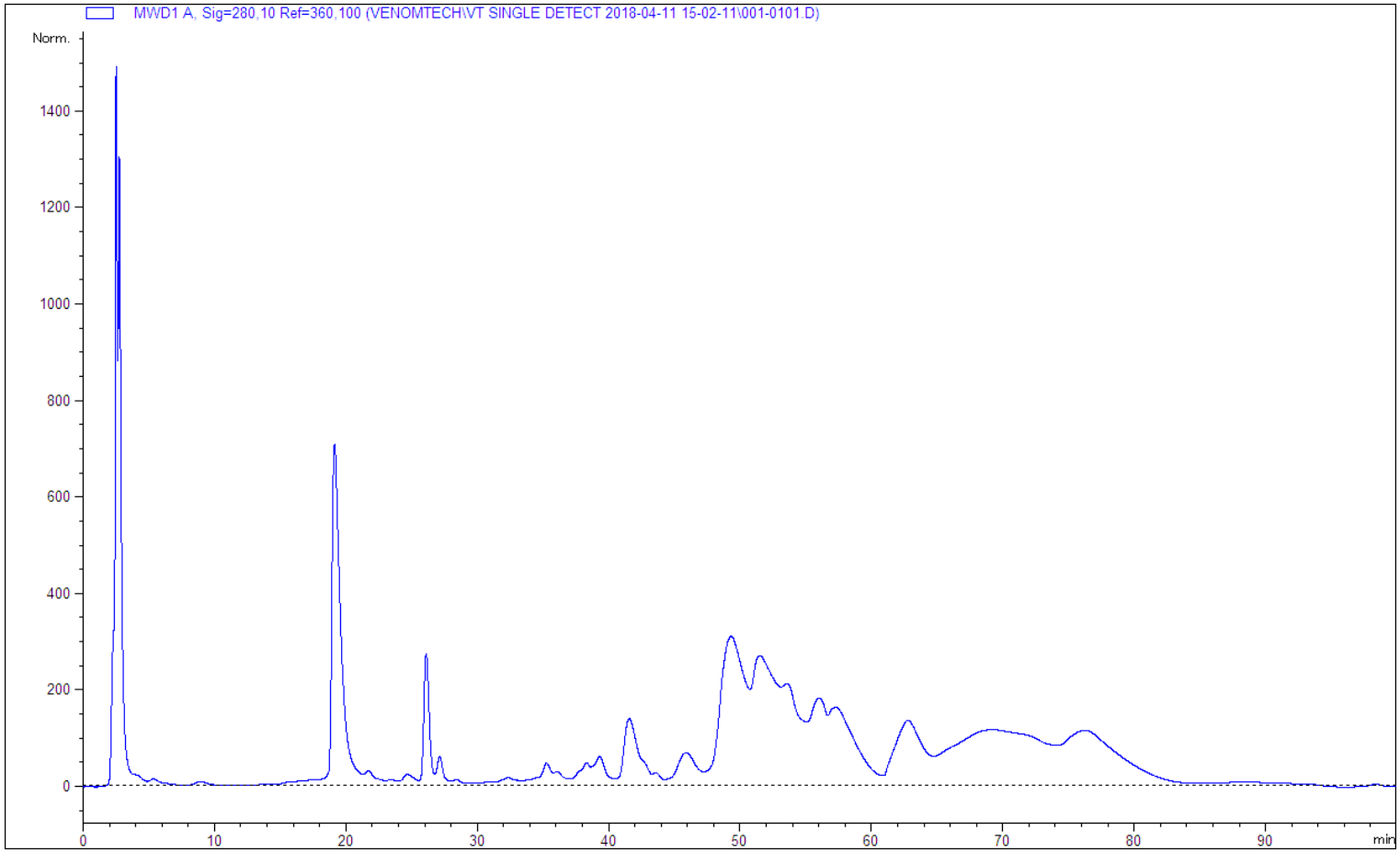
| | N.nub | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | BxPC-3 | | | | SW620 | | | |
| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value | Mean rank diff. | Significant? | Summary | Adjusted P Value |
| Neg con vs. i17r2 | -55.29 | Yes | ** | 0.0098 | 78.54 | Yes | *** | 0.001 |

| | N.nct | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | BxPC-3 | | | | SW620 | | | |
| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value | Mean rank diff. | Significant? | Summary | Adjusted P Value |
| i10r2 vs. Neg con | 57.27 | No | ns | 0.1171 | 38.32 | No | ns | 0.74 |
| i18r1 vs. Neg con | 55.61 | No | ns | 0.1413 | 33.65 | No | ns | >0.9999 |
| i18r2 vs. Neg con | 86.94 | Yes | ** | 0.002 | 74.32 | Yes | * | 0.0119 |
| i19r1 vs. Neg con | 73.61 | Yes | * | 0.0147 | 66.98 | Yes | * | 0.0328 |
| i20r1 vs. Neg con | 48.61 | No | ns | 0.2961 | 33.4 | No | ns | >0.9999 |
| i20r2 vs. Neg con | 65.61 | Yes | * | 0.0429 | 64.65 | Yes | * | 0.0444 |
| i20r5 vs. Neg con | 62.61 | No | ns | 0.0624 | 64.98 | Yes | * | 0.0425 |

| | N.pal | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | BxPC-3 | | | | SW620 | | | |
| Dunn's multiple comparisons test | Mean rank diff. | Significant? | Summary | Adjusted P Value | Mean rank diff. | Significant? | Summary | Adjusted P Value |
| i15r4 vs. Neg con | 61.91 | Yes | * | 0.0158 | 52.82 | No | ns | 0.0658 |
| i17r2 vs. Neg con | 67.91 | Yes | ** | 0.006 | 53.15 | No | ns | 0.063 |
| i18r2 vs. Neg con | 68.57 | Yes | ** | 0.0054 | 59.49 | Yes | * | 0.0262 |
| i18r3 vs. Neg con | 73.57 | Yes | ** | 0.0023 | 61.82 | Yes | * | 0.0186 |
| i20r1 vs. Neg con | 69.91 | Yes | ** | 0.0043 | 62.15 | Yes | * | 0.0177 |

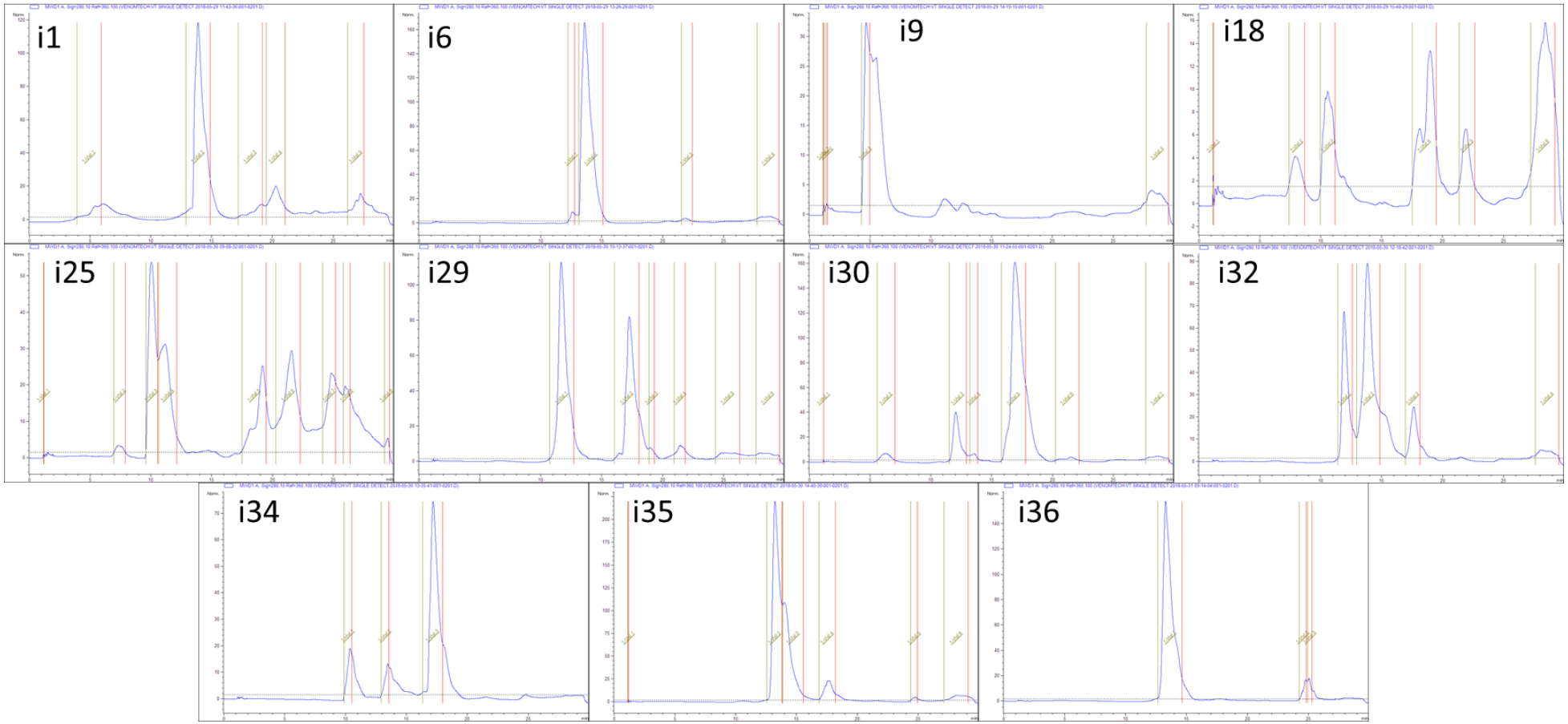
# Appendix III: HPLC Traces

*Raw HPLC Traces - First Dimension Ion Exchange Traces -N.mos 1D Ion Exchange Trace:*

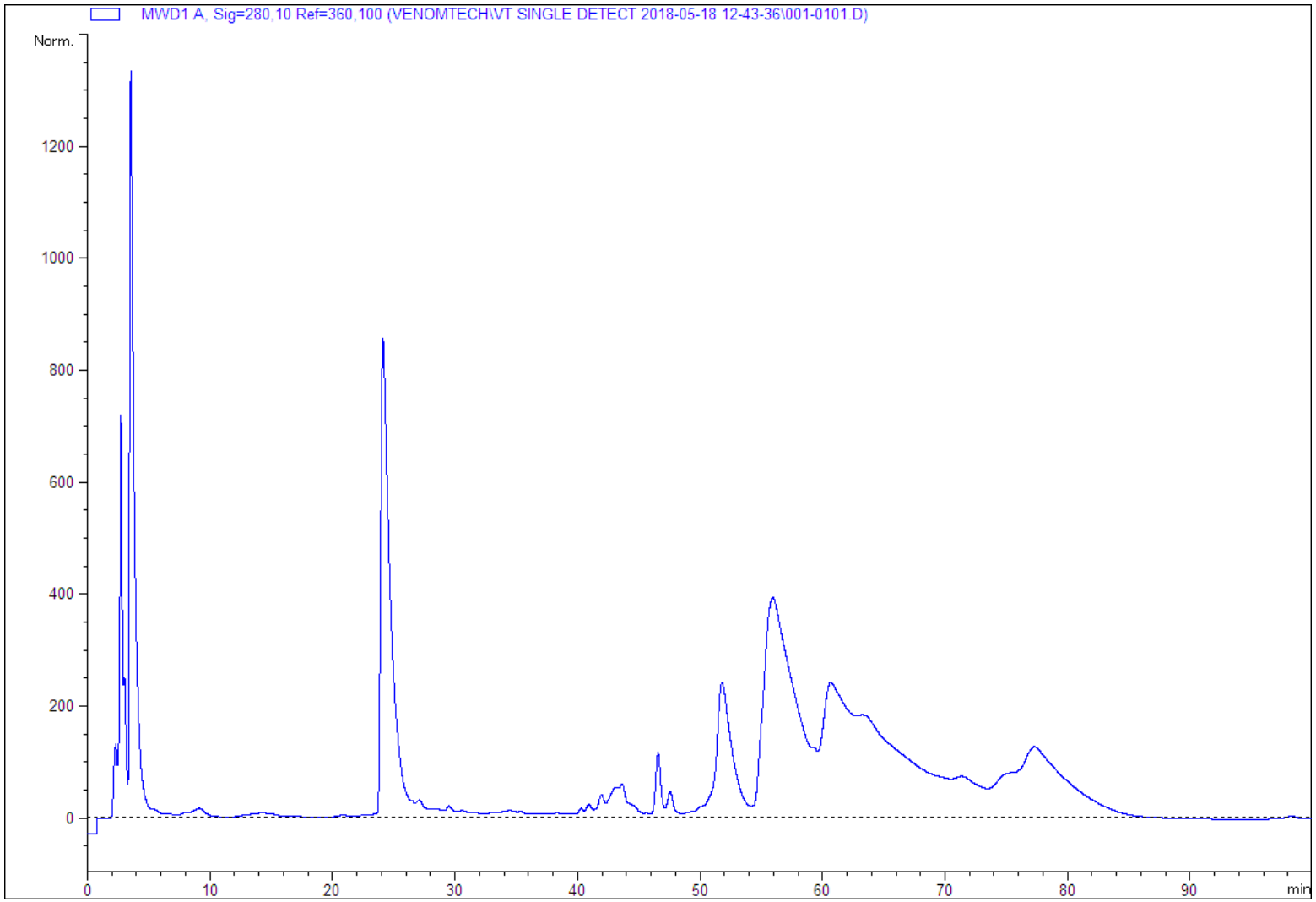



C

*N.mos 2D HPLC traces:*

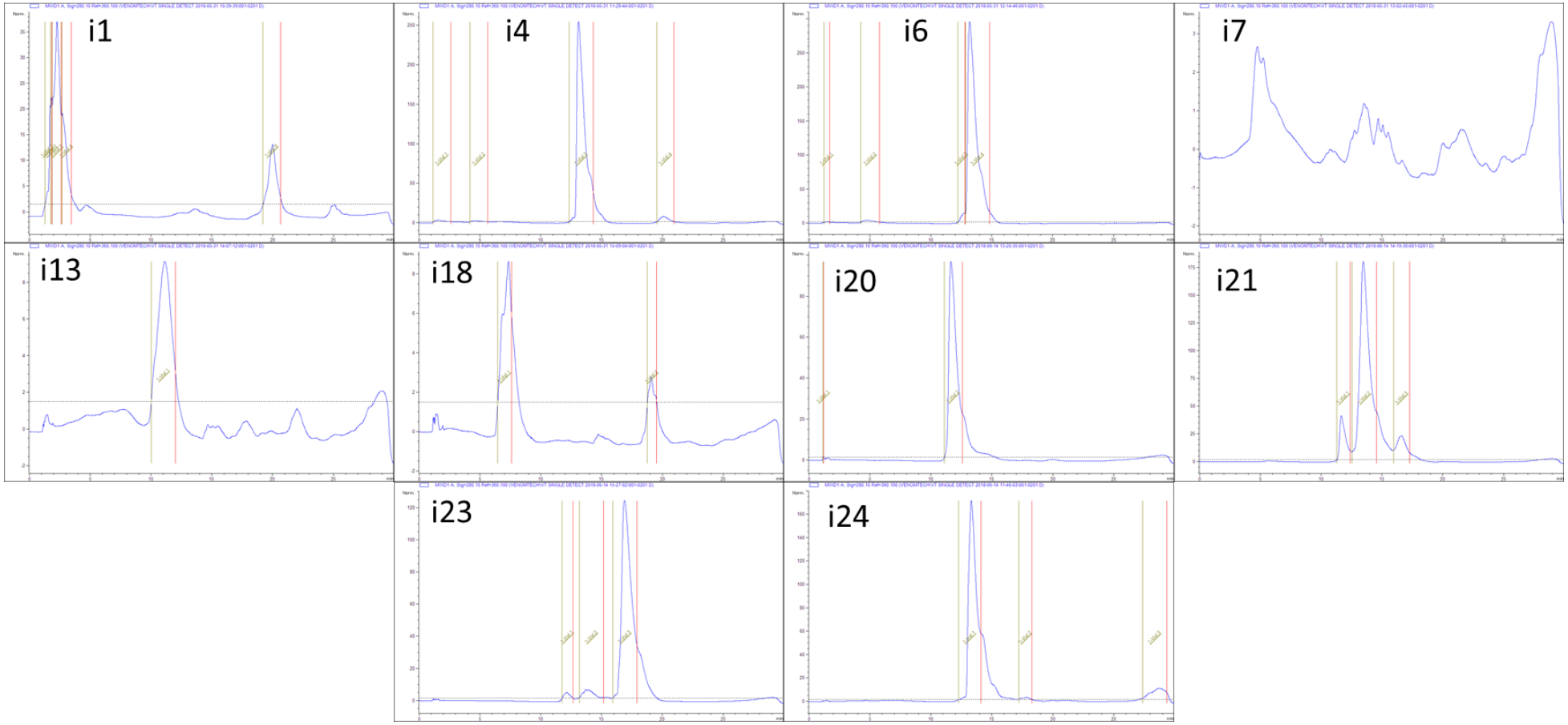


*N.nct 1D Ion Exchange Trace:*

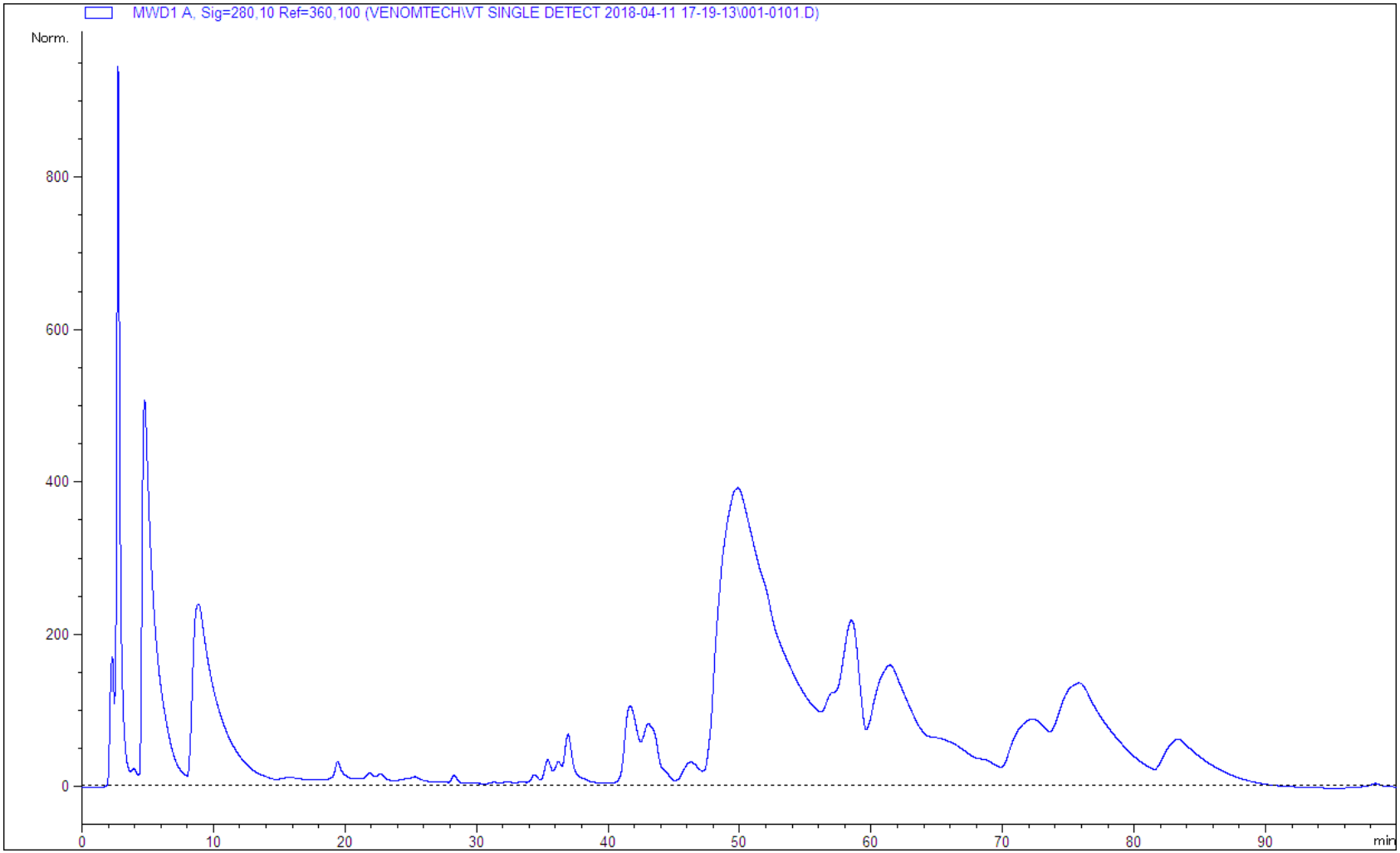



CII

*N.nct 2D HPLC traces:*

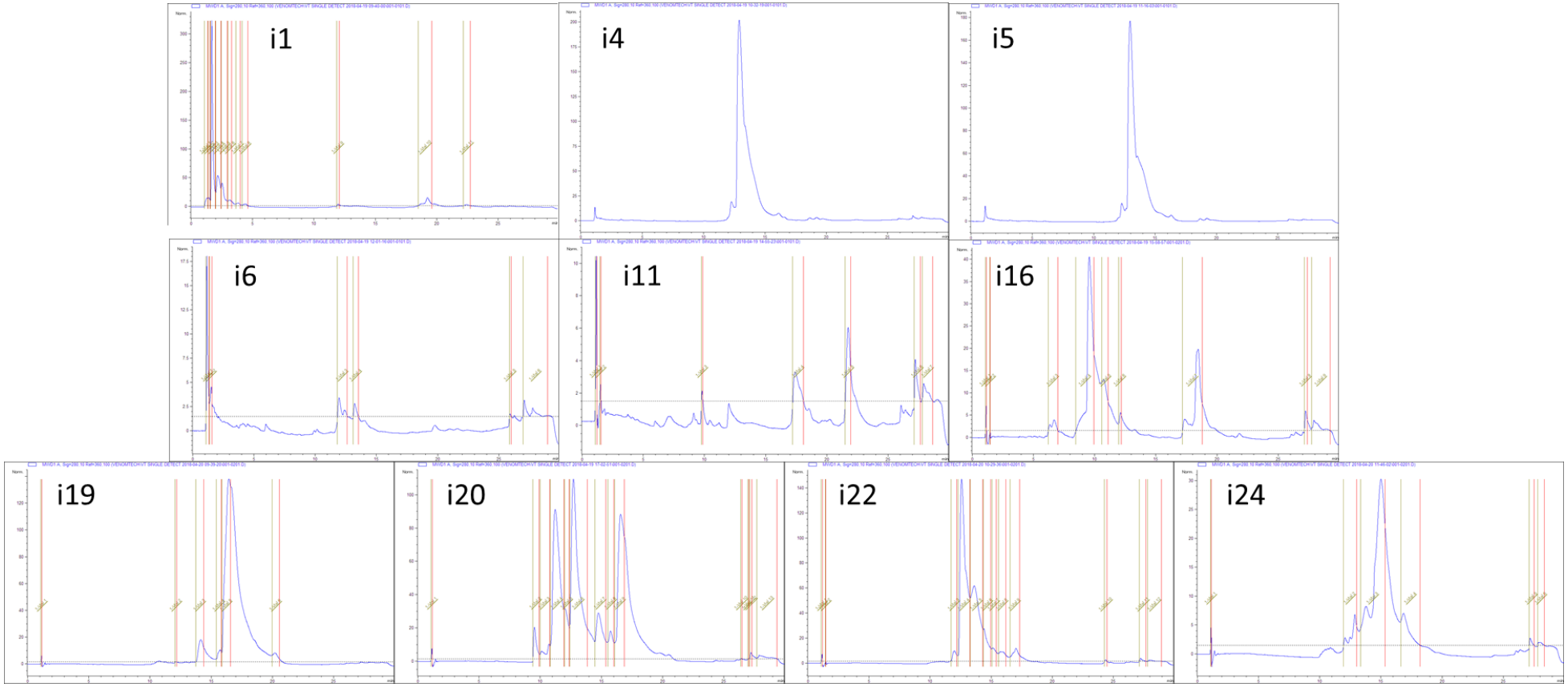


*N.nig 1D Ion Exchange Trace:*

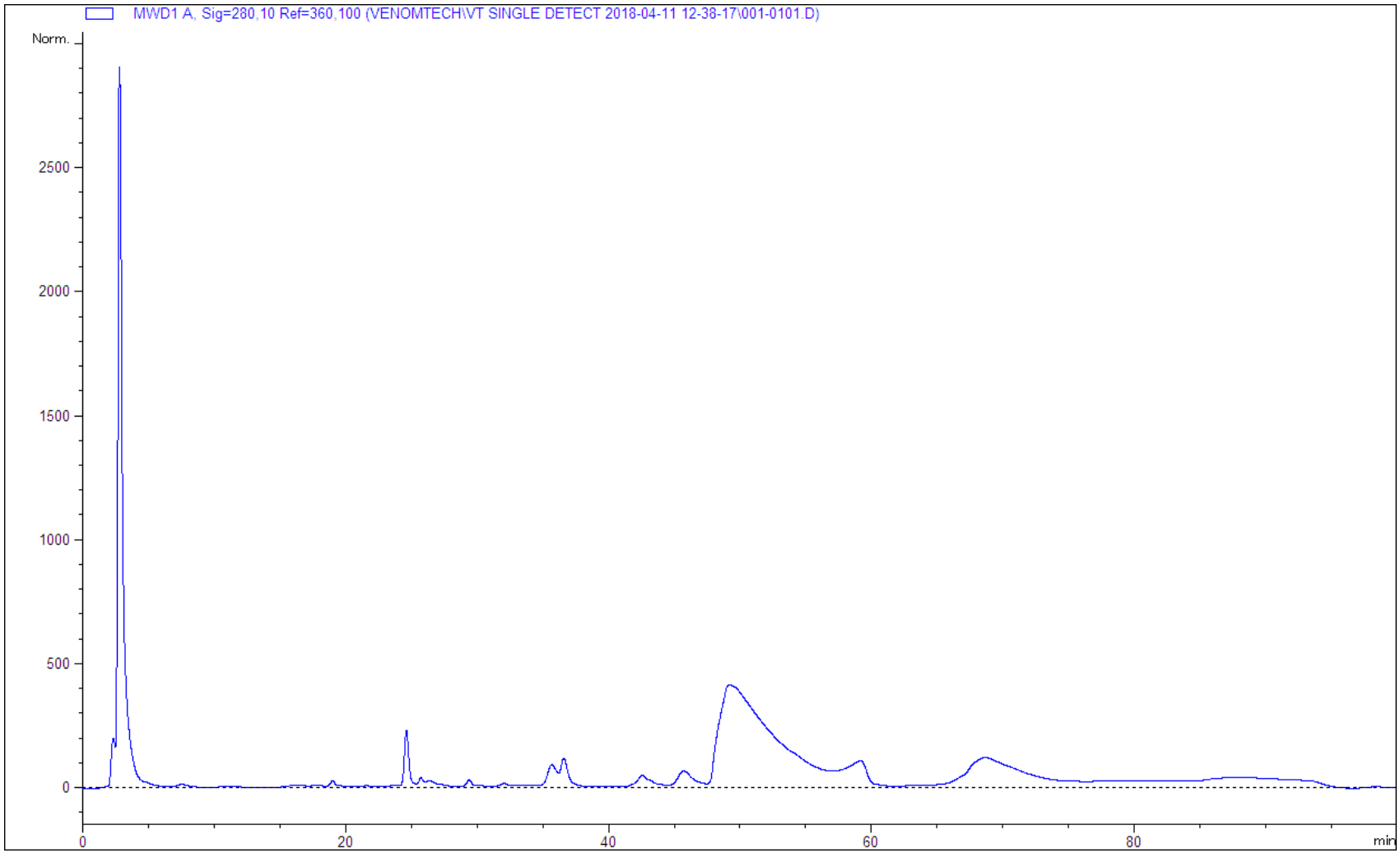



CIV

*N.nig 2D HPLC traces:*

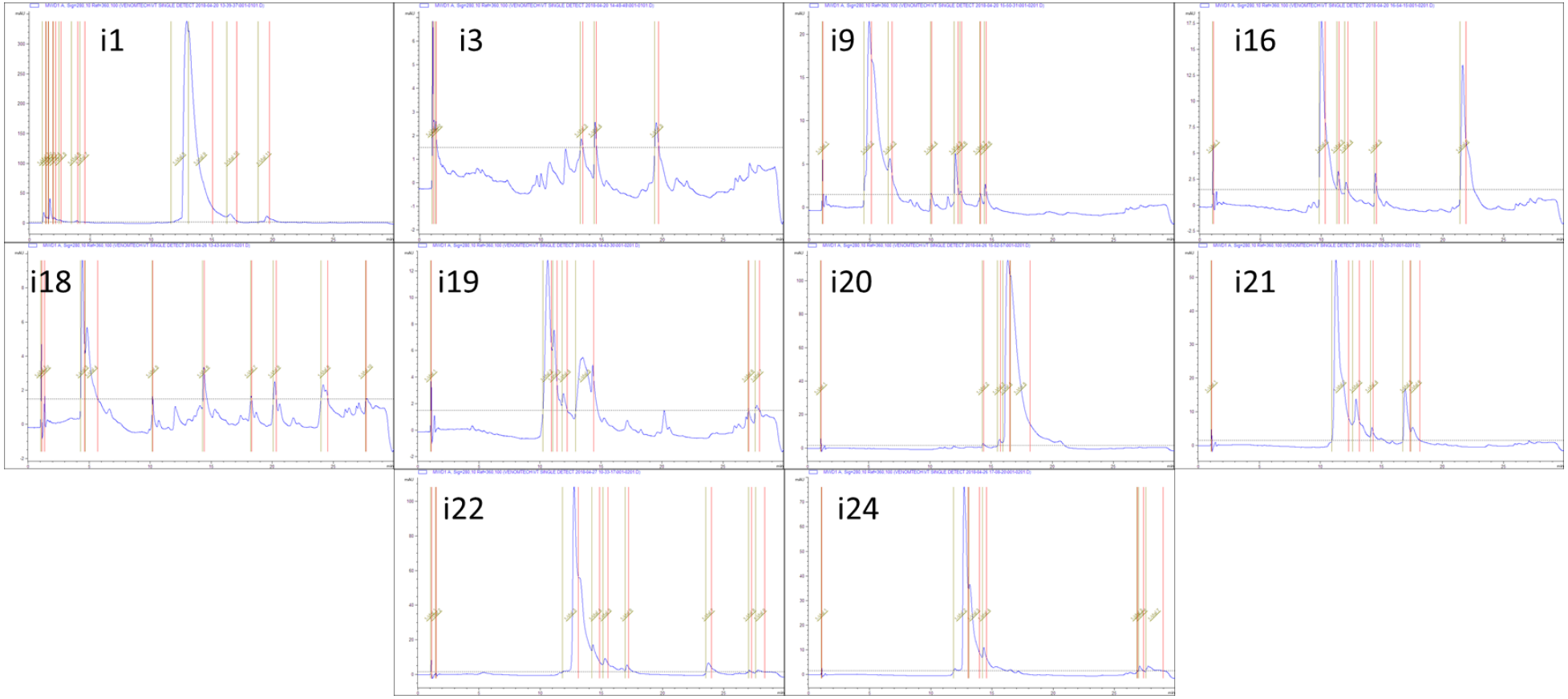


*N.nub 1D Ion Exchange Trace:*

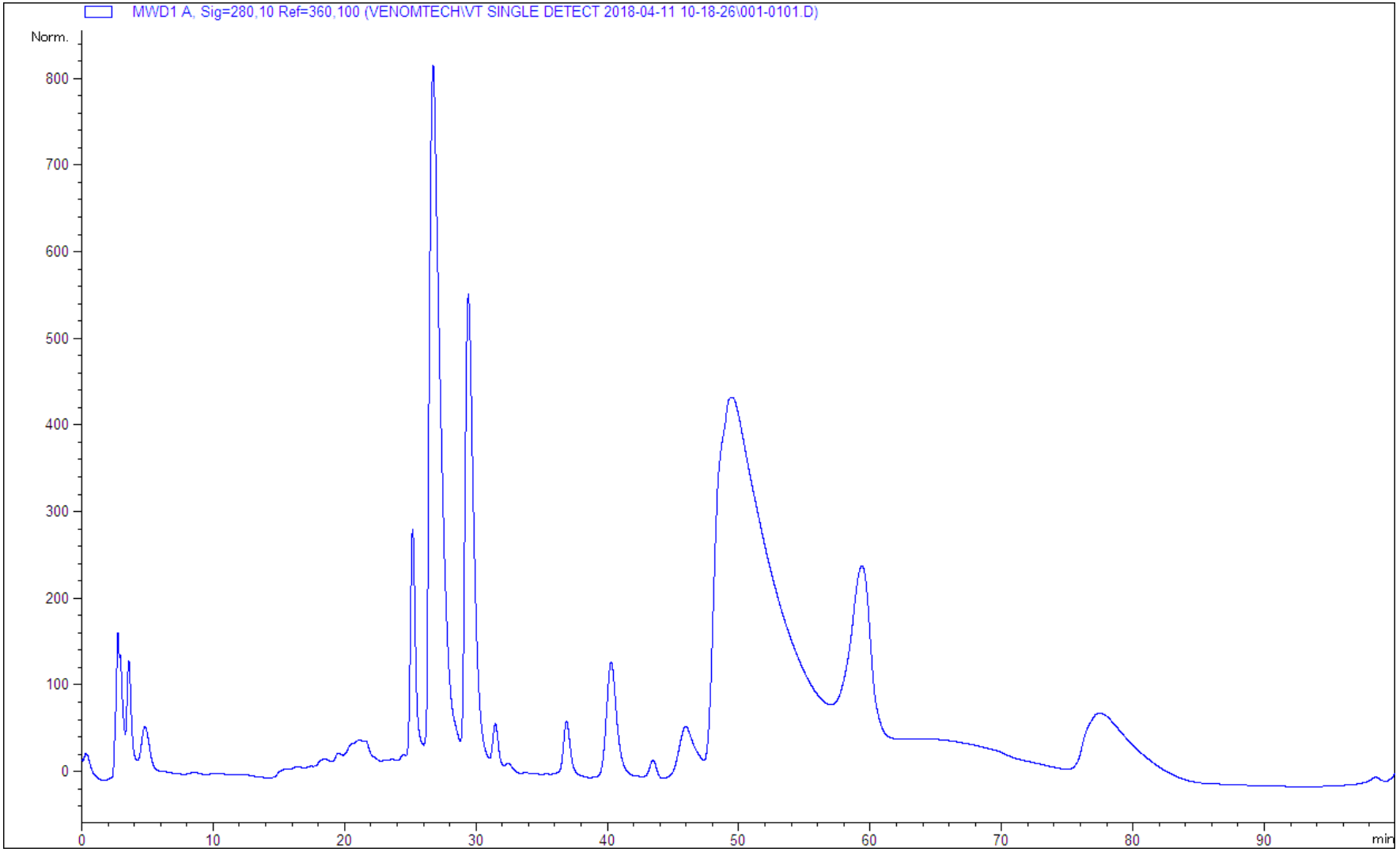



CVI

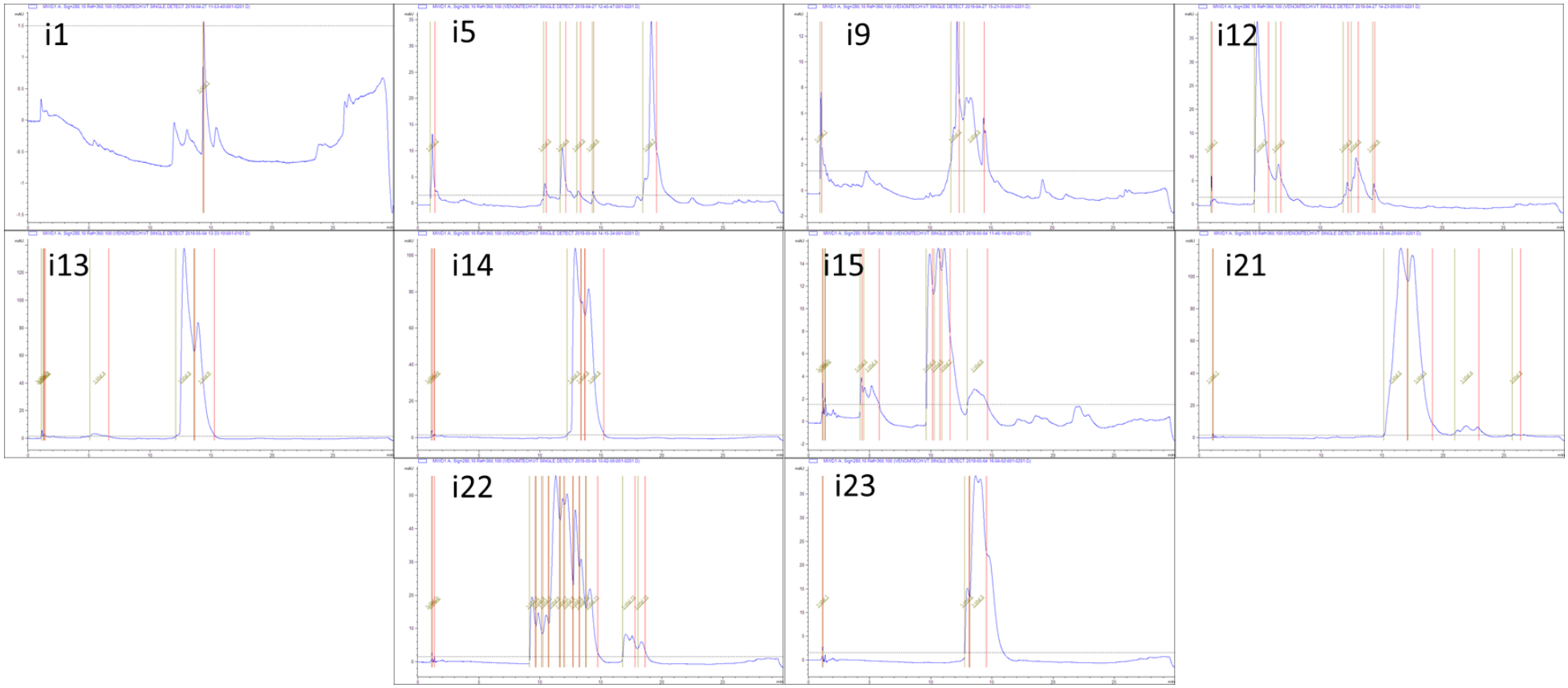
*N.nub 2D HPLC traces:*



*N.pal 1D Ion Exchange Trace:*



*N.pal 2D HPLC traces:*



# Appendix IV: Mass Spectroscopy Data

## *MS Raw Data*

### *N.pal_i15r4:*

**Peptide Mapping:**

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| ☑ 1.1 | 3SA1_NAJPA | 242 | 7279 | 17 (13) | 7 (5) | 35.11 | Cytotoxin 1 OS=Naja pallida OX=8658 PE=1 SV=1 |
| ☑ 1.2 | 3SA1_NAJNA | 62 | 7243 | 6 (4) | 3 (2) | 1.83 | Cytotoxin 1 OS=Naja naja OX=35670 PE=1 SV=1 |
| | ▼4 samesets of 3SA1_NAJNA | | | | | | |
| | 3SA3_NAJNA | 62 | 7197 | 6 (4) | 3 (2) | | Cytotoxin 3 OS=Naja naja OX=35670 PE=1 SV=1 |
| | 3SA7A_NAJKA | 62 | 7197 | 6 (4) | 3 (2) | | Cytotoxin 2 OS=Naja kaouthia OX=8649 PE=1 SV=1 |
| | 3SA7_NAJNA | 62 | 7244 | 6 (4) | 3 (2) | | Cytotoxin 7 OS=Naja naja OX=35670 PE=1 SV=1 |
| | 3SA8_NAJNA | 62 | 7245 | 6 (4) | 3 (2) | | Cytotoxin 8 OS=Naja naja OX=35670 PE=1 SV=2 |

Redisplay | All | None

▼18 peptide matches (11 non-duplicate, 7 duplicate)

☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | 1 | 2 | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 105 | ▶1 | 406.7400 | 811.4654 | 811.4626 | 3.49 | 0 | 54 | 0.00053 | ▶1 | U | ■ | | R.AAPMVPVK.R |
| 109 | | 409.2634 | 816.5123 | 816.5109 | 1.68 | 0 | 36 | 0.012 | ▶1 | U | | ■ | K.LIPLAYK.T |
| 121 | | 414.7361 | 827.4577 | 827.4575 | 0.22 | 0 | 40 | 0.01 | ▶1 | U | ■ | | R.AAPMVPVK.R + Oxidation (M) |
| 234 | ▶2 | 474.7172 | 947.4198 | 947.4205 | -0.68 | 0 | 43 | 0.0022 | ▶1 | | ■ | ■ | R.GCIDVCPK |
| 249 | | 484.7899 | 967.5652 | 967.5637 | 1.54 | 1 | 36 | 0.0051 | ▶1 | U | ■ | | R.AAPMVPVKR.G |
| 250 | | 323.5293 | 967.5660 | 967.5637 | 2.39 | 1 | 30 | 0.0062 | ▶1 | U | ■ | | R.AAPMVPVKR.G |
| 264 | ▶1 | 492.7879 | 983.5613 | 983.5586 | 2.72 | 1 | 49 | 0.00028 | ▶1 | U | ■ | | R.AAPMVPVKR.G + Oxidation (M) |
| 379 | ▶1 | 552.7702 | 1103.5259 | 1103.5216 | 3.87 | 1 | 25 | 0.27 | ▶1 | | ■ | ■ | K.RGCIDVCPK |
| 484 | ▶1 | 651.8366 | 1301.6586 | 1301.6591 | -0.35 | 0 | 55 | 0.00013 | ▶1 | U | ■ | | K.CNQLIPPFWK.T |
| 532 | | 683.2389 | 1364.4633 | 1364.4618 | 1.10 | 1 | 73 | 8.3e-006 | ▶1 | U | ■ | | K.YMCCNTDKCN.- |
| 614 | ▶1 | 515.2870 | 1542.8392 | 1542.8381 | 0.74 | 1 | 21 | 0.23 | ▶1 | U | ■ | | -.LKCNQLIPPFWK.T |

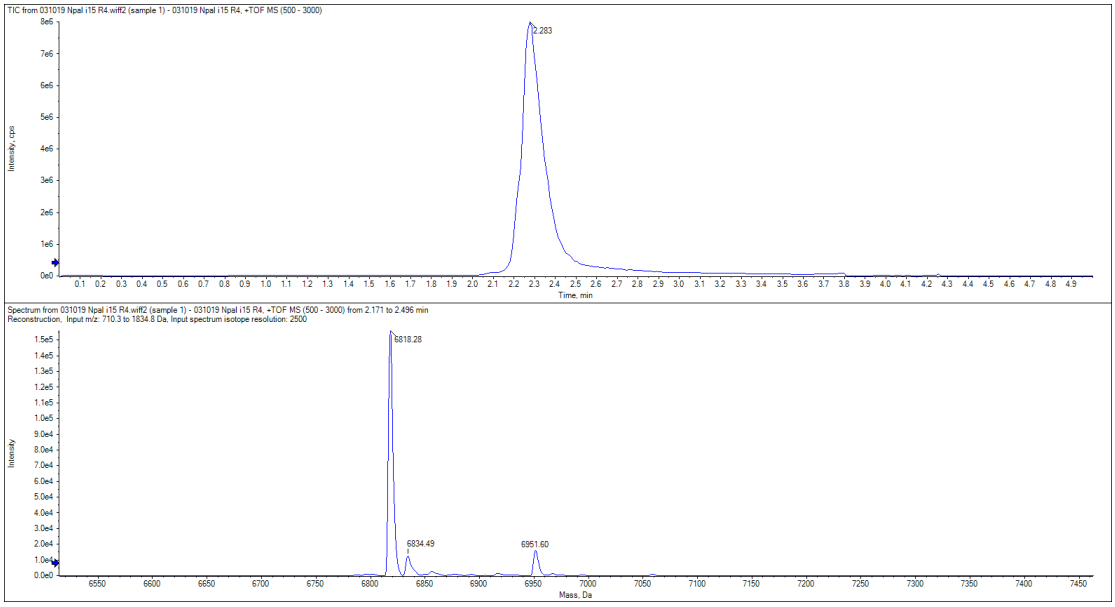
| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| 6.1 | PA2B1_NAJME | 44 | 14262 | 1 (1) | 1 (1) | 0.32 | Basic phospholipase A2 1 OS=Naja melanoleuca OX=8643 PE=1 SV=1 |
| | ▼4 samesets of PA2B1_NAJME | | | | | | |
| | PA2B3_NAJMO | 44 | 14091 | 1 (1) | 1 (1) | | Basic phospholipase A2 CM-III OS=Naja mossambica OX=8644 PE=1 SV=1 |
| | PA2B4_NAJNG | 44 | 14057 | 1 (1) | 1 (1) | | Phospholipase A2 ""basic"" OS=Naja nigricollis OX=8654 PE=1 SV=1 |
| | PA2TC_OXYSC | 44 | 14102 | 1 (1) | 1 (1) | | Neutral phospholipase A2 homolog taipoxin beta chain 2 OS=Oxyuranus scutellatus scutell |
| | PA22_OXYSC | 44 | 16948 | 1 (1) | 1 (1) | | Phospholipase A2 OS2 OS=Oxyuranus scutellatus scutellatus OX=8667 PE=1 SV=2 |

▼1 peptide matches (1 non-duplicate, 0 duplicate)

☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 267 | | 494.2407 | 986.4668 | 986.4669 | -0.12 | 0 | 44 | 0.00031 | ▶1 | U | K.GTPVDDLDR.C |

**Intact Mass:** Main peak at 6818.28Da and minor peak at 6951.60Da




N.pal_i15_R4

### Cytotoxin 1 OS=Naja pallida OX=8658 PE=1 SV=1

**Database:** SwissProt
**Score:** 242
**Nominal mass (M,):** 7279
**Calculated pI:** 9.41
**Taxonomy:** Naja pallida

Sequence similarity is available as an NCBI BLAST search of 3SA1_NAJPA against nr.

### Search parameters

**MS data file:** C:\Rachel\031019 Npal i15 R4 pm.mgf
**Enzyme:** Trypsin: cuts C-term side of KR unless next residue is P.
**Fixed modifications:** Carbamidomethyl (C)
**Variable modifications:** Oxidation (M), Phospho (ST), Phospho (Y)

### Protein sequence coverage: 65%

Matched peptides shown in ***bold red***.

```
  1 LKCNQLIPPF WKTCPKGKNL CYKMTMRAAP MVPVKRGCID VCPKSSLLIK
 51 YMCCNTDKCN
```

Unformatted sequence string: 60 residues (for pasting into other applications).

Sort peptides by ◉ Residue Number ○ Increasing Mass ○ Decreasing Mass

Show predicted peptides also

| Query | Start – End | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 613 | 1 – 12 | 515.2847 | 1542.8324 | 1542.8381 | -3.68 | 1 | 10 | 2.4 | 3 | U | -.LKCNQLIPPFWK.T |
| 614 | 1 – 12 | 515.2870 | 1542.8392 | 1542.8381 | 0.74 | 1 | 21 | 0.23 | 1 | U | -.LKCNQLIPPFWK.T |
| 483 | 3 – 12 | 651.8339 | 1301.6532 | 1301.6591 | -4.51 | 0 | 55 | 0.00038 | 1 | U | K.CNQLIPPFWK.T |
| 484 | 3 – 12 | 651.8366 | 1301.6586 | 1301.6591 | -0.35 | 0 | 55 | 0.00013 | 1 | U | K.CNQLIPPFWK.T |
| 104 | 28 – 35 | 406.7399 | 811.4653 | 811.4626 | 3.33 | 0 | 46 | 0.0033 | 1 | U | R.AAPMVPVK.R |
| 105 | 28 – 35 | 406.7400 | 811.4654 | 811.4626 | 3.49 | 0 | 54 | 0.00053 | 1 | U | R.AAPMVPVK.R |
| 121 | 28 – 35 | 414.7361 | 827.4577 | 827.4575 | 0.22 | 0 | 40 | 0.01 | 1 | U | R.AAPMVPVK.R + Oxidation (M) |
| 249 | 28 – 36 | 484.7899 | 967.5652 | 967.5637 | 1.54 | 1 | 36 | 0.0051 | 1 | U | R.AAPMVPVKR.G |
| 250 | 28 – 36 | 323.5293 | 967.5660 | 967.5637 | 2.39 | 1 | 30 | 0.0062 | 1 | U | R.AAPMVPVKR.G |
| 263 | 28 – 36 | 492.7879 | 983.5612 | 983.5586 | 2.61 | 1 | 39 | 0.0013 | 1 | U | R.AAPMVPVKR.G + Oxidation (M) |
| 264 | 28 – 36 | 492.7879 | 983.5613 | 983.5586 | 2.72 | 1 | 49 | 0.00028 | 1 | U | R.AAPMVPVKR.G + Oxidation (M) |
| 378 | 36 – 44 | 552.7691 | 1103.5237 | 1103.5216 | 1.87 | 1 | 19 | 0.24 | 1 | | K.RGCIDVCPK.S |
| 379 | 36 – 44 | 552.7702 | 1103.5259 | 1103.5216 | 3.87 | 1 | 25 | 0.27 | 1 | | K.RGCIDVCPK.S |
| 234 | 37 – 44 | 474.7172 | 947.4198 | 947.4205 | -0.68 | 0 | 43 | 0.0022 | 1 | | R.GCIDVCPK.S |
| 235 | 37 – 44 | 474.7175 | 947.4205 | 947.4205 | 0.021 | 0 | 38 | 0.0045 | 1 | | R.GCIDVCPK.S |
| 236 | 37 – 44 | 474.7186 | 947.4227 | 947.4205 | 2.30 | 0 | 31 | 0.03 | 1 | | R.GCIDVCPK.S |
| 532 | 51 – 60 | 683.2389 | 1364.4633 | 1364.4618 | 1.10 | 1 | 73 | 8.3e-006 | 1 | U | K.YMCCNTDKCN.- |

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| ☑ 1.1 | 3SA1_NAJPA | 242 | 7279 | 17 (13) | 7 (5) | 35.11 | Cytotoxin 1 OS=Naja pallida OX=8658 PE=1 SV=1 |
| ☑ 1.2 | 3SA1_NAJNA | 62 | 7243 | 6 (4) | 3 (2) | 1.83 | Cytotoxin 1 OS=Naja naja OX=35670 PE=1 SV=1 |
| | ▸ 4 samesets of 3SA1_NAJNA | | | | | | |

Redisplay | All | None

▾ *18 peptide matches (11 non-duplicate, 7 duplicate)*

☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | 1 | 2 | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 105 | ▸1 | 406.7400 | 811.4654 | 811.4626 | 3.49 | 0 | 54 | 0.00053 | ▸1 | U | ■ | | R.AAPMVPVK.R |
| 109 | | 409.2634 | 816.5123 | 816.5109 | 1.68 | 0 | 36 | 0.012 | ▸1 | U | | ■ | K.LIPLAYK.T |
| 121 | | 414.7361 | 827.4577 | 827.4575 | 0.22 | 0 | 40 | 0.01 | ▸1 | U | ■ | | R.AAPMVPVK.R + Oxidation (M) |
| 234 | ▸2 | 474.7172 | 947.4198 | 947.4205 | -0.68 | 0 | 43 | 0.0022 | ▸1 | | ■ | ■ | R.GCIDVCPK |
| 249 | | 484.7899 | 967.5652 | 967.5637 | 1.54 | 1 | 36 | 0.0051 | ▸1 | U | ■ | | R.AAPMVPVKR.G |
| 250 | | 323.5293 | 967.5660 | 967.5637 | 2.39 | 1 | 30 | 0.0062 | ▸1 | U | ■ | | R.AAPMVPVKR.G |
| 264 | ▸1 | 492.7879 | 983.5613 | 983.5586 | 2.72 | 1 | 49 | 0.00028 | ▸1 | U | ■ | | R.AAPMVPVKR.G + Oxidation (M) |
| 379 | ▸1 | 552.7702 | 1103.5259 | 1103.5216 | 3.87 | 1 | 25 | 0.27 | ▸1 | | ■ | ■ | K.RGCIDVCPK |
| 484 | ▸1 | 651.8366 | 1301.6586 | 1301.6591 | -0.35 | 0 | 55 | 0.00013 | ▸1 | U | ■ | | K.CNQLIPPFWK.T |
| 532 | | 683.2389 | 1364.4633 | 1364.4618 | 1.10 | 1 | 73 | 8.3e-006 | ▸1 | U | ■ | | K.YMCCNTDKCN.- |
| 614 | ▸1 | 515.2870 | 1542.8392 | 1542.8381 | 0.74 | 1 | 21 | 0.23 | ▸1 | U | ■ | | -.LKCNQLIPPFWK.T |

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| 6.1 | PA2B1_NAJME | 44 | 14262 | 1 (1) | 1 (1) | 0.32 | Basic phospholipase A2 1 OS=Naja melanoleuca OX=8643 PE=1 SV=1 |
| | ▾ *4 samesets of PA2B1_NAJME* | | | | | | |
| | PA2B3_NAJMO | 44 | 14091 | 1 (1) | 1 (1) | | Basic phospholipase A2 CM-III OS=Naja mossambica OX=8644 PE=1 SV=1 |
| | PA2B4_NAJNG | 44 | 14057 | 1 (1) | 1 (1) | | Phospholipase A2 ""basic"" OS=Naja nigricollis OX=8654 PE=1 SV=1 |
| | PA2TC_OXYSC | 44 | 14102 | 1 (1) | 1 (1) | | Neutral phospholipase A2 homolog taipoxin beta chain 2 OS=Oxyuranus scutellatus scutellatus OX=8667 PE=1 SV=1 |
| | PA22_OXYSC | 44 | 16948 | 1 (1) | 1 (1) | | Phospholipase A2 OS2 OS=Oxyuranus scutellatus scutellatus OX=8667 PE=1 SV=2 |

▾ *1 peptide matches (1 non-duplicate, 0 duplicate)*

☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 267 | | 494.2407 | 986.4668 | 986.4669 | -0.12 | 0 | 44 | 0.00031 | ▸1 | U | K.GTPVDDLDR.C |

*N.pal_i17r2:*

N.pal_i17_R2

## Peptide Mapping:

### Cytotoxin 1 OS=Naja pallida OX=8658 PE=1 SV=1

**Database:** SwissProt
**Score:** 44
**Nominal mass ($M_r$):** 7279
**Calculated pI:** 9.41
**Taxonomy:** Naja pallida

Sequence similarity is available as an NCBI BLAST search of 3SA1_NAJPA against nr.

### Search parameters

**MS data file:** C:\Rachel\031019 Npal i17 R2 pm.mgf
**Enzyme:** Trypsin: cuts C-term side of KR unless next residue is P.
**Fixed modifications:** Carbamidomethyl (C)
**Variable modifications:** Oxidation (M), Phospho (ST), Phospho (Y)

### Protein sequence coverage: 15%

Matched peptides shown in ***bold red***.

```
  1 LKCNQLIPPF WKTCPKGKNL CYKMTMRAAP MVPVKRGCID VCPKSSLLIK
 51 YMCCNTDKCN
```

Unformatted sequence string: 60 residues (for pasting into other applications).

Sort peptides by ◉ Residue Number ○ Increasing Mass ○ Decreasing Mass

Show predicted peptides also

| Query | Start – End | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 362 | 36 – 44 | 552.7686 | 1103.5226 | 1103.5216 | 0.93 | 1 | 32 | 0.065 | 1 | U | K.RGCIDVCPK.S |
| 363 | 36 – 44 | 552.7686 | 1103.5227 | 1103.5216 | 0.97 | 1 | 36 | 0.013 | 1 | U | K.RGCIDVCPK.S |
| 243 | 37 – 44 | 474.7175 | 947.4205 | 947.4205 | -0.023 | 0 | 40 | 0.032 | 1 | U | R.GCIDVCPK.S |
| 244 | 37 – 44 | 474.7181 | 947.4217 | 947.4205 | 1.28 | 0 | 29 | 0.021 | 1 | U | R.GCIDVCPK.S |

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| ☑ 2.1 | 3SA1_NAJNI | 82 | 7148 | 6 (4) | 4 (3) | 5.18 | Cytotoxin 1 OS=Naja nivea OX=8655 PE=1 SV=1 |
| | ▼*3 samesets of 3SA1_NAJNI* | | | | | | |
| | 3SA2_NAJHA | 82 | 7309 | 6 (4) | 4 (3) | | Cytotoxin 2 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| | 3SA6_NAJHA | 82 | 7308 | 6 (4) | 4 (3) | | Cytotoxin 6 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| | 3SA9_NAJHA | 82 | 7120 | 6 (4) | 4 (3) | | Cytotoxin 9 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| ☑ 2.2 | IHH_CHICK | 34 | 45200 | 2 (2) | 2 (2) | 0.23 | Indian hedgehog protein OS=Gallus gallus OX=9031 GN=IHH PE=2 SV=1 |

Redisplay | All | None

▼*8 peptide matches (6 non-duplicate, 2 duplicate)*

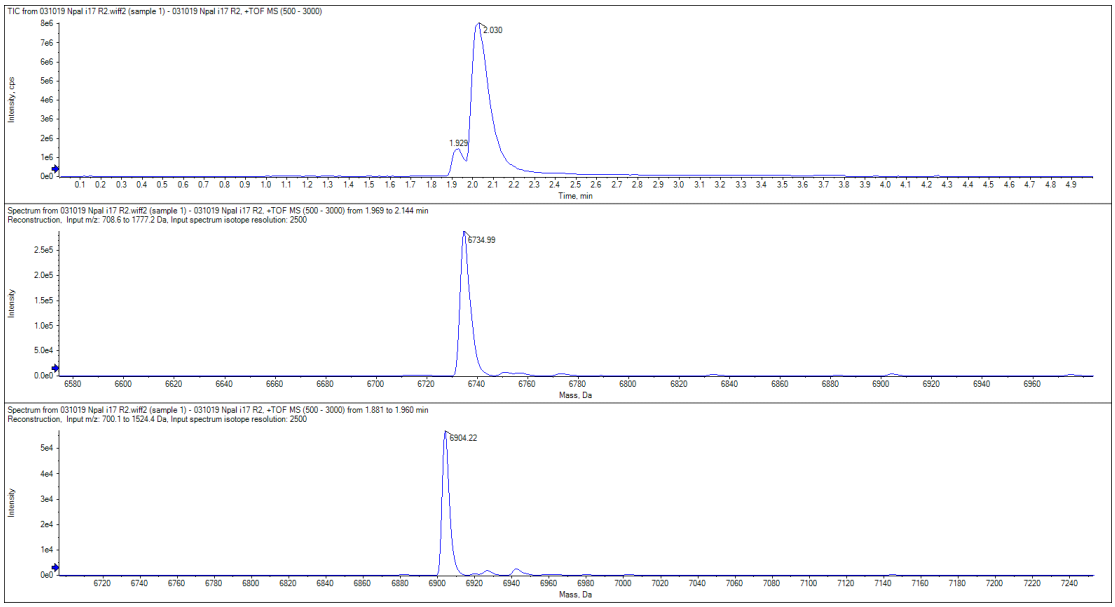
☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | 1 | 2 | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | | 409.2638 | 816.5131 | 816.5109 | 2.76 | 0 | 35 | 0.015 | ▶1 | U | ■ | | K.LIPLAYK.Q |
| 238 | | 473.3078 | 944.6011 | 944.6059 | -5.03 | 1 | 26 | 0.058 | ▶1 | U | ■ | | R.KLIPLAYK.Q |
| 243 | ▶1 | 474.7175 | 947.4205 | 947.4205 | -0.023 | 0 | 40 | 0.032 | ▶1 | U | ■ | | R.GCIDVCPK.D |
| 363 | ▶1 | 552.7686 | 1103.5227 | 1103.5216 | 0.97 | 1 | 36 | 0.013 | ▶1 | U | ■ | | K.RGCIDVCPK.D |
| 507 | | 457.2129 | 1368.6168 | 1368.6166 | 0.15 | 1 | 21 | 0.077 | ▶1 | U | ■ | | K.TCPEGKNLCYK.M |
| 574 | | 781.3922 | 1560.7699 | 1560.7640 | 3.78 | 1 | 67 | 6.8e-006 | ▶1 | U | ■ | | R.GCIDVCPKDSALVK.Y |

▼*3 subsets and intersections (72 subset proteins in total)*

| | Score | Mass | Subset of | |
|---|---|---|---|---|
| 3SA1_NAJNA | 49 | 7243 | 2.1, 2.2 | Cytotoxin 1 OS=Naja naja OX=35670 PE=1 SV=1 |
| ▼*4 samesets of 3SA1_NAJNA* | | | | |
| 3SA3_NAJNA | 49 | 7197 | | Cytotoxin 3 OS=Naja naja OX=35670 PE=1 SV=1 |
| 3SA7A_NAJKA | 49 | 7197 | | Cytotoxin 2 OS=Naja kaouthia OX=8649 PE=1 SV=1 |
| 3SA7_NAJNA | 49 | 7244 | | Cytotoxin 7 OS=Naja naja OX=35670 PE=1 SV=1 |
| 3SA8_NAJNA | 49 | 7245 | | Cytotoxin 8 OS=Naja naja OX=35670 PE=1 SV=2 |
| 3SA0_NAJSP | 44 | 9605 | 2.1 | Cytotoxin OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| ▶62 samesets of 3SA0_NAJSP | | | | |
| 3SA3A_NAJNA | 35 | 0 | 2.1, 2.2 | |
| ▶3 samesets of 3SA3A_NAJNA | | | | |

**Intact Mass:** Main peak at 6734.99Da and minor peak at 6904.22Da




## N.pal_i17_R2

### Cytotoxin 1 OS=Naja nivea OX=8655 PE=1 SV=1

**Database:** SwissProt
**Score:** 82
**Nominal mass ($M_r$):** 7148
**Calculated pI:** 9.00
**Taxonomy:** Naja nivea

Sequence similarity is available as an NCBI BLAST search of 3SA1_NAJNI against nr.

### Search parameters

**MS data file:** C:\Rachel\031019 Npal i17 R2 pm.mgf
**Enzyme:** Trypsin: cuts C-term side of KR unless next residue is P.
**Fixed modifications:** Carbamidomethyl (C)
**Variable modifications:** Oxidation (M), Phospho (ST), Phospho (Y)

### Protein sequence coverage: 43%

Matched peptides shown in ***bold red***.

```
  1 LKCHKLVPPV WKTCPEGKNL CYKMFMVSTS TVPVKRGCID VCPKDSALVK
 51 YVCCSTDKCN
```

Unformatted sequence string: 60 residues (for pasting into other applications).

Sort peptides by (•) Residue Number ( ) Increasing Mass ( ) Decreasing Mass

Show predicted peptides also

| Query | Start – End | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 507 | 13 – 23 | 457.2129 | 1368.6168 | 1368.6166 | 0.15 | 1 | 21 | 0.077 | 1 | U | K.TCPEGKNLCYK.M |
| 362 | 36 – 44 | 552.7686 | 1103.5226 | 1103.5216 | 0.93 | 1 | 32 | 0.065 | 1 | U | K.RGCIDVCPK.D |
| 363 | 36 – 44 | 552.7686 | 1103.5227 | 1103.5216 | 0.97 | 1 | 36 | 0.013 | 1 | U | K.RGCIDVCPK.D |
| 243 | 37 – 44 | 474.7175 | 947.4205 | 947.4205 | -0.023 | 0 | 40 | 0.032 | 1 | U | R.GCIDVCPK.D |
| 244 | 37 – 44 | 474.7181 | 947.4217 | 947.4205 | 1.28 | 0 | 29 | 0.021 | 1 | U | R.GCIDVCPK.D |
| 574 | 37 – 50 | 781.3922 | 1560.7699 | 1560.7640 | 3.78 | 1 | 67 | 6.8e-006 | 1 | U | R.GCIDVCPKDSALVK.Y |

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| ☑ 2.1 | 3SA1_NAJNI | 82 | 7148 | 6 (4) | 4 (3) | 5.18 | Cytotoxin 1 OS=Naja nivea OX=8655 PE=1 SV=1 |
| | ▼3 samesets of 3SA1_NAJNI | | | | | | |
| | 3SA2_NAJHA | 82 | 7309 | 6 (4) | 4 (3) | | Cytotoxin 2 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| | 3SA6_NAJHA | 82 | 7308 | 6 (4) | 4 (3) | | Cytotoxin 6 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| | 3SA9_NAJHA | 82 | 7120 | 6 (4) | 4 (3) | | Cytotoxin 9 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| ☑ 2.2 | IHH_CHICK | 34 | 45200 | 2 (2) | 2 (2) | 0.23 | Indian hedgehog protein OS=Gallus gallus OX=9031 GN=IHH PE=2 SV=1 |

Redisplay | All | None

▼8 peptide matches (6 non-duplicate, 2 duplicate)

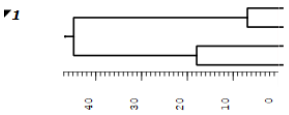
☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | 1 | 2 | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | | 409.2638 | 816.5131 | 816.5109 | 2.76 | 0 | 35 | 0.015 | ▶1 | U | | ■ | K.LIPLAYK.Q |
| 238 | | 473.3078 | 944.6011 | 944.6059 | -5.03 | 1 | 26 | 0.058 | ▶1 | U | | ■ | R.KLIPLAYK.Q |
| 243 | ▶1 | 474.7175 | 947.4205 | 947.4205 | -0.023 | 0 | 40 | 0.032 | ▶1 | U | ■ | | R.GCIDVCPK.D |
| 363 | ▶1 | 552.7686 | 1103.5227 | 1103.5216 | 0.97 | 1 | 36 | 0.013 | ▶1 | U | ■ | | K.RGCIDVCPK.D |
| 507 | | 457.2129 | 1368.6168 | 1368.6166 | 0.15 | 1 | 21 | 0.077 | ▶1 | U | ■ | | K.TCPEGKNLCYK.M |
| 574 | | 781.3922 | 1560.7699 | 1560.7640 | 3.78 | 1 | 67 | 6.8e-006 | ▶1 | U | ■ | | R.GCIDVCPKDSALVK.Y |

## *N.nct_i18r2:*

N.nct_i18_R2

## Peptide Mapping:




| | | | |
|---|---|---|---|
| 1 | 3SA4_NAJMO | 155 | Cytotoxin 4 OS=Naja mossambica OX=8644 PE=1 SV=1 |
| 4 | IHH_CHICK | 39 | Indian hedgehog protein OS=Gallus gallus OX=9031 GN=IHH PE=2 SV=1 |
| 2 | 3SA1_NAJPA | 118 | Cytotoxin 1 OS=Naja pallida OX=8658 PE=1 SV=1 |
| 3 | 3SA2_NAJHA | 93 | Cytotoxin 2 OS=Naja annulifera OX=96794 PE=1 SV=1 |

Threshold (0): 0 Cut

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| ☑ 1.1 | 3SA4_NAJMO | 155 | 7167 | 12 (8) | 5 (4) | 6.63 | Cytotoxin 4 OS=Naja mossambica OX=8644 PE=1 SV=1 |
| ☑ 1.2 | 3SA1_NAJPA | 118 | 7279 | 7 (5) | 4 (4) | 6.41 | Cytotoxin 1 OS=Naja pallida OX=8658 PE=1 SV=1 |
| | ▼1 sameset of 3SA1_NAJPA | | | | | | |
| | 3SA2_NAJMO | 118 | 7252 | 7 (5) | 4 (4) | | Cytotoxin 2 OS=Naja mossambica OX=8644 PE=1 SV=1 |
| ☑ 1.3 | 3SA2_NAJHA | 93 | 7309 | 6 (4) | 4 (4) | 6.41 | Cytotoxin 2 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| ☑ 1.4 | IHH_CHICK | 39 | 45200 | 5 (2) | 2 (2) | 0.19 | Indian hedgehog protein OS=Gallus gallus OX=9031 GN=IHH PE=2 SV=1 |

▼3 PA2B3_NAJMO 79 Basic phospholipase A2 CM-III OS=Naja mossambica OX=8644 PE=1 SV=1

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| 3.1 | PA2B3_NAJMO | 79 | 14091 | 3 (2) | 3 (2) | 0.72 | Basic phospholipase A2 CM-III OS=Naja mossambica OX=8644 PE=1 SV=1 |
| | ▼1 sameset of PA2B3_NAJMO | | | | | | |
| | PA2B4_NAJNG | 79 | 14057 | 3 (2) | 3 (2) | | Phospholipase A2 ""basic"" OS=Naja nigricollis OX=8654 PE=1 SV=1 |

▼3 peptide matches (3 non-duplicate, 0 duplicate)

☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 387 | | 494.2405 | 986.4665 | 986.4669 | -0.38 | 0 | 43 | 0.00091 | ▶1 | U | K.GTPVDDLDR.C |
| 767 | | 687.8357 | 1373.6569 | 1373.6615 | -3.35 | 0 | 5 | 0.82 | ▶1 | U | R.YIDANYNINFK.K |
| 948 | | 719.6377 | 2155.8912 | 2155.9020 | -4.98 | 0 | 56 | 5.3e-006 | ▶1 | U | K.CGAAVCNCDLVAANCFAGAR.Y |

▼5 3NOJ6_DENJA 44 Toxin S6C6 OS=Dendroaspis jamesoni kaimosae OX=8619 PE=1 SV=1

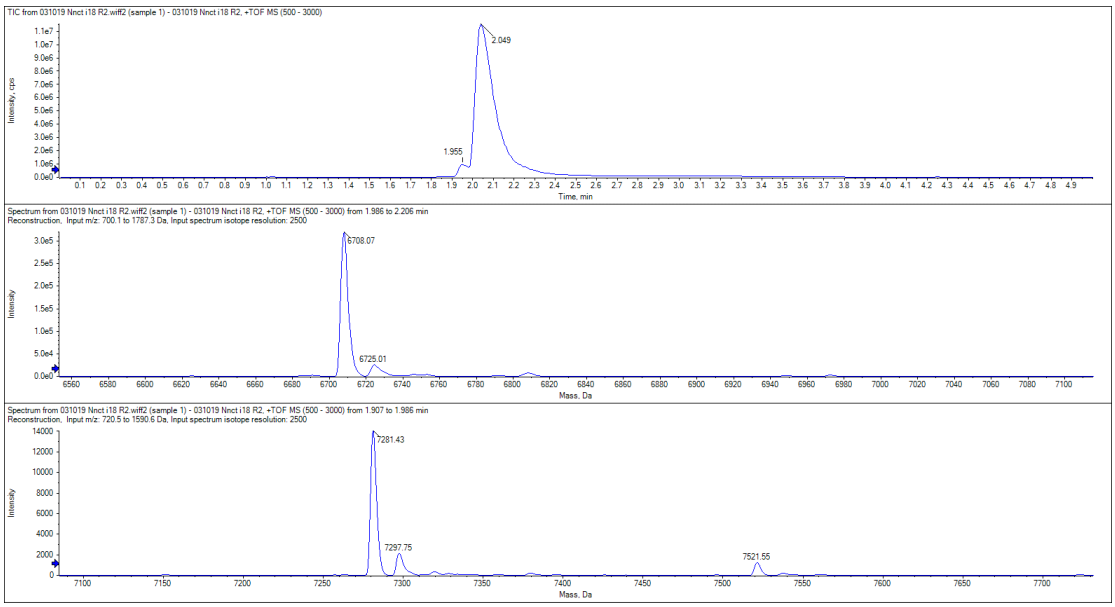
| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| 5.1 | 3NOJ6_DENJA | 44 | 7963 | 2 (1) | 1 (1) | 0.59 | Toxin S6C6 OS=Dendroaspis jamesoni kaimosae OX=8619 PE=1 SV=1 |

▼2 peptide matches (1 non-duplicate, 1 duplicate)

☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 573 | ▶1 | 578.2331 | 1154.4516 | 1154.4519 | -0.23 | 0 | 44 | 0.00041 | ▶1 | U | K.CCTTNLCNI.- |

**Intact Mass:** Main peak at 6708.07Da, small peak at 7281.43Da and minor peak at 7521.55Da




N.nct_i18_R2

## Cytotoxin 4 OS=Naja mossambica OX=8644 PE=1 SV=1

**Database:** SwissProt
**Score:** 155
**Nominal mass ($M_r$):** 7167
**Calculated pI:** 9.48
**Taxonomy:** Naja mossambica

Sequence similarity is available as an NCBI BLAST search of 3SA4_NAJMO against nr.

### Search parameters

**MS data file:** C:\Rachel\031019 Nnct i18 R2 pm.mgf
**Enzyme:** Trypsin: cuts C-term side of KR unless next residue is P.
**Fixed modifications:** Carbamidomethyl (C)
**Variable modifications:** Oxidation (M), Phospho (ST), Phospho (Y)

### Protein sequence coverage: 40%

Matched peptides shown in ***bold red***.

```
  1 LKCNKLIPIA YKTCPEGKNL CYKMMLASKK MVPVKRGCIN VCPKNSALVK
 51 YVCCSTDRCN
```

Unformatted sequence string: 60 residues (for pasting into other applications).

Sort peptides by ◉ Residue Number ○ Increasing Mass ○ Decreasing Mass

Show predicted peptides also

| Query | Start – End | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 152 | 6 – 12 | 409.2621 | 816.5096 | 816.5109 | -1.63 | 0 | 34 | 0.019 | 1 | U | K.LIPIAYK.T |
| 153 | 6 – 12 | 409.2627 | 816.5108 | 816.5109 | -0.15 | 0 | 22 | 0.25 | 1 | U | K.LIPIAYK.T |
| 154 | 6 – 12 | 409.2639 | 816.5133 | 816.5109 | 2.97 | 0 | 24 | 0.18 | 1 | U | K.LIPIAYK.T |
| 138 | 24 – 30 | 404.7250 | 807.4355 | 807.4346 | 1.04 | 1 | 17 | 0.038 | 1 | U | K.MMLASKK.M |
| 468 | 51 – 58 | 530.7139 | 1059.4133 | 1059.4114 | 1.82 | 0 | 39 | 0.0045 | 1 | U | K.YVCCSTDR.C |
| 469 | 51 – 58 | 530.7148 | 1059.4150 | 1059.4114 | 3.38 | 0 | 33 | 0.0039 | 1 | U | K.YVCCSTDR.C |
| 734 | 51 – 60 | 667.7491 | 1333.4836 | 1333.4850 | -1.04 | 1 | 64 | 8.5e-005 | 1 | U | K.YVCCSTDRCN.- |
| 735 | 51 – 60 | 667.7498 | 1333.4851 | 1333.4850 | 0.11 | 1 | 62 | 0.00012 | 1 | U | K.YVCCSTDRCN.- |
| 736 | 51 – 60 | 445.5029 | 1333.4869 | 1333.4850 | 1.43 | 1 | 37 | 0.00033 | 1 | U | K.YVCCSTDRCN.- |
| 737 | 51 – 60 | 667.7508 | 1333.4870 | 1333.4850 | 1.54 | 1 | 60 | 0.00024 | 1 | U | K.YVCCSTDRCN.- |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ▼3 | | PA2B3_NAJMO | | | 79 | Basic phospholipase A2 CM-III OS=Naja mossambica OX=8644 PE=1 SV=1 | |
| | | Score | Mass | Matches | Sequences | emPAI | |
| 3.1 | PA2B3_NAJMO | 79 | 14091 | 2 (2) | 2 (2) | 0.72 | Basic phospholipase A2 CM-III OS=Naja mossambica OX=8644 PE=1 SV=1 |
| | ▼1 sameset of PA2B3_NAJMO | | | | | | |
| | PA2B4_NAJNG | 79 | 14057 | 2 (2) | 2 (2) | | Phospholipase A2 ""basic"" OS=Naja nigricollis OX=8654 PE=1 SV=1 |

▼2 peptide matches (2 non-duplicate, 0 duplicate)

☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 387 | | 494.2405 | 986.4665 | 986.4669 | -0.38 | 0 | 43 | 0.00091 | ▶1 | U | K.GTFVDDLDR.C |
| 948 | | 719.6377 | 2155.8912 | 2155.9020 | -4.98 | 0 | 56 | 5.3e-006 | ▶1 | U | K.CGAAVCNCDLVAANCFAGAR.Y |

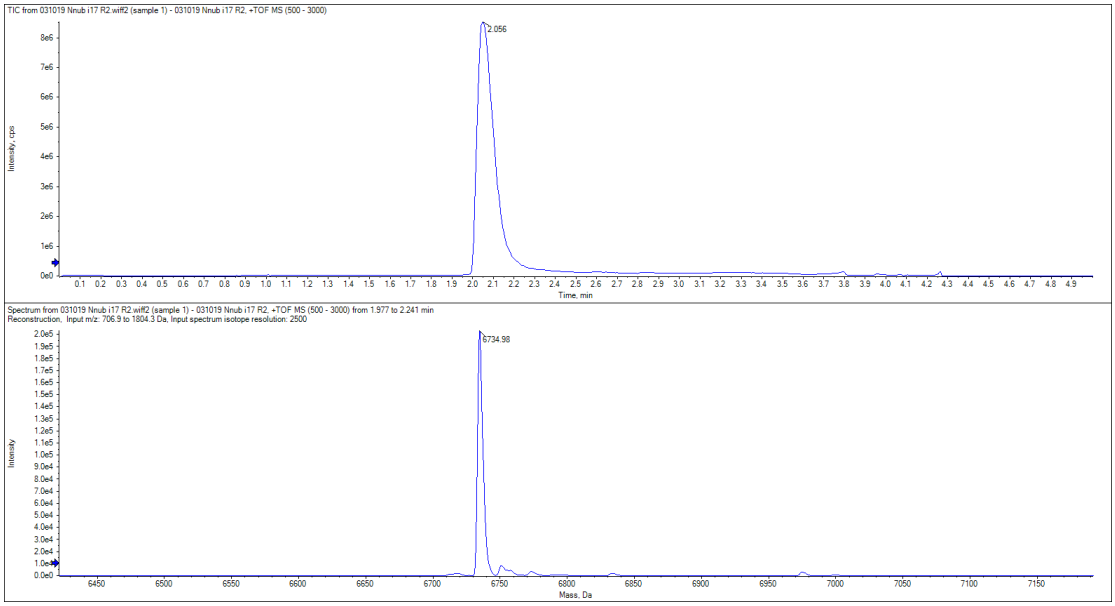
## *N.nub_i17r2:*

## N.nub_i17_R2

## Peptide Mapping:

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| ☑ 2.1 | 3SA1_NAJNI | 76 | 7148 | 6 (4) | 4 (3) | 5.08 | Cytotoxin 1 OS=Naja nivea OX=8655 PE=1 SV=1 |
| | ▼3 samesets of 3SA1_NAJNI | | | | | | |
| | 3SA2_NAJHA | 76 | 7309 | 6 (4) | 4 (3) | | Cytotoxin 2 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| | 3SA6_NAJHA | 76 | 7308 | 6 (4) | 4 (3) | | Cytotoxin 6 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| | 3SA9_NAJHA | 76 | 7120 | 6 (4) | 4 (3) | | Cytotoxin 9 OS=Naja annulifera OX=96794 PE=1 SV=1 |
| ☑ 2.2 | 3SA0_NAJSP | 56 | 9605 | 5 (4) | 3 (3) | 2.88 | Cytotoxin OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| | ▼25 samesets of 3SA0_NAJSP | | | | | | |
| | 3SA1_NAJKA | 56 | 7259 | 5 (4) | 3 (3) | | Cytotoxin 1 OS=Naja kaouthia OX=8649 PE=1 SV=1 |
| | 3SA1_NAJSP | 56 | 9527 | 5 (4) | 3 (3) | | Cytotoxin 1 OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| | 3SA2A_NAJSP | 56 | 9561 | 5 (4) | 3 (3) | | Cytotoxin 2a OS=Naja sputatrix OX=33626 PE=2 SV=1 |
| | 3SA2B_NAJSP | 56 | 9549 | 5 (4) | 3 (3) | | Cytotoxin 2b OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| | 3SA2_NAJAT | 56 | 9548 | 5 (4) | 3 (3) | | Cytotoxin 2 OS=Naja atra OX=8656 PE=1 SV=2 |
| | 3SA2_NAJKA | 56 | 9512 | 5 (4) | 3 (3) | | Cytotoxin 2 OS=Naja kaouthia OX=8649 PE=3 SV=1 |
| | 3SA2_NAJNA | 56 | 7215 | 5 (4) | 3 (3) | | Cytotoxin 2 OS=Naja naja OX=35670 PE=1 SV=1 |
| | 3SA3A_NAJAT | 56 | 9572 | 5 (4) | 3 (3) | | Cytotoxin 3a OS=Naja atra OX=8656 PE=3 SV=1 |
| | 3SA3B_NAJAT | 56 | 9530 | 5 (4) | 3 (3) | | Cytotoxin 3b OS=Naja atra OX=8656 PE=3 SV=2 |
| | 3SA3_NAJAT | 56 | 9545 | 5 (4) | 3 (3) | | Cytotoxin 3 OS=Naja atra OX=8656 PE=1 SV=1 |
| | 3SA3_NAJSP | 56 | 9545 | 5 (4) | 3 (3) | | Cytotoxin 3 OS=Naja sputatrix OX=33626 PE=1 SV=1 |
| | 3SA4A_NAJSP | 56 | 9575 | 5 (4) | 3 (3) | | Cytotoxin 4a OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| | 3SA4B_NAJSP | 56 | 9591 | 5 (4) | 3 (3) | | Cytotoxin 4b OS=Naja sputatrix OX=33626 PE=1 SV=2 |
| | 3SA4N_NAJAT | 56 | 9606 | 5 (4) | 3 (3) | | Cytotoxin 4N OS=Naja atra OX=8656 PE=3 SV=1 |
| | 3SA4_NAJAT | 56 | 9591 | 5 (4) | 3 (3) | | Cytotoxin 4 OS=Naja atra OX=8656 PE=1 SV=2 |
| | 3SA4_NAJKA | 56 | 9545 | 5 (4) | 3 (3) | | Cytotoxin 4 OS=Naja kaouthia OX=8649 PE=1 SV=1 |
| | 3SA5A_NAJSP | 56 | 9548 | 5 (4) | 3 (3) | | Cytotoxin 5a OS=Naja sputatrix OX=33626 PE=1 SV=2 |
| | 3SA5B_NAJSP | 56 | 9562 | 5 (4) | 3 (3) | | Cytotoxin 5b OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| | 3SA5_NAJAT | 56 | 9593 | 5 (4) | 3 (3) | | Cytotoxin 5 OS=Naja atra OX=8656 PE=3 SV=1 |
| | 3SA6_NAJSP | 56 | 7481 | 5 (4) | 3 (3) | | Cytotoxin 6 (Fragment) OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| | 3SA7_NAJSP | 56 | 7513 | 5 (4) | 3 (3) | | Cytotoxin 7 (Fragment) OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| | 3SA9_NAJNA | 56 | 8548 | 5 (4) | 3 (3) | | Cytotoxin 9 (Fragment) OS=Naja naja OX=35670 PE=1 SV=1 |
| | 3SAA_NAJNA | 56 | 7216 | 5 (4) | 3 (3) | | Cytotoxin 10 OS=Naja naja OX=35670 PE=1 SV=2 |
| | 3SAC3_NAJSP | 56 | 7204 | 5 (4) | 3 (3) | | Cytotoxin KJC3 OS=Naja sputatrix OX=33626 PE=1 SV=1 |
| | 3SAT_NAJAT | 56 | 7262 | 5 (4) | 3 (3) | | Cytotoxin 5 OS=Naja atra OX=8656 PE=1 SV=1 |
| ☑ 2.3 | IHH_CHICK | 40 | 45200 | 2 (2) | 2 (2) | 0.23 | Indian hedgehog protein OS=Gallus gallus OX=9031 GN=IHH PE=2 SV=1 |

**Intact Mass:** Main peak at 6734.98Da



N.nub_i17_R2

### Cytotoxin 1 OS=Naja nivea OX=8655 PE=1 SV=1

**Database:** SwissProt
**Score:** 76
**Nominal mass ($M_r$):** 7148
**Calculated pI:** 9.00
**Taxonomy:** Naja nivea

Sequence similarity is available as an NCBI BLAST search of 3SA1_NAJNI against nr.

### Search parameters

**MS data file:** C:\Rachel\031019 Nnub i17 R2 pm.mgf
**Enzyme:** Trypsin: cuts C-term side of KR unless next residue is P.
**Fixed modifications:** Carbamidomethyl (C)
**Variable modifications:** Oxidation (M), Phospho (ST), Phospho (Y)

### Protein sequence coverage: 43%

Matched peptides shown in ***bold red***.

```
  1 LKCHKLVPPV WKTCPEGKNL CYKMFMVSTS TVPVKRGCID VCPKDSALVK
 51 YVCCSTDKCN
```

Unformatted sequence string: 60 residues (for pasting into other applications).

Sort peptides by ◉ Residue Number ○ Increasing Mass ○ Decreasing Mass

Show predicted peptides also

| Query | Start – End | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 403 | 13 – 23 | 457.2111 | 1368.6115 | 1368.6166 | -3.74 | 1 | 11 | 0.43 | 1 | U | K.TCPEGKNLCYK.M |
| 294 | 36 – 44 | 552.7664 | 1103.5182 | 1103.5216 | -3.06 | 1 | 29 | 0.028 | 1 | | K.RGCIDVCPK.D |
| 295 | 36 – 44 | 552.7681 | 1103.5216 | 1103.5216 | 0.035 | 1 | 23 | 0.19 | 1 | | K.RGCIDVCPK.D |
| 190 | 37 – 44 | 474.7168 | 947.4190 | 947.4205 | -1.53 | 0 | 44 | 0.0018 | 1 | | R.GCIDVCPK.D |
| 191 | 37 – 44 | 474.7181 | 947.4217 | 947.4205 | 1.32 | 0 | 32 | 0.023 | 1 | | R.GCIDVCPK.D |
| 454 | 37 – 50 | 521.2609 | 1560.7609 | 1560.7640 | -1.96 | 1 | 52 | 0.0001 | 1 | U | R.GCIDVCPKDSALVK.Y |

| | | Score | Mass | Matches | Sequences | emPAI | |
|---|---|---|---|---|---|---|---|
| ☑ 2.1 | 3SA1_NAJNI | 76 | 7148 | 6 (4) | 4 (3) | 5.08 | Cytotoxin 1 OS=Naja nivea OX=8655 PE=1 SV=1 |
| | ▶ 3 samesets of 3SA1_NAJNI | | | | | | |
| ☑ 2.2 | 3SA0_NAJSP | 56 | 9605 | 5 (4) | 3 (3) | 2.88 | Cytotoxin OS=Naja sputatrix OX=33626 PE=3 SV=1 |
| | ▶ 25 samesets of 3SA0_NAJSP | | | | | | |
| ☑ 2.3 | IHH_CHICK | 40 | 45200 | 2 (2) | 2 (2) | 0.23 | Indian hedgehog protein OS=Gallus gallus OX=9031 GN=IHH PE=2 SV=1 |

Redisplay | All | None

▼ *9 peptide matches (7 non-duplicate, 2 duplicate)*

☑ Auto-fit to window

| Query | Dupes | Observed | Mr(expt) | Mr(calc) | ppm | M | Score | Expect | Rank | U | 1 | 2 | 3 | Peptide |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 74 | | 409.2648 | 816.5150 | 816.5109 | 5.01 | 0 | 36 | 0.013 | ▶1 | U | | | ■ | K.LIPLAYK.Q |
| 130 | | 440.2714 | 878.5282 | 878.5266 | 1.88 | 0 | 37 | 0.014 | ▶1 | U | | ■ | | K.LVPLFYK.T |
| 187 | | 473.3098 | 944.6050 | 944.6059 | -0.90 | 1 | 32 | 0.014 | ▶1 | U | | | ■ | R.KLIPLAYK.Q |
| 190 | ▶1 | 474.7168 | 947.4190 | 947.4205 | -1.53 | 0 | 44 | 0.0018 | ▶1 | | ■ | ■ | | R.GCIDVCPK |
| 294 | ▶1 | 552.7664 | 1103.5182 | 1103.5216 | -3.06 | 1 | 29 | 0.028 | ▶1 | | ■ | ■ | | K.RGCIDVCPK |
| 403 | | 457.2111 | 1368.6115 | 1368.6166 | -3.74 | 1 | 11 | 0.43 | ▶1 | U | ■ | | | K.TCPEGKNLCYK.M |
| 454 | | 521.2609 | 1560.7609 | 1560.7640 | -1.96 | 1 | 52 | 0.0001 | ▶1 | U | ■ | | | R.GCIDVCPKDSALVK.Y |

## Appendix V: Amino Acid Sequence Replacement and SAR including 3SOFB

Amino acid matrix to predict the unknown amino acids present in the N.pal_i15r4 sequence following partial mass spec identification. Amino acids giving a combined molecular weight of between 230-234 are highlighted in green. The exact mass required to give a final mw of 6818.28 Da is 232.2 Da.

| Amino Acid | Molecular Weight (g/mol) | Ala | Arg | Asn | Asp | Cys | Glu | Gln | Gly | His | Ile | Leu | Lys | Met | Phe | Pro | Ser | Thr | Trp | Tyr | Val |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 89.1 | 174.2 | 132.1 | 133.1 | 121.2 | 147.1 | 146.2 | 75.1 | 155.2 | 131.2 | 131.2 | 146.2 | 149.2 | 165.2 | 115.1 | 105.1 | 119.1 | 204.2 | 181.2 | 117.1 |
| Ala | 89.1 | 178.2 | 263.3 | 221.2 | 222.2 | 210.3 | 236.2 | 235.3 | 164.2 | 244.3 | 220.3 | 220.3 | 235.3 | 238.3 | 254.3 | 204.2 | 194.2 | 208.2 | 293.3 | 270.3 | 206.2 |
| Arg | 174.2 | 263.3 | 348.4 | 306.3 | 307.3 | 295.4 | 321.3 | 320.4 | 249.3 | 329.4 | 305.4 | 305.4 | 320.4 | 323.4 | 339.4 | 289.3 | 279.3 | 293.3 | 378.4 | 355.4 | 291.3 |
| Asn | 132.1 | 221.2 | 306.3 | 264.2 | 265.2 | 253.3 | 279.2 | 278.3 | 207.2 | 287.3 | 263.3 | 263.3 | 278.3 | 281.3 | 297.3 | 247.2 | 237.2 | 251.2 | 336.3 | 313.3 | 249.2 |
| Asp | 133.1 | 222.2 | 307.3 | 265.2 | 266.2 | 254.3 | 280.2 | 279.3 | 208.2 | 288.3 | 264.3 | 264.3 | 279.3 | 282.3 | 298.3 | 248.2 | 238.2 | 252.2 | 337.3 | 314.3 | 250.2 |
| Cys | 121.2 | 210.3 | 295.4 | 253.3 | 254.3 | 242.4 | 268.3 | 267.4 | 196.3 | 276.4 | 252.4 | 252.4 | 267.4 | 270.4 | 286.4 | 236.3 | 226.3 | 240.3 | 325.4 | 302.4 | 238.3 |
| Glu | 147.1 | 236.2 | 321.3 | 279.2 | 280.2 | 268.3 | 294.2 | 293.3 | 222.2 | 302.3 | 278.3 | 278.3 | 293.3 | 296.3 | 312.3 | 262.2 | 252.2 | 266.2 | 351.3 | 328.3 | 264.2 |
| Gln | 146.2 | 235.3 | 320.4 | 278.3 | 279.3 | 267.4 | 293.3 | 292.4 | 221.3 | 301.4 | 277.4 | 277.4 | 292.4 | 295.4 | 311.4 | 261.3 | 251.3 | 265.3 | 350.4 | 327.4 | 263.3 |
| Gly | 75.1 | 164.2 | 249.3 | 207.2 | 208.2 | 196.3 | 222.2 | 221.3 | 150.2 | 230.3 | 206.3 | 206.3 | 221.3 | 224.3 | 240.3 | 190.2 | 180.2 | 194.2 | 279.3 | 256.3 | 192.2 |
| His | 155.2 | 244.3 | 329.4 | 287.3 | 288.3 | 276.4 | 302.3 | 301.4 | 230.3 | 310.4 | 286.4 | 286.4 | 301.4 | 304.4 | 320.4 | 270.3 | 260.3 | 274.3 | 359.4 | 336.4 | 272.3 |
| Ile | 131.2 | 220.3 | 305.4 | 263.3 | 264.3 | 252.4 | 278.3 | 277.4 | 206.3 | 286.4 | 262.4 | 262.4 | 277.4 | 280.4 | 296.4 | 246.3 | 236.3 | 250.3 | 335.4 | 312.4 | 248.3 |
| Leu | 131.2 | 220.3 | 305.4 | 263.3 | 264.3 | 252.4 | 278.3 | 277.4 | 206.3 | 286.4 | 262.4 | 262.4 | 277.4 | 280.4 | 296.4 | 246.3 | 236.3 | 250.3 | 335.4 | 312.4 | 248.3 |
| Lys | 146.2 | 235.3 | 320.4 | 278.3 | 279.3 | 267.4 | 293.3 | 292.4 | 221.3 | 301.4 | 277.4 | 277.4 | 292.4 | 295.4 | 311.4 | 261.3 | 251.3 | 265.3 | 350.4 | 327.4 | 263.3 |
| Met | 149.2 | 238.3 | 323.4 | 281.3 | 282.3 | 270.4 | 296.3 | 295.4 | 224.3 | 304.4 | 280.4 | 280.4 | 295.4 | 298.4 | 314.4 | 264.3 | 254.3 | 268.3 | 353.4 | 330.4 | 266.3 |
| Phe | 165.2 | 254.3 | 339.4 | 297.3 | 298.3 | 286.4 | 312.3 | 311.4 | 240.3 | 320.4 | 296.4 | 296.4 | 311.4 | 314.4 | 330.4 | 280.3 | 270.3 | 284.3 | 369.4 | 346.4 | 282.3 |
| Pro | 115.1 | 204.2 | 289.3 | 247.2 | 248.2 | 236.3 | 262.2 | 261.3 | 190.2 | 270.3 | 246.3 | 246.3 | 261.3 | 264.3 | 280.3 | 230.2 | 220.2 | 234.2 | 319.3 | 296.3 | 232.2 |
| Ser | 105.1 | 194.2 | 279.3 | 237.2 | 238.2 | 226.3 | 252.2 | 251.3 | 180.2 | 260.3 | 236.3 | 236.3 | 251.3 | 254.3 | 270.3 | 220.2 | 210.2 | 224.2 | 309.3 | 286.3 | 222.2 |
| Thr | 119.1 | 208.2 | 293.3 | 251.2 | 252.2 | 240.3 | 266.2 | 265.3 | 194.2 | 274.3 | 250.3 | 250.3 | 265.3 | 268.3 | 284.3 | 234.2 | 224.2 | 238.2 | 323.3 | 300.3 | 236.2 |
| Trp | 204.2 | 293.3 | 378.4 | 336.3 | 337.3 | 325.4 | 351.3 | 350.4 | 279.3 | 359.4 | 335.4 | 335.4 | 350.4 | 353.4 | 369.4 | 319.3 | 309.3 | 323.3 | 408.4 | 385.4 | 321.3 |
| Tyr | 181.2 | 270.3 | 355.4 | 313.3 | 314.3 | 302.4 | 328.3 | 327.4 | 256.3 | 336.4 | 312.4 | 312.4 | 327.4 | 330.4 | 346.4 | 296.3 | 286.3 | 300.3 | 385.4 | 362.4 | 298.3 |
| Val | 117.1 | 206.2 | 291.3 | 249.2 | 250.2 | 238.3 | 264.2 | 263.3 | 192.2 | 272.3 | 248.3 | 248.3 | 263.3 | 266.3 | 282.3 | 232.2 | 222.2 | 236.2 | 321.3 | 298.3 | 234.2 |

SAR table including 3SOFB from the UniProt database

| Code | Cytotoxin | Species | LD50 | 1 | 2 | 3 | | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3SA1 | 1 | *Naja mossambica* | 0.83 | L | K | C | - | N | Q | L | I | P | P | F | W | K | T | C | P | K | G | K | N | L | C | Y | K | M | T | M | R | A | - | A | P | M | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | I | K | Y | M | C | C | N | T | N | K | C | N |
| 3SA2 | 2 | *Naja mossambica* | 1.11 | L | K | C | - | N | Q | L | I | P | P | F | W | K | T | C | P | K | G | K | N | L | C | Y | K | M | T | M | R | G | - | A | S | K | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | I | K | Y | M | C | C | N | T | D | K | C | N |
| 3SA2 | 2 | *Naja kaouthia* | 1.2 | L | K | C | - | N | K | L | I | P | L | A | Y | K | T | C | P | A | G | K | N | L | C | Y | K | M | F | M | V | S | - | N | K | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | L | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA1 | 1 | *Maja melanoleuca* | 1.37 | L | E | C | - | N | K | L | V | P | I | A | H | K | T | C | P | A | G | K | N | L | C | Y | Q | M | Y | M | V | S | - | K | S | T | I | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA2 | 2 | *Naja nivea* | 1.5 | L | K | C | - | H | Q | L | I | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | - | T | P | M | I | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | M | C | C | N | T | D | K | C | N |
| 3SA3 | 3 | *Naja nivea* | 1.6 | L | K | C | - | N | Q | L | I | P | P | F | W | K | T | C | P | K | G | K | N | L | C | Y | N | M | Y | M | V | S | - | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA3 | 3 | *Naja mossambica* | 1.82 | L | K | C | - | N | R | L | I | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | T | M | R | L | - | A | P | K | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | L | L | I | K | Y | M | C | C | N | T | N | K | C | N |
| 3SA5 | 5 | *Naja annulifera* | 1.82 | L | K | C | - | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | - | T | P | M | I | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | M | C | C | N | T | N | K | C | N |
| 3SA4 | 4 | *Naja mossambica* | 1.97 | L | K | C | - | N | K | L | I | P | I | A | Y | K | T | C | P | E | G | K | N | L | C | Y | K | M | M | L | A | S | - | K | K | M | V | P | V | K | R | G | C | I | N | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | S | T | D | R | C | N |
| 3SA2 | 2 | *Naja annulifera* | 1.98 | L | K | C | - | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | - | T | P | M | L | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | M | C | C | N | T | D | K | C | N |
| 3SA2 | 2 | *Naja atra* | 2.1 | L | K | C | - | N | K | L | V | P | L | F | Y | K | T | C | P | A | G | K | N | L | C | Y | K | M | F | M | V | S | - | N | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA4 | 4 | *Naja atra* | 2.1 | R | K | C | - | N | K | L | V | P | L | F | Y | K | T | C | P | A | G | K | N | L | C | Y | K | M | F | M | V | S | - | N | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA1 | 1 | *Naja kaouthia* | 2.5 | L | K | C | - | N | K | L | V | P | L | F | Y | K | T | C | P | A | G | K | N | L | C | Y | K | M | F | M | V | S | - | N | K | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | L | V | L | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA5 | 5 | *Naja haje* | 2.6 | L | K | C | - | H | Q | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | S | - | S | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | K | C | N |
| 3SA8 | 8 | *Naja annulifera* | 2.6 | L | K | C | - | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | S | - | T | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | N | K | C | N |
| 3SA7 | 7 | *Naja annulifera* | 2.61 | L | K | C | - | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | - | T | P | M | L | P | V | K | R | G | C | I | N | V | C | P | K | D | S | A | L | V | K | Y | M | C | C | N | T | N | K | C | N |
| 3SA5 | 5 | *Naja mossambica* | 2.85 | L | K | C | - | K | K | L | I | P | L | F | S | K | T | C | P | E | G | K | N | L | C | Y | K | M | T | M | R | L | - | A | P | K | V | P | V | K | R | G | C | I | D | V | C | P | K | S | S | F | L | V | K | Y | E | C | C | D | T | D | R | C | N |
| 3SA1 | 1 | *Naja nivea* | 2.9 | L | K | C | - | H | K | L | V | P | P | V | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | F | M | V | S | - | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | V | C | C | S | T | D | K | C | N |
| 3SA3 | 3 | *Naja annulifera* | 2.98 | L | K | C | - | Y | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | S | - | T | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | K | C | N |
| 3SA1 | 1 | *Naja annulifera* | 3 | L | K | C | - | H | K | L | V | P | P | V | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | F | M | V | S | - | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | S | T | D | K | C | N |
| 3SA4 | 4 | *Naja annulifera* | 4 | L | K | C | - | N | K | L | I | P | P | F | W | K | T | C | P | K | G | K | N | L | C | Y | K | M | Y | M | V | S | - | T | L | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | N | K | C | N |
| 3SA10 | 10 | *Naja annulifera* | 4.14 | L | E | C | - | N | Q | L | I | P | I | A | H | K | T | C | P | E | G | K | N | L | C | Y | K | M | F | M | V | S | - | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | N | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA9 | 9 | *Naja annulifera* | 4.9 | L | E | C | - | N | K | L | V | P | I | A | H | K | T | C | P | E | G | K | N | L | C | Y | K | M | F | M | V | S | - | T | S | T | V | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | V | C | C | N | T | D | R | C | N |
| 3SA6 | 6 | *Naja annulifera* | 5.09 | L | K | C | - | H | K | L | V | P | P | F | W | K | T | C | P | E | G | K | N | L | C | Y | K | M | Y | M | V | A | - | T | P | M | L | P | V | K | R | G | C | I | D | V | C | P | K | D | S | A | L | V | K | Y | M | C | C | N | T | N | K | C | N |
| 3SOFB | | *Naja annulifera* | 25.2 | L | K | C | H | N | T | Q | L | P | F | I | Y | K | T | C | P | E | G | K | N | L | C | F | K | T | T | L | K | K | L | P | L | K | I | P | I | K | R | G | C | A | A | T | C | P | K | S | S | A | L | L | K | V | V | C | C | S | T | D | K | C | N |

| Key: | high | 0.8-1 | medium | 0.6-0.8 | low | 0.4-0.6 | very low | 0.2-0.4 | none | 0-0.2 |
|---|---|---|---|---|---|---|---|---|---|---|