Please cite this publication as follows:

Link to official URL (if available):

Contact: create.library@canterbury.ac.uk

BMC Genomics

Research article

Genomic structure and cloning of two transcript isoforms of human Sp8

Maria-athina Milona[1], Julie E Gough[2] and Alasdair J Edgar*[3]

Address: [1]Department of Cell Biology and Genetics, Faculty of Medicine, Erasmus Medical Center Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands, [2]Manchester Materials Science Centre, University of Manchester and UMIST, Grosvenor St., Manchester, M1 7HS, United Kingdom and [3]Department of Craniofacial Development, GKT Dental Institute, King's College, Guy's Hospital, London, SE1 9RT, United Kingdom

Email: Maria-athina Milona - m.mylona@erasmusmc.nl; Julie E Gough - j.gough@umist.ac.uk; Alasdair J Edgar* - alasdair.edgar@kcl.ac.uk

* Corresponding author

This article is available from: http://www.biomedcentral.com/1471-2164/5/86

Abstract

Background: The Specificity proteins (Sp) are a family of transcription factors that have three highly conserved zinc-fingers located towards the carboxy-terminal that bind GC-boxes and assist in the initiation of gene transcription. Human Sp1-7 genes have been characterized. Recently, the phenotype of Sp8 null mice has been described, being tailless and having severe truncation of both fore and hind limbs. They also have malformed brains with defective closure of the anterior and posterior neuropore during brain development.

Results: The human Sp8 gene is a three-exon gene that maps to 7p21.3, close to the related Sp4 gene. From an osteosarcoma cell line we cloned two transcript variants that use two different first exons and have a common second exon. One clone encodes a 508-residue protein, Sp8L (isoform 1) and the other a shorter 490-residue protein, Sp8S (isoform 2). These two isoforms are conserved being found also in mice and zebrafish. Analysis of the Sp8L protein sequence reveals an amino-terminal hydrophobic Sp-motif that is disrupted in Sp8S, a buttonhead box and three C_2H_2 zinc-fingers. Sp8 mRNA expression was detected in a wide range of tissues at a low level, with the highest levels being found in brain. Treatment of the murine pluripotent cell line C3H10T1/2 with 100 ng/mL BMP-2 induced Sp8 mRNA after 24 hours.

Conclusions: There is conservation of the two Sp8 protein isoforms between primates, rodents and fish, suggesting that the isoforms have differing roles in gene regulation. Sp8 may play a role in chondrogenic/osteoblastic differentiation in addition to its role in brain and limb development.

Background

We are interested in understanding how Specificity proteins (Sp) govern extracellular matrix deposition during bone formation. The Sp proteins are a family of transcription factors with three zinc-fingers that bind GC-boxes and assist the further binding of the multiprotein complex TFIID promoting the initiation of gene transcription [1-4].

GC-boxes have the consensus sequence GGGCGGG. Sp1 and Sp3 are associated with chondrocytic differentiation by regulating the alpha1(II) procollagen gene (COL2A1), a chondrocytic marker. Sp1 has been shown to be an activator and Sp3 a repressor of COL2A1 transcription [5]. Despite the ubiquitous expression of Sp3, Sp3 null mice suffer from specific skeletal defects associated with a delay

in osteoblast differentiation, having reduced ossification and impaired bone formation and show a decrease in the expression of osteocalcin, an osteoblast specific gene [6]. Sp7, also known as Osterix, is required for osteoblast differentiation since in Sp7 null mice bone formation does not occur [7]. Previously, we cloned two isoforms of human Sp7 that are generated from distinct promoters upstream of exon 1a and exon 1b [8]. We identified another human Sp gene that is closely related to Sp7 and named it Sp8. Recently, Sp8 null mice have been described [9,10]. Note that the mouse mBtd gene [10] is Sp8 and not a specific homologue of the *Drosophila* gene buttonhead (Btd) [11]. However, Btd is a member of the Sp family [12]. Sp8 null mice have a well-defined phenotype, having severe truncation of all limbs and no tail. Their malformed brains have defective closure of the anterior and posterior neuropores that leads to exencephaly and spina bifida. Sp8 is required for proper maintenance and maturation of the apical ectodermal ridge, a signalling centre, which forms at the limb bud apex and governs the outgrowth of the limb during development. Here we describe the sequence of two isoforms of human Sp8.

Results
In our search for additional Sp transcription factors, we performed a TblastN search of the human genome with the sequence of Sp7 [8]. The most closely related sequences were found on genomic DNA BAC clones CTA-324D18 and CTB-86D3 that were in the draft stage (GenBank accession Nos. AC005251 and AC005060 Genome Sequencing Center, Washington University, USA). This new gene was named Sp8 since it is closely related to Sp7.

Human Sp8 cDNAs
Utilising the sequences of the two BAC clones, PCR primers were designed to amplify probable cDNAs. Two different cDNA clones were obtained by RT-PCR from the human MG-63 osteosarcoma cell line and sequenced (Fig. 1). Successful amplification and sequencing required methods designed to overcome problematic GC-rich regions. These transcript variants differed at their 5' termini. Clone 1 comprises exon 1a and 2, and together they encode the long protein isoform of Sp8 (Sp8L)(Fig. 1)(GenBank accession No. AY167047). Clone 2 comprises the untranslated exon 1b and exon 2. Exon 1b contains an in-frame stop codon and no initiation methionine codon. Together they encode the short protein isoform (Sp8S)(Fig. 1)(GenBank accession No. AY167048). The first ATG codon of Sp8L is in excellent sequence consensus for an initiation methionine having a guanosine residue at positions -3 and +4, and that of Sp8S is in good sequence consensus having a guanosine at -3 [13]. The ORF is GC rich (70%). Another Sp8 cDNA clone was isolated from the osteoblasts of a patient with osteoporosis. This has a glycine-165 deletion; having lost one

of the five sequential GGC codons (nucleotides 569–583) (Fig. 1).

Human Sp8 gene
After the completion of the human genome the location and structure of the human Sp8 gene was identified by mapping the sequences of Sp8L and Sp8S cDNA clones. The human Sp8 gene maps to 7p21.3 between the multidrug resistance protein (P-glycoprotein) and ribosomal protein L23 genes and is separated by 0.65 Mbp from the related Sp4 gene by seven genes encoding two ribosomal proteins, four putative genes of unknown function and a gene similar to argininosuccinate synthetase. The Sp8 and Sp4 genes are linked to the homeobox gene cluster HOX A, which is 5.6 Mbp away. The human Sp8 gene spans 4.6 kb and has three exons (Fig. 2). Exons 1a and 1b encode different transcript variants that are likely to be initiated from distinct promoters. The first two exons are separated by a 733 bp intron 1a, followed by a shorter, 152 bp intron 1b. All splice donor/acceptor sites contained consensus GT/AG dinucleotides. The gene possesses a CpG island upstream of exon 1a (438 bp, 63% GC). Computer analysis of the immediate human Sp8 promoter preceding exon 1a contains a number of RNA polymerase II promoter elements including a TATA-box, CCAAT-box and two GC-boxes that are conserved in rodents and fish.

Comparison of the human Sp8 protein with those of other species
The ORF of Sp8L clone 1 encodes a 508-residue protein with a 50,500 Da molecular mass and an isoelectric point 9.02. The first ATG codon of clone 2 corresponds to the second ATG of clone 1, thereby omitting the first 18 residues of the ORF, encoding a 490-residue protein with a 48,674 Da molecular mass and an isoelectric point 9.10.

The human Sp8 protein belongs to the Sp/Krüppel like factor (KLF) family of proteins that are characterised by three Cys2-His2 zinc-fingers [14]. The zinc-fingers are located towards the COOH terminus and are involved in binding DNA in a sequence specific manner. However, the Sp proteins are distinguished from the KLF family by an amino-terminal hydrophobic domain called the Sp-motif with the consensus sequence PLALLA and a buttonhead box (BTD) with the consensus sequence CxCP(N/Y)C prior to the zinc finger domain.

Database searches identified complete Sp8 genes and protein sequences from other vertebrate species. The chimpanzee gene is located on chromosome 7 (clone RP43-37F3); the mouse gene at 12F2 (BAC clone RP23-161L22); the rat gene at 6q33 on BAC clone CH230-1K24 and the zebrafish within BAC clone CH211-180P9. A comparison of the protein sequences shows that Sp8 is well conserved through vertebrate evolution (Fig. 3). The

```
(A) Exon 1a
 -88   attgtattgcacacctctaaaaaaaacactgcctctgatttatcaatataaaaaagatcctctgagaggaggagggcacttttgtgtg
   1 atggcaacttcacttctaggg
     M  A  T  S  L  L  G
(B) Exon 1b
 -97 gcgtggtgcttgctcccgaatcagacgccggcagccaaacttgtcccctcctgtagagtaggaagcggccgggcgccggggctgttgggg
     gtgccag
(C) Exon 2
                        gaagaaccgaggttgggatcgactcctctggccatgcttgccgctacctgtaataagataggcagcccc
                        E  E  P  R  L  G  S  T  P  L  A  M  L  A  A  T  C  N  K  I  G  S  P
  91 agcccgtctccctcctccctctcggacagctcttcttccttcggcaaaggcttccacccctggaaacgctcctcgtcctcttcttccgcc
     S  P  S  P  S  S  L  S  D  S  S  S  S  F  G  K  G  F  H  P  W  K  R  S  S  S  S  S  S  A
 181 agctgcaacgtagtgggttccagtctctcaagcttcggcgtgtccggggcctccaggaacggcggctcgtcctcggcggctgcggcggcc
     S  C  N  V  V  G  S  S  L  S  S  F  G  V  S  G  A  S  R  N  G  G  S  S  S  A  A  A  A
 271 gcggcagcagccgcggctgccgcggccctggtgtccgactcgttcagctgcggcggctcgcctggctccagcgccttctccctcacctcc
     A  A  A  A  A  A  A  A  A  L  V  S  D  S  F  S  C  G  G  S  P  G  S  S  A  F  S  L  T  S
 361 agcagcgccgcagccgccgccgccgccgccgcagccgccgcctccagctcgcccttcgccaacgactactctgttttccaggcccccgga
     S  S  A  A  A  A  A  A  A  A  A  A  A  A  S  S  S  P  F  A  N  D  Y  S  V  F  Q  A  P  G
 451 gtttccggggggcagcggcggcggcggcggggggcggcggcggcggctcctccgcgcactcgcaggacggctcccaccagccggtgttcatc
     V  S  G  G  S  G  G  G  G  G  G  G  G  G  S  S  A  H  S  Q  D  G  S  H  Q  P  V  F  I
 541 tccaaggtgcacacctctgtggacgggctgcagggcatctacccgcgggtgggcatggcgcacccgtacgagtcgtggtttaagccctcg
     S  K  V  H  T  S  V  D  G  L  Q  G  I  Y  P  R  V  G  M  A  H  P  Y  E  S  W  F  K  P  S
 631 cacccgggcctgggtgctgcgggcgaggtgggctcggccggcgcctccagctggtgggacgtgggggccggctggatcgacgtgcagaac
     H  P  G  L  G  A  A  G  E  V  G  S  A  G  A  S  S  W  W  D  V  G  A  G  W  I  D  V  Q  N
 721 ccgaacagcgcggctgcgctgcccggctcgctgcaccctgccgccggggggctccaaacctcgctgcactcgccgctcggaggctacaac
     P  N  S  A  A  A  L  P  G  S  L  H  P  A  A  G  G  L  Q  T  S  L  H  S  P  L  G  G  Y  N
 811 tcggattactcgggcctgagtcactcggccttcagcagcggcgcctcctcgcacctgctcagccccgccgggcagcacctcatggacggc
     S  D  Y  S  G  L  S  H  S  A  F  S  S  G  A  S  S  H  L  L  S  P  A  G  Q  H  L  M  D  G
 901 ttcaagccagtgctaccgggctcctacccgggactcggccccgtcgccgctggccggcgcggggggctccatgttgagcgctgggccttcg
     F  K  P  V  L  P  G  S  Y  P  D  S  A  P  S  P  L  A  G  A  G  G  S  M  L  S  A  G  P  S
 991 gcgccgctggggggctccccgcgctcctcagctcgccgctactccggccgcgccacctgcgactgccccaactgccaggaggcagagcgg
     A  P  L  G  G  S  P  R  S  S  A  R  R  Y  S  G  R  A  T  C  D  C  P  N  C  Q  E  A  E  R
1081 ctgggccctgccggggcgagcttgcggcgcaagggcctgcacagctgccacatcccgggctgcggcaaggtgtacggcaagacttcgcac
     L  G  P  A  G  A  S  L  R  R  K  G  L  H  S  C  H  I  P  G  C  G  K  V  Y  G  K  T  S  H
1171 ctcaaggcgcacctgcgctggcacacgggcgagcggcccttcgtgtgcaactggctttttctgcggcaagcgcttcacgcgctccgacgag
     L  K  A  H  L  R  W  H  T  G  E  R  P  F  V  C  N  W  L  F  C  G  K  R  F  T  R  S  D  E
1261 ctgcagcggcacctgcggacccacaccggcgagaagcgcttcgcctgtccagtttgcaacaagcgcttcatgcgcagcgaccacctcagc
     L  Q  R  H  L  R  T  H  T  G  E  K  R  F  A  C  P  V  C  N  K  R  F  M  R  S  D  H  L  S
1351 aagcacgtgaagacgcacagtggcggcggcggcggcggcggcggctcggcgggctcgggcagcggcggcaagaagggcagcgacaccgacagc
     K  H  V  K  T  H  S  G  G  G  G  G  G  G  S  A  G  S  G  S  G  G  K  K  G  S  D  T  D  S
1441 gagcacagcgccgcgggcagcccgccctgccactccccagagctgctgcagcccccgagccgggcaccgcaacggcctagagtgacgc
     E  H  S  A  A  G  S  P  P  C  H  S  P  E  L  L  Q  P  P  E  P  G  H  R  N  G  L  E  *
```

Figure 1

Human Sp8 cDNA sequences encompassing the ORF and translation. (A) Sequence encoding exon 1a encoding the start of Sp8L protein isoform. (B) Sequence encoding the untranslated exon 1b encoding the Sp8S protein isoform. (C) Sequence encoding the common exon 2. The atg/methionine start codon for the long protein isoform is shown in bold and coloured red and that of the short protein isoform coloured pink. The ORF stop codon, **tga**, is shown in bold and coloured red and is indicated by a red asterisk. An in-frame stop codon, **tag**, in the 5'UTR of exon 1b is shown in bold and coloured red. The exon/exon boundaries were determined by comparison with the sequence of genomic DNA clone CTA-324D18. The nucleotides underlined and shown in bold type indicate the location of the exon/exon boundaries. The loss of a GGC (glycine) codon found in an additional Sp8 clone is shown in green and bolded.

human and chimpanzee proteins are very similar with one conservative substitution and an additional glycine residue. The mouse Sp8 protein [9,10,15] has 97% identity and the zebrafish 79% identity to human Sp8. An amphibian Sp8 EST from *Xenopus laevis* was identified by our database search (accession No. BI313193). The mammalian Sp8 proteins differ from that of the zebrafish by the insertion of two poly-alanine and a poly-glycine region in the amino-half of the protein.

Database searches also showed that homologues of the vertebrate Sp8 genes are found in insects, but not in nematodes, yeasts and higher plants (Fig. 4). The fruit fly, *Drosophila melanogaster*, D-Sp1 gene is a homologue of Sp8

Figure 2
Gene structure of human Sp8. The human genomic sequences corresponding to the Sp8 cDNAs are located on clones CTA-324D18 and CTB-86D3. The location of a CpG island is shown in green. The first exon encoding the start of the Sp8L transcript is labelled 1a and that of Sp8S, 1b. The ORF is indicated by closed boxes. The sequences around the initiation methionine codons (underlined) are shown with those nucleotides corresponding to Kozak consensus sequence coloured red. The location of the TGA stop codon and a potential polyadenylation signal, AATAAA are also shown. The sizes of the exons and introns, in base pairs, are indicated.

and not, as the name implies, Sp1 nor Sp4 as previously stated [16]. The red flour beetle, *Tribolium castaneum*, Sp8 protein [17] has 43% identity and 80% similarity to human Sp8. However, D-Sp1 differs from the vertebrate and beetle Sp8 proteins possessing a much longer, glutamine and histidine-rich carboxy-terminus. The possession of Sp, BTD and zinc-fingers domains in a protein is characteristic of Sp proteins in general; these domains in these three Sp8 proteins are very highly conserved. Outside these domains, a motif of unknown function is located towards the amino-terminal with the sequence GKGFHPWKKS and is unique to Sp8 proteins.

By comparison with other Sp proteins, the Sp8 protein can be divided up into 5 domains. The amino-terminal domain A (residues 1–53) is highly conserved in Sp8 proteins from other species. Two alpha-helices, residues 3–7 and 15–27, are predicted in this domain; whereas the rest of the protein appears to lack any secondary structure except for the zinc-finger domain. Domain A is also found in the Sp7 proteins. Domain B, residues 54–167, is a low-complexity region, having poly- serine, alanine and glycine stretches that are not present in fish Sp8 proteins. Domain C, residues 168–373, is similar to the same region of Sp7 and Sp6 that has been shown to be involved

in transcriptional activation [7] and contains a basic region (residues 360–371) that probably contributes to a nuclear localization signal and a BTD-box, a highly conserved region that is present in all Sp family members. The BTD-box was first described as a 10-amino-acid region in the Drosophila zygotic gene, buttonhead [11]. Domain D, residues 374–456 contains three classical zinc finger structures (residues 374–456) of the Cys2-His2 type where the conserved cysteine residues in two short beta-sheets and the histidine residues in an alpha-helix tetrahedrally coordinate a zinc ion [18]. Between the fingers are five-residue linker sequences with the consensus sequence TGE+x. The negatively charged carboxy-terminal domain E, residues 457–508, contains a glycine-rich region (458–475).

Tissue distribution of Sp8 mRNA
The expression of Sp8 in human and murine adult tissues was examined by reverse transcriptase PCR using primers located in the 3'UTR. Human Sp8 mRNA expression was detected in a wide range of tissues at a low level, being found in heart, brain, placenta, liver, pancreas, prostate, testis, ovary and colon, but was below the limit of detection in lung, skeletal muscle, kidney, spleen, thymus, small intestine and peripheral blood leukocytes (Fig. 5A). Murine Sp8 mRNA expression, relative to the expression
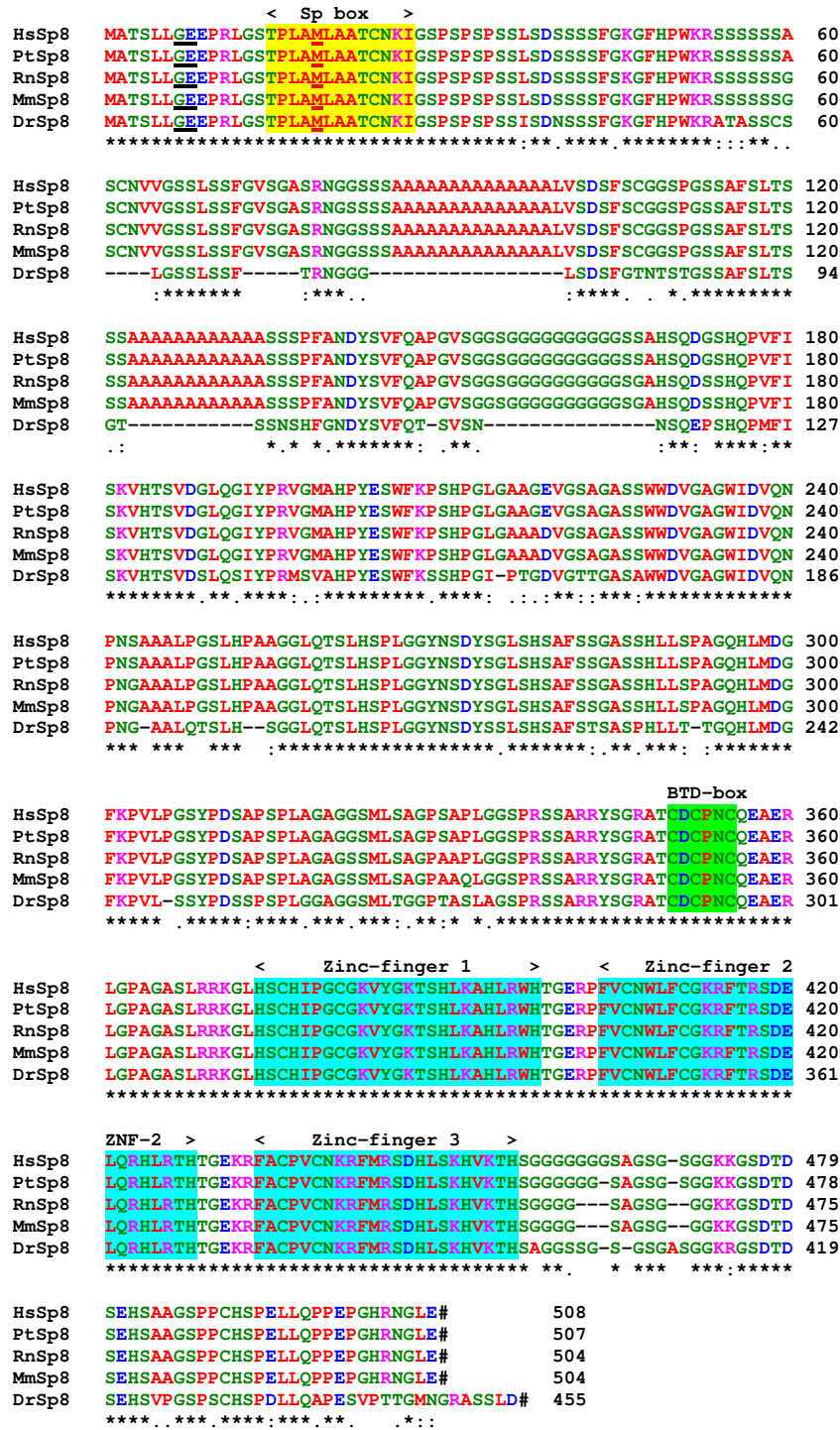
```
                              <   Sp box   >
HsSp8   MATSLLGEEPRLGSTPLAMLAATCNKIGSPSPSPSSLSDSSSSFGKGFHPWKRSSSSSSA  60
PtSp8   MATSLLGEEPRLGSTPLAMLAATCNKIGSPSPSPSSLSDSSSSFGKGFHPWKRSSSSSSA  60
RnSp8   MATSLLGEEPRLGSTPLAMLAATCNKIGSPSPSPSSLSDSSSSFSKGFHPWKRSSSSSSG  60
MmSp8   MATSLLGEEPRLGSTPLAMLAATCNKIGSPSPSPSSLSDSSSSFGKGFHPWKRSSSSSSG  60
DrSp8   MATSLLGEEPRLGSTPLAMLAATCNKIGSPSPSPSSISDNSSSFGKGFHPWKRATASSCS  60
        *********************************:**.****.*******::**..

HsSp8   SCNVVGSSLSSFGVSGASRNGGSSSAAAAAAAAAAAAAAALVSDSFSCGGSPGSSAFSLTS 120
PtSp8   SCNVVGSSLSSFGVSGASRNGSSSSAAAAAAAAAAAAAAALVSDSFSCGGSPGSSAFSLTS 120
RnSp8   SCNVVGSSLSSFGVSGASRNGGSSSAAAAAAAAAAAAAAALVSDSFSCGGSPGSSAFSLTS 120
MmSp8   SCNVVGSSLSSFGVSGASRNGGSSSAAAAAAAAAAAAAAALVSDSFSCGGSPGSSAFSLTS 120
DrSp8   ----LGSSLSSF-----TRNGGG----------------LSDSFGTNTSTGSSAFSLTS    94
            :.*******     :***..               :****. . *.*********

HsSp8   SSAAAAAAAAAAAASSSPFANDYSVFQAPGVSGGSGGGGGGGGGGSSAHSQDGSHQPVFI 180
PtSp8   SSAAAAAAAAAAAASSSPFANDYSVFQAPGVSGGSGGGGGGGGGGSSAHSQDGSHQPVFI 180
RnSp8   SSAAAAAAAAAAAASSSPFANDYSVFQAPGVSGGSGGGGGGGGGGSGAHSQDSSHQPVFI 180
MmSp8   SSAAAAAAAAAAAASSSPFANDYSVFQAPGVSGGSGGGGGGGGGGSGAHSQDSSHQPVFI 180
DrSp8   GT-----------SSNSHFGNDYSVFQT-SVSN---------------NSQEPSHQPMFI 127
        .:               *.* *.*******: .**.               :**: ****:**

HsSp8   SKVHTSVDGLQGIYPRVGMAHPYESWFKPSHPGLGAAGEVGSAGASSWWDVGAGWIDVQN 240
PtSp8   SKVHTSVDGLQGIYPRVGMAHPYESWFKPSHPGLGAAGEVGSAGASSWWDVGAGWIDVQN 240
RnSp8   SKVHTSVDGLQGIYPRVGMAHPYESWFKPSHPGLGAAADVGSAGASSWWDVGAGWIDVQN 240
MmSp8   SKVHTSVDGLQGIYPRVGMAHPYESWFKPSHPGLGAAADVGSAGASSWWDVGAGWIDVQN 240
DrSp8   SKVHTSVDSLQSIYPRMSVAHPYESWFKSSHPGI-PTGDVGTTGASAWWDVGAGWIDVQN 186
        ********.**.****:.:*********.****:  .::.:**::***:************

HsSp8   PNSAAALPGSLHPAAGGLQTSLHSPLGGYNSDYSGLSHSAFSSGASSHLLSPAGQHLMDG 300
PtSp8   PNSAAALPGSLHPAAGGLQTSLHSPLGGYNSDYSGLSHSAFSSGASSHLLSPAGQHLMDG 300
RnSp8   PNGAAALPGSLHPAAGGLQTSLHSPLGGYNSDYSGLSHSAFSSGASSHLLSPAGQHLMDG 300
MmSp8   PNGAAALPGSLHPAAGGLQTSLHSPLGGYNSDYSGLSHSAFSSGASSHLLSPAGQHLMDG 300
DrSp8   PNG-AALQTSLH--SGGLQTSLHSPLGGYNSDYSSLSHSAFSTSASPHLLT-TGQHLMDG 242
        *** ***  ***   :.******************.*******:.**.***: :*******

                                                            BTD-box
HsSp8   FKPVLPGSYPDSAPSPLAGAGGSMLSAGPSAPLGGSPRSSARRYSGRATCDCPNCQEAER 360
PtSp8   FKPVLPGSYPDSAPSPLAGAGGSMLSAGPSAPLGGSPRSSARRYSGRATCDCPNCQEAER 360
RnSp8   FKPVLPGSYPDSAPSPLAGAGSSMLSAGPAAPLGGSPRSSARRYSGRATCDCPNCQEAER 360
MmSp8   FKPVLPGSYPDSAPSPLAGAGSSMLSAGPAAQLGGSPRSSARRYSGRATCDCPNCQEAER 360
DrSp8   FKPVL-SSYPDSSPSPLGGAGGSMLTGGPTASLAGSPRSSARRYSGRATCDCPNCQEAER 301
        ***** .*****.****.***.***:.**:.**:* *.************************

                     <       Zinc-finger 1       >   <    Zinc-finger 2
HsSp8   LGPAGASLRRKGLHSCHIPGCGKVYGKTSHLKAHLRWHTGERPFVCNWLFCGKRFTRSDE 420
PtSp8   LGPAGASLRRKGLHSCHIPGCGKVYGKTSHLKAHLRWHTGERPFVCNWLFCGKRFTRSDE 420
RnSp8   LGPAGASLRRKGLHSCHIPGCGKVYGKTSHLKAHLRWHTGERPFVCNWLFCGKRFTRSDE 420
MmSp8   LGPAGASLRRKGLHSCHIPGCGKVYGKTSHLKAHLRWHTGERPFVCNWLFCGKRFTRSDE 420
DrSp8   LGPAGASLRRKGLHSCHIPGCGKVYGKTSHLKAHLRWHTGERPFVCNWLFCGKRFTRSDE 361
        ************************************************************

        ZNF-2  >      <     Zinc-finger 3      >
HsSp8   LQRHLRTHTGEKRFACPVCNKRFMRSDHLSKHVKTHSGGGGGGGGSAGSG-SGGKKGSDTD 479
PtSp8   LQRHLRTHTGEKRFACPVCNKRFMRSDHLSKHVKTHSGGGGGG-SAGSG-SGGKKGSDTD 478
RnSp8   LQRHLRTHTGEKRFACPVCNKRFMRSDHLSKHVKTHSGGGG---SAGSG--GGKKGSDTD 475
MmSp8   LQRHLRTHTGEKRFACPVCNKRFMRSDHLSKHVKTHSGGGG---SAGSG--GGKKGSDTD 475
DrSp8   LQRHLRTHTGEKRFACPVCNKRFMRSDHLSKHVKTHSAGGSSG-S-GSGASGGKRGSDTD 419
        ************************************** **.  * ***  ***:*****

HsSp8   SEHSAAGSPPCHSPELLQPPEPGHRNGLE#       508
PtSp8   SEHSAAGSPPCHSPELLQPPEPGHRNGLE#       507
RnSp8   SEHSAAGSPPCHSPELLQPPEPGHRNGLE#       504
MmSp8   SEHSAAGSPPCHSPELLQPPEPGHRNGLE#       504
DrSp8   SEHSVPGSPSCHSPDLLQAPESVPTTGMNGRASSLD#  455
        ****..***.****:***.**.    .*::
```

Figure 3

Sequence comparison of the human Sp8 protein with those from other vertebrate species. The species are man, *Homo sapiens*, Hs; mouse, *Mus musculus*, Mm; chimpanzee, *Pan troglodytes*, Pt; rat *Rattus norvegicus*, Rn and Zebrafish, *Danio rerio*, Dr. The location of the exon/exon boundaries are underlined in black on the protein sequences. The methionine residues at the start of translation of Sp8S are underlined in red. The positions of the Sp, BTD-box and three zinc-finger domains found in most Sp protein family members are indicated. The conserved residues are shown by (*), strongly conserved residues by (:) and weakly conserved residues by (.). Stop codons are indicated by a hash. Residues are colour coded: basic, DE, blue; acidic,_KR, pink; polar, CGHNQSTY, green and hydrophobic, AFILMPVW, red.
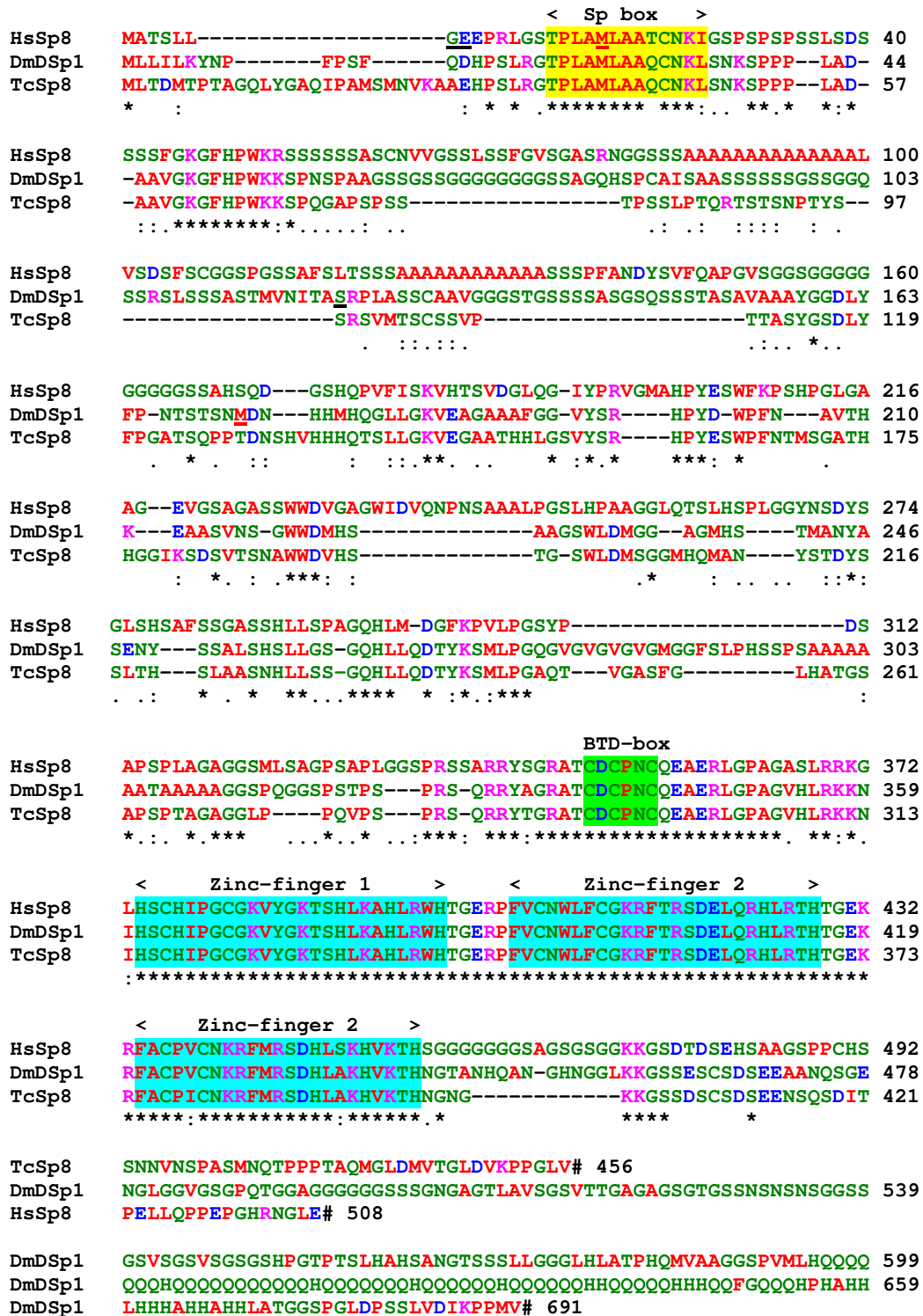
```
                                                < Sp box    >
HsSp8    MATSLL------------------GEEPRLGSTPLAMLAATCNKIGSPSPSPSSLSDS 40
DmDSp1   MLLILKYNP-------FPSF-----QDHPSLRGTPLAMLAAQCNKLSNKSPPP--LAD- 44
TcSp8    MLTDMTPTAGQLYGAQIPAMSMNVKAAEHPSLRGTPLAMLAAQCNKLSNKSPPP--LAD- 57
         *   :                   : * * .******** ***:.. **.*  *:*

HsSp8    SSSFGKGFHPWKRSSSSSSASCNVVGSSLSSFGVSGASRNGGSSSAAAAAAAAAAAAAAL 100
DmDSp1   -AAVGKGFHPWKKSPNSPAAGSSGSSSGGGGGGGGSSAGQHSPCAISAASSSSSSGSSGGQ 103
TcSp8    -AAVGKGFHPWKKSPQGAPSPSS------------------TPSSLPTQRTSTSNPTYS-- 97
          ::.********:*.....: ..                  .: .:  ::::  : .

HsSp8    VSDSFSCGGSPGSSAFSLTSSSAAAAAAAAAAAAASSSPFANDYSVFQAPGVSGGSGGGGG 160
DmDSp1   SSRSLSSSASTMVNITASRPLASSCAAVGGGSTGSSSSASGSQSSSTASAVAAAYGGDLY 163
TcSp8    -----------------SRSVMTSCSSVP--------------------TTASYGSDLY 119
                          .  ::.::.                      .:.. *..

HsSp8    GGGGGSSAHSQD---GSHQPVFISKVHTSVDGLQG-IYPRVGMAHPYESWFKPSHPGLGA 216
DmDSp1   FP-NTSTSNMDN---HHMHQGLLGKVEAGAAAFGG-VYSR----HPYD-WPFN---AVTH 210
TcSp8    FPGATSQPPTDNSHVHHHQTSLLGKVEGAATHHLGSVYSR----HPYESWPFNTMSGATH 175
         .  * . ::     :  ::.**. ..    * :*.*    ***: *       .

HsSp8    AG--EVGSAGASSWWDVGAGWIDVQNPNSAAALPGSLHPAAGGLQTSLHSPLGGYNSDYS 274
DmDSp1   K---EAASVNS-GWWDMHS--------------AAGSWLDMGG--AGMHS----TMANYA 246
TcSp8    HGGIKSDSVTSNAWWDVHS--------------TG-SWLDMSGGMHQMAN----YSTDYS 216
          :   *.   : .***: :               .*     : . .. . ::*:

HsSp8    GLSHSAFSSGASSHLLSPAGQHLM-DGFKPVLPGSYP---------------------DS 312
DmDSp1   SENY---SSALSHSLLGS-GQHLLQDTYKSMLPGQGVGVGVGVGMGGFSLPHSSPSAAAAA 303
TcSp8    SLTH---SLAASNHLLSS-GQHLLQDTYKSMLPGAQT---VGASFG---------LHATGS 261
         . .:   * . *  **...****  * :*.:***                          :

                                              BTD-box
HsSp8    APSPLAGAGGSMLSAGPSAPLGGSPRSSARRYSGRATCDCPNCQEAERLGPAGASLRRKG 372
DmDSp1   AATAAAAAGGSPQGGSPSTPS---PRS-QRRYAGRATCDCPNCQEAERLGPAGVHLRKKN 359
TcSp8    APSPTAGAGGLP----PQVPS---PRS-QRRYTGRATCDCPNCQEAERLGPAGVHLRKKN 313
         *.:. *.***    ...*.* ..:***: ***:******************** **:*.

         <     Zinc-finger 1     >    <     Zinc-finger 2     >
HsSp8    LHSCHIPGCGKVYGKTSHLKAHLRWHTGERPFVCNWLFCGKRFTRSDELQRHLRTHTGEK 432
DmDSp1   IHSCHIPGCGKVYGKTSHLKAHLRWHTGERPFVCNWLFCGKRFTRSDELQRHLRTHTGEK 419
TcSp8    IHSCHIPGCGKVYGKTSHLKAHLRWHTGERPFVCNWLFCGKRFTRSDELQRHLRTHTGEK 373
         :*********************************************************

         <   Zinc-finger 2    >
HsSp8    RFACPVCNKRFMRSDHLSKHVKTHSGGGGGGGSAGSGSGGKKGSDTDSEHSAAGSPPCHS 492
DmDSp1   RFACPVCNKRFMRSDHLAKHVKTHNGTANHQAN-GHNGGLKKGSSESCSDSEEAANQSGE 478
TcSp8    RFACPICNKRFMRSDHLAKHVKTHNGNG------------KKGSSDSCSDSEENSQSDIT 421
         *****:***********:******.*           ****        *

TcSp8    SNNVNSPASMNQTPPPTAQMGLDMVTGLDVKPPGLV# 456
DmDSp1   NGLGGVGSGPQTGGAGGGGGGSSSGNGAGTLAVSGSVTTGAGAGSGTGSSNSNSNSGGSS 539
HsSp8    PELLQPPEPGHRNGLE# 508

DmDSp1   GSVSGSVSGSGSHPGTPTSLHAHSANGTSSSLLGGGLHLATPHQMVAAGGSPVMLHQQQQ 599
DmDSp1   QQQHQQQQQQQQQQHQQQQQQHQQQQQHQQQQQHHQQQQQHHHQQFGQQQHPHAHH 659
DmDSp1   LHHHAHHAHHLATGGSPGLDPSSLVDIKPPMV# 691
```

Figure 4

Sequence comparison of the human Sp8 protein with those from other invertebrate species; red flour beetle, *Tribolium casta-neum*, TcSp8 and fruit fly, *Drosophila melanogaster* Sp8, DmD-Sp1 proteins. The locations of the exon/exon boundaries are underlined in black on the protein sequences (GE and S). The methionine residues at the start of translation for the short protein isoforms are underlined in red (M). The positions of the Sp, BTD-box and three zinc-fingers motifs are shown. Stop codons are indicated by a hash.
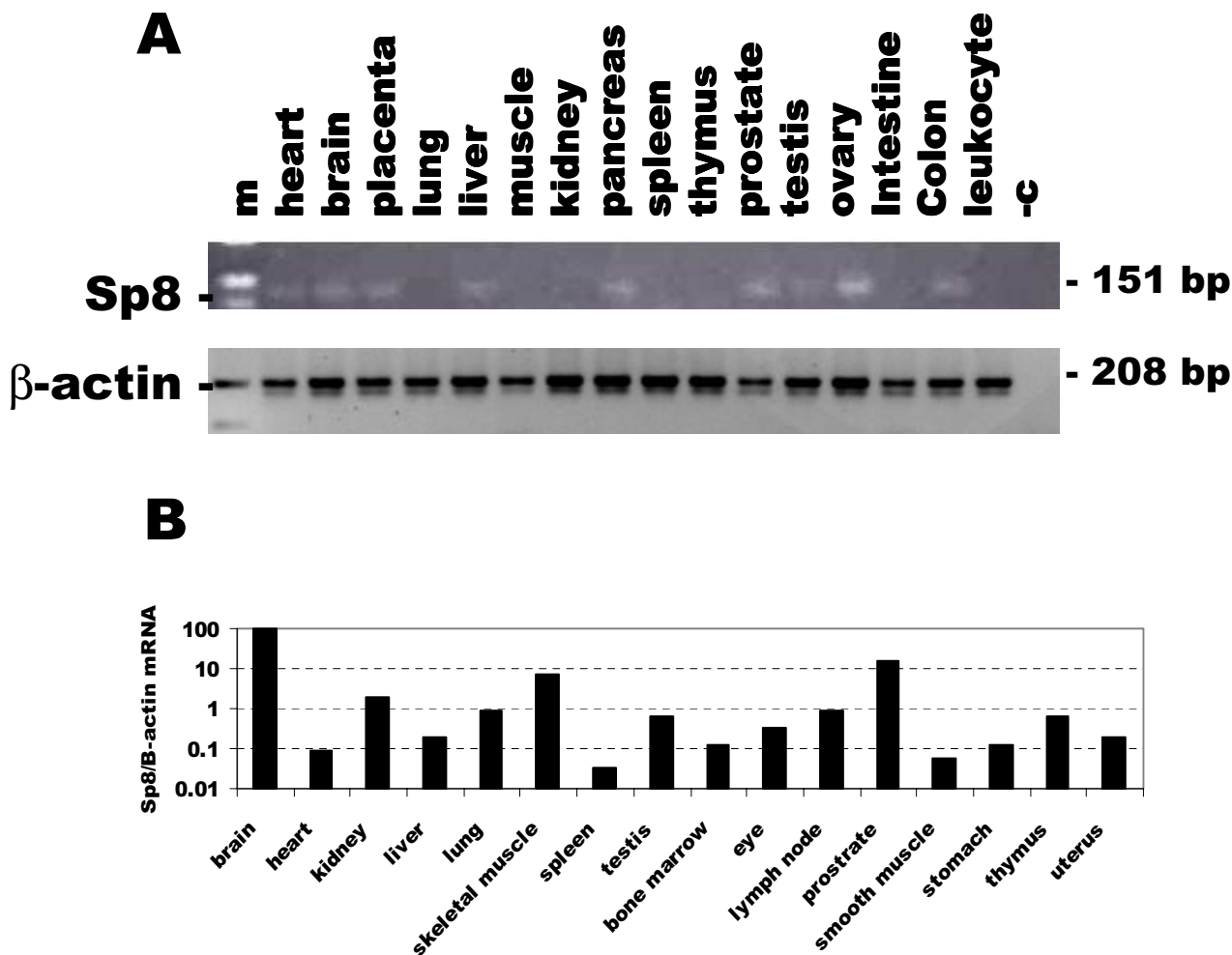
Figure 5
Expression of Sp8 mRNA in human and mouse adult tissues determined by RT-PCR using primers located in the 3'UTR. **A**. Ethidium bromide stained agarose gel of PCR products from human tissues. The lanes are: skeletal muscle, muscle; small intestine, intestine; peripheral blood leukocytes, leukocyte; negative control, -c and 100 bp ladder, m. The expression of beta-actin mRNA, a housekeeping gene, in the same samples is shown below. **B**. RT real time PCR of murine Sp8 mRNA expression normalised to that of beta-actin.

of beta-actin mRNA, was detected in all tissues examined, with the highest levels being found in brain and prostate and the lowest in spleen (Fig. 5B).

Expression of human Sp8 mRNA in osteoblast-like cells
The expression of both transcript variants mRNAs encoding Sp8L and Sp8S in osteoblast-like cells was examined by reverse transcriptase PCR using primers specific for each isoform (Fig. 6). The Sp8L was most abundant in the osteosarcoma cell lines HOS and MG63 and was present in adult and craniofacial osteoblasts, but was below the limit of detection in foetal osteoblasts and chondrocytes.

The Sp8S was most abundant also in the osteosarcoma cell lines HOS and MG63 and was present in craniofacial and foetal osteoblasts, but was below the limit of detection in adult osteoblasts and chondrocytes. Runx2, a transcription factor important for osteoblast differentiation, and beta-actin, a housekeeping gene, were expressed in all the osteoblast-like cells.

BMP-2 regulation of Sp8 mRNA expression in the murine cell line in C3H10T1/2
BMPs are potent secreted factors that promote osteoblast differentiation during development and bone
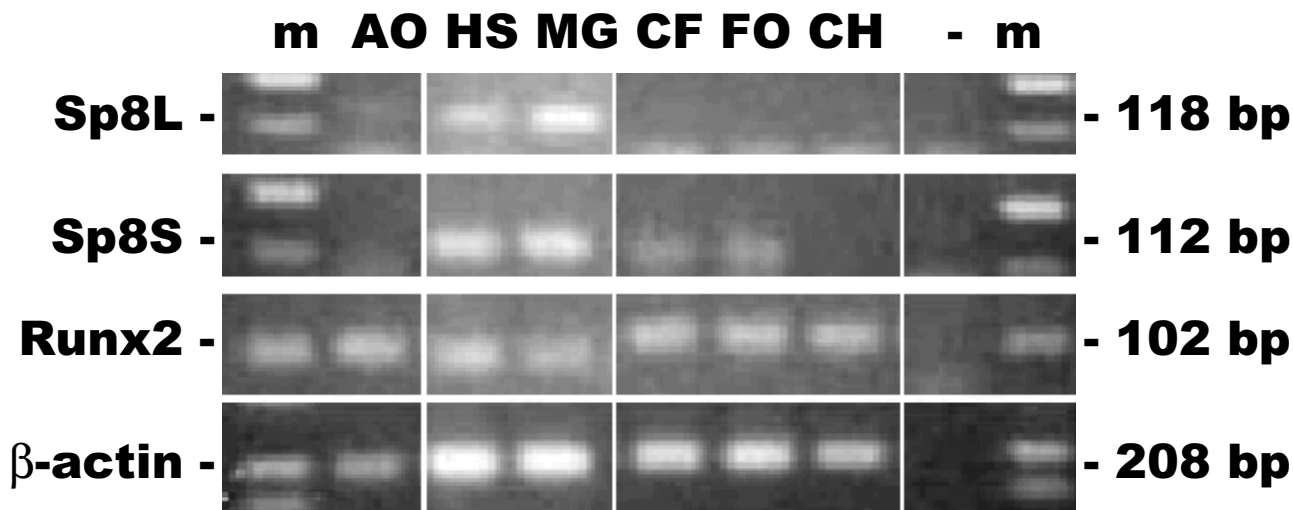
Figure 6
Expression of Sp8 mRNAs in human osteoblast-like cells by reverse transcriptase PCR. Amplicons were run on an ethidium bromide stained agarose gel. The amplicons were: Sp8 long protein isoform, Sp8L; Sp8 short protein isoform, Sp8S; Runx2/ Cbfa1, Runx2 and the housekeeping gene beta-actin. The sizes of the amplicons are indicated. The lanes are: primary adult osteoblasts, AO; HOS osteosarcoma cell line, HS; MG63 osteosarcoma cell line, MG; primary craniofacial osteoblasts, CF; primary foetal osteoblasts, FO; primary adult chondrocytes, CH; negative control, (-) and 100 bp markers, m.
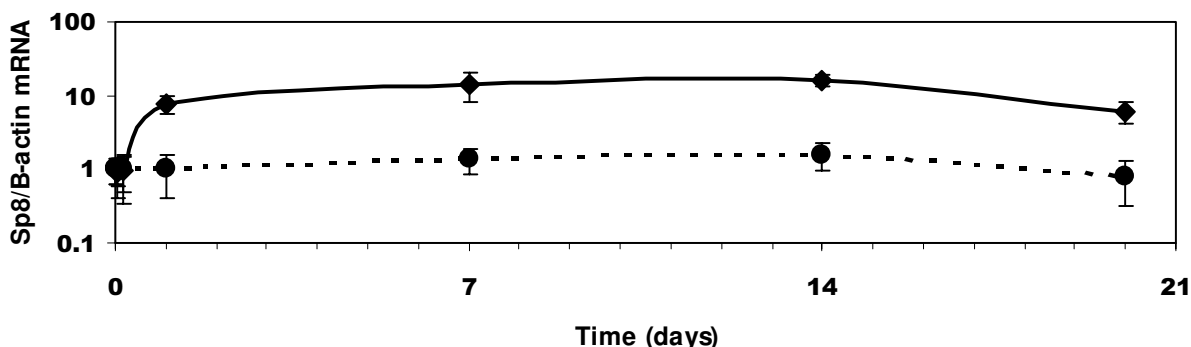
Figure 7
Time course of effect of 100 ng/mL BMP-2 treatment on C3H10T1/2 cells. Murine Sp8 and beta-actin mRNA expression was determined by RT real time PCR. The relative levels of expression in treated (•) and untreated cells (♦) were examined after 0, 7, 14 and 20 days treatment. Significant increases in Sp8 expression were seen at 7, 14 and 20 days ($p < 0.05$).

remodelling. Since Sp8 was found in human osteosarcoma cell lines and is important in skeletal development [9,10] we examined the regulation of Sp8 by BMP-2 in the murine pluripotent embryonic cell line, C3H10T1/2 that

can be induced to form a chondrogenic/osteoblastic phenotype by treatment with BMPs [19-21]. Sp8 mRNA expression, normalised to that of beta-actin, was measured by RT-real time PCR and was induced by 100 ng/mL BMP-2 after one day and remained upregulated at 20 days (Fig. 7). Data normalised to the housekeeping gene G3PDH gave similar results (data not shown) indicating that Sp8 was upregulated and that beta-actin was not down-regulated.

Discussion

In the human genome three Sp gene pairs had been described previously; being maximally separated by 3.2 Mbp, transcribed in opposite directions, being orientated 5' to 5' manner and colocalized with a specific HOX gene clusters [8]. Sp3 and Sp5 are located on chromosome 2q31.1; Sp1 and Sp7 on 12q13.13 and Sp2-Sp6 on 17q21.3-q22. A search for a partner to Sp4 located on 7q21.3-q22 revealed a putative eighth Sp gene most similar to Sp7 suggesting that Sp8, like Sp7, may play a role also in skeletal development. Consequently, we used RT-PCR to amplify Sp8 cDNAs from a human osteosarcoma cell line and isolated two transcript variants clones that encode a full-length protein (long isoform) Sp8L and an amino-terminal truncated protein (short isoform) Sp8S, with 508 and 490 residues respectively. We found a glycine-165 deletion mutation in an additional cDNA clone isolated from a patient with osteoporosis. The length of this poly-glycine region is conserved in primates and rodents, but is absent in fish. We speculate that this Sp8 mutation, and other mutations yet to be discovered, may play a role in susceptibility to osteoporosis, and since Sp8 plays an important role in neuropore closure be a candidate gene for spina bifida [9].

The sequence of our Sp8L clone is similar to that of two other human clones, but the other two clones lack 170 and 128 bp of the low-complexity, B domain (GenBank accession Nos. BAB71297 and AAH38669). These deletions may be caused by incomplete reverse transcriptase reactions occurring during cDNA library synthesis, presumably because the high GC content (81%) generates strong secondary structure in the mRNA in this region. These deletion-carrying clones are unlikely to be generated by the presence of cryptic intron(s) since no suitable donor or acceptor splice sites are present in the genomic DNA sequence. One of the deletion-carrying clones, AAH38669, possesses a polymorphism, cac>cgc, at nucleotide 1263 that results in a His>Arg mutation at residue 448. His-448 is located in the third zinc- finger and is conserved in all Sp proteins and, by homology with Sp1, is likely to contact DNA [18]. This mutated protein would be expected to be deleterious, having reduced affinity for GC-box binding sites in promoter regions.

We found that in man Sp8 has two transcript variants utilising two different first exons. Sequence analyses of mouse and zebrafish genomic and EST data support this gene structure for Sp8 in vertebrates and not a gene structure with two untranslated 5' exons that has been suggested for murine Sp8 [10]. Interestingly, the drosophila D-Sp1 gene, a Sp8 homologue, also has two transcript variants, encoding long and short protein isoforms (accession No. AE003448) [22]. Both Sp8 and Sp7 have a similar exon structure, being three-exon genes with two transcript variants with different first exons [8]. The long protein isoforms use translated first exons that encodes seven residues and the short protein isoforms use untranslated first exons; overall, the proteins have 39.5% identity and 65.9% similarity. Both the Sp8S and Sp7S proteins lack an 18-residue amino-terminus thereby disrupting a hydrophobic region termed the Sp domain that is conserved in other Sp proteins. The conservation of the two protein isoforms through evolution suggests that they have differing roles. Although the function of this hydrophobic region is unknown, other zinc-finger transcription proteins often have a conserved protein-protein interaction domain at their amino-terminus (e.g. BTB or kelch domains). This indicates that the Sp motif may also be involved in a protein-protein interaction and that the short protein isoforms do not have this protein interaction domain.

Amino-terminal spliced variants expressed from separate promoters are a feature of other important transcription factors that regulate skeletal development. An oestrogen regulated protein in osteoblasts, KLF10 is another member of the Sp/Krüppel-like factor family that has two amino-terminal variant isoforms generated in a similar way to those of Sp8. These isoforms are named the TGFβ-inducible early gene (TIEG1) and the early growth response gene-alpha [23-25]. Runx2 has two major amino-terminal isoforms that exert different functions during the process of osteoblast differentiation; the Runx2-type-I isoform is widely expressed in osteoprogenitor cells and active osteoblasts, whereas the Runx2-type-II isoform is restricted to cells lining mineralised bones [26] and BMP-2 preferentially upregulates the Runx2-type-II isoform [27].

Sp8 does not initiate BMP-mediated signalling during apical ectodermal ridge formation, but may function downstream of the BMP receptor-1a in the signalling cascade [9,10]. The signalling events downstream of the BMP receptor that result in tissue-specific gene expression and skeletal development have been only partially elucidated. We found that Sp8 expression was induced by 100 ng/mL BMP-2 only after 24 hours in C3H10T1/2 cells. BMP-2 has been shown to induce several other transcription factors that promote differentiation such as Runx2/Cbfa1, Sp7/

osterix and ZNF450 in addition to the negative regulator Id1 [7,28-30]. Induction of these genes in C3H10T1/2 cells occurs within 4 hours, preceding that of Sp8, suggesting that Sp8 is not directly regulated by BMP signalling and that it may be induced by one or more of the BMP-early-induced genes.

The Sp8 gene has a wider phylogenetic distribution than Sp7, being found in coelomates, whereas Sp7 is limited to vertebrates. In view of the close similarities between human Sp8 and Sp7 in gene structure and amino acid sequence we speculate that they evolved from an ancestral Sp8 gene during a duplication of a Sp/Hox gene cluster [31]. The Sp8 gene has retained its function in regulating appendage/limb growth [9,10,12,17] and that Sp7 has subsequently evolved a novel function, namely, regulating cartilage/bone formation [7].

Conclusions

In humans, there are two transcript variants of Sp8 that utilise different first exons encoding long and short protein isoforms. These two isoforms are conserved being found also in the zebrafish suggesting that they have distinct functions. Sp8 is upregulated by BMP-2 in the murine pluripotent cell line C3H10T1/2 and may play a role in mesenchymal differentiation.

Methods

Cell culture and RNA extraction

Two cell lines derived from osteosarcomas, HOS and MG-63, were cultured at 37°C in 5% CO_2 using Dulbecco's Modified Eagle's Medium, containing 4 mM L-glutamine, 4500 mg glucose/L, 1500mg bicarbonate/L (Invitrogen, UK) with the addition of 10% foetal bovine serum, 10 μM ascorbic acid, 100 IU/mL penicillin and 50 μg/ml streptomycin. Primary human osteoblasts were isolated from trabecular bone of femoral heads taken during total hip arthroplasty and cultured as previously described [32]. Primary human craniofacial osteoblasts were obtained from paediatric skull and cultured as previously described [33]. Human primary foetal osteoblasts were obtained and cultured as previously described [34]. Human primary articular chondrocytes were obtained from isolated femoral heads and cultured as previously described [35].

The murine cell line C3H10T1/2 was cultured in Dulbecco's modified Eagle's medium supplemented with 50 U/mL penicillin, 50 μg/mL streptomycin (Invitrogen) and 10% new born bovine serum (Sigma) at 37°C under 5% CO_2 in a humidified incubator. Prior to BMP-2 treatment, cells were seeded at a density of approximately 40,000 cells/cm², left for 24 hours, then treated with fresh media with or without 100 ng/mL human recombinant BMP-2 (Wyeth Corporation) and the media replaced every three days.

Total RNA was extracted from cells using guanidine thiocyanate and treated with DNase-I to remove any contaminating genomic DNA (SV Total RNA isolation system, Promega, UK). The concentration and purity of eluted RNA was determined spectrophotometrically and the quality of the RNA was verified by non-denaturating agarose gel electrophoresis. For Sp8 cloning, total RNA was reverse transcribed with an oligo-dT primer using ThermoScript, an AMV RNase H- reverse transcriptase at an elevated temperature of 60°C (Invitrogen, UK).

Molecular cloning of human Sp8

The human Sp7 protein sequence [8] was used to search the human genome sequence and the EST database to identify closely related genes. PCR primers designed from the sequence of genomic DNA clones CTA-324D18, CTB-86D3 and EST sequence from Image clone 2721342 (accession No. AW207154) were used to clone two transcript variants spliced isoforms of the human Sp8 cDNA by PCR from MG-63 osteosarcoma cell line cDNA. The PCR primers were: Sp8L forward, ATTGTATTGCACAC-CTCTAAAAAAAACA; Sp8S forward, GCGTGGTGCTT-GCTCCC and common reverse, GCGTCACTCTAGGCCGTTG (Helena Biosciences, France). The cDNAs were amplified by PCR with an annealing temperature of 60°C using the Advantage-GC Advantage kit (Clontech, UK) with the addition of 0.5 M GC-Melt since the DNA sequence of Sp8 is GC-rich. PCR products were excised from agarose gels stained with ethidium bromide and eluted from the agarose using a DNA extraction kit (Qiagen, UK). The PCR products were cloned into the T-A vector pCR4-TOPO (Invitrogen, The Netherlands). Transformed colonies were picked and vectors containing inserts were extracted using the Wizard Plus SV minipreps DNA purification system (Promega, UK) and sequenced in both directions using ThermoFidelase 2 (Fidelity Systems Inc., USA).

Tissue and cellular distribution of human Sp8 mRNA by reverse transcriptase PCR

Human cDNA was analysed for the relative expression of the Sp8, Runx2/Cbfa1 and beta-actin mRNA by PCR. Sixteen adult tissue cDNAs (BD Clontech, UK) were generated from polyA+ selected RNA and reverse transcribed using an oligo-dT primer. Total RNA from cell type cDNAs were reversed transcribed using random hexamer primers using an AMV RNase H- reverse transcriptase (Thermo-Script, Invitrogen, UK) at 60°C. Approximately 4.0 ng of cDNA from each tissue, and cDNA derived from 50 ng of total RNA from each cell type were amplified by PCR using Taq Gold polymerase. Tissue master mixes were divided into gene specific mixes with the addition of PCR primers to a final concentration of 200 μM. The primers were: Sp8L forward primer, GCAACTTCACTTCTAG-GG**GAAGA**(exon 1a/**2**); Sp8S forward primer,

TGGGGGTGCCAG**GAAGAAC**(exon 1b/**2**) and a common reverse primer, AGCTGTCCGAGAGGGAGGA (exon 2), producing a 118 and 112 bp amplicons respectively; Sp8 3'UTR, TTAGTCCGGCCGTCAATTGT and TGGTATTTAAACTACAGCCTCGTCTGA, producing a 151 bp amplicon; Runx2, AGAAGAGCCAGGCAG**GTGCT**(exon 6/7) and TTCGTGGGTTGGAGAAGCG (exon 7), producing a 102 bp amplicon, as measure of the sum of all Runx2 isoforms, and beta-actin, GGCCACGGCTGCTTC and GTTGGCGTACAGGTCTTTGC, producing a 208 bp amplicon. The amplification conditions were; a 10 min hot start to activate the polymerase followed by 50 cycles of 95°C for 15 sec and 60°C for 1 min. Amplification specificity was confirmed by direct sequencing of the amplicons.

Expression of murine Sp8 mRNA by real time PCR

Utilising a CAS-1200 robotic precision liquid handling system, PCR was carried out on a Rotor Gene 3000 (Corbett Research, Australia) using a SYBR Green I double-stranded DNA binding dye assay. Copy DNA derived from 200 ng of total RNA from each sample was amplified by PCR using Taq Gold polymerase using PCR primers to a final concentration of 50 nM. The primers were: mouse Sp8 3'UTR, CCATTCAGCTCTGGCTAGGTCTT and GATTCCCGTTCGCAGAACTC producing a 67 bp amplicon. Beta-actin mRNA was used as a control gene as previously described [36]. The amplification conditions were; a 10 min hot start to activate the polymerase followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min. The number of cycles required for the fluorescence to become significantly higher than background fluorescence (termed cycle threshold [C_t]) was used as a measure of abundance. A comparative C_t method was used to determine gene expression. Expression levels in each cDNA sample were normalised to the expression levels of the control gene (ΔC_t). The ratios of Sp8 mRNA/control gene RNA from each cDNA were standardised to that of the untreated cells on day 0 that was taken as 100% ($\Delta\Delta C_t$). The formula $E^{-\Delta\Delta Ct}$ was used to calculate relative expression levels where E is the efficiency of the PCR per cycle. Statistically significant changes in gene expression were determined using the t-test on data from three replicate experiments. The amplification specificity was confirmed by melting curve analysis and agarose gel electrophoresis.

Authors' contributions

AJE conceived of the study, and participated in its design and coordination. MAM, JEG and AJE carried out the cell culture work, MAM and AJE carried out the expression studies and MAM and AJE carried out the gene cloning. All authors participated in writing the manuscript, read, and approved the final manuscript.

Acknowledgements

References

1. Dynan WS, Tjian R: **Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II.** *Cell* 1983, **32**:669-680.
2. Yoo J, Jeong MJ, Kwon BM, Hur MW, Park YM, Han MY: **Activation of dynamin I gene expression by Sp1 and Sp3 is required for neuronal differentiation of N1E-115 cells.** *J Biol Chem* 2002, **277**:11904-11909.
3. Kolell KJ, Crawford DL: **Evolution of Sp transcription factors.** *Mol Biol Evol* 2002, **19**:216-222.
4. Yu B, Datta PK, Bagchi S: **Stability of the Sp3-DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites.** *Nucleic Acids Res* 2003, **31**:5368-5376.
5. Chadjichristos C, Ghayor C, Kypriotou M, Martin G, Renard E, Ala-Kokko L, Suske G, de Crombrugghe B, Pujol JP, Galera P: **Sp1 and Sp3 transcription factors mediate interleukin-1 beta down-regulation of human type II collagen gene expression in articular chondrocytes.** *J Biol Chem* 2003, **278**:39762-39772.
6. Gollner H, Dani C, Phillips B, Philipsen S, Suske G: **Impaired ossification in mice lacking the transcription factor Sp3.** *Mech Dev* 2001, **106**:77-83.
7. Nakashima K, Zhou X, Kunkel G, Zhang Z, Deng JM, Behringer RR, de Crombrugghe B: **The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation.** *Cell* 2002, **108**:17-29.
8. Milona MA, Gough JE, Edgar AJ: **Expression of alternatively spliced isoforms of human Sp7 in osteoblast-like cells.** *BMC Genomics* 2003, **4**:43.
9. Bell SM, Schreiner CM, Waclaw RR, Campbell K, Potter SS, Scott WJ: **Sp8 is crucial for limb outgrowth and neuropore closure.** *Proc Natl Acad Sci U S A* 2003, **100**:12195-12200.
10. Treichel D, Schock F, Jackle H, Gruss P, Mansouri A: **mBtd is required to maintain signaling during murine limb development.** *Genes Dev* 2003, **17**:2630-2635.
11. Cohen SM, Jurgens G: **Mediation of Drosophila head development by gap-like segmentation genes.** *Nature* 1990, **346**:482-485.
12. Wimmer EA, Jackle H, Pfeifle C, Cohen SM: **A Drosophila homologue of human Sp1 is a head-specific segmentation gene.** *Nature* 1993, **366**:690-694.
13. Kozak M: **An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs.** *Nucleic Acids Res* 1987, **15**:8125-8148.
14. Philipsen S, Suske G: **A tale of three fingers: the family of mammalian Sp/XKLF transcription factors.** *Nucleic Acids Res* 1999, **27**:2991-3000.
15. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schonbach C, Gojobori T, Baldarelli R, Hill DP, Bult C, Hume DA, Quackenbush J, Schriml LM, Kanapin A, Matsuda H, Batalov S, Beisel KW, Blake JA, Bradt D, Brusic V, Chothia C, Corbani LE, Cousins S, Dalla E, Dragani TA, Fletcher CF, Forrest A, Frazer KS, Gaasterland T, Gariboldi M, Gissi C, Godzik A, Gough J, Grimmond S, Gustincich S, Hirokawa N, Jackson IJ, Jarvis ED, Kanai A, Kawaji H, Kawasawa Y, Kedzierski RM, King BL, Konagaya A, Kurochkin IV, Lee Y, Lenhard B, Lyons PA, Maglott DR, Maltais L, Marchionni L, McKenzie L, Miki H, Nagashima T, Numata K, Okido T, Pavan WJ, Pertea G, Pesole G, Petrovsky N, Pillai R, Pontius JU, Qi D, Ramachandran S, Ravasi T, Reed JC, Reed DJ, Reid J, Ring BZ, Ringwald M, Sandelin A, Schneider C, Semple CA, Setou M, Shimada K, Sultana R, Takenaka Y, Taylor MS, Teasdale RD, Tomita M, Verardo R, Wagner L, Wahlestedt C, Wang Y, Watanabe Y, Wells C, Wilming LG, Wynshaw-Boris A, Yanagisawa M, Yang I, Yang L, Yuan Z, Zavolan M, Zhu Y, Zimmer A, Carninci P, Hayatsu N, Hirozane-Kishikawa T, Konno H, Nakamura M, Sakazume N, Sato K, Shiraki T, Waki K, Kawai J, Aizawa K, Arakawa T, Fukuda S, Hara A, Hashizume W, Imotani K, Ishii Y, Itoh M, Kagawa I, Miyazaki A, Sakai K, Sasaki D, Shibata K, Shinagawa A, Yasunishi A, Yoshino M, Waterston R, Lander ES, Rogers J, Birney E, Hayashizaki Y: **Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.** *Nature* 2002, **420**:563-573.

16. Wimmer EA, Frommer G, Purnell BA, Jackle H: **buttonhead and D-Sp1: a novel Drosophila gene pair.** *Mech Dev* 1996, **59:**53-62.

17. Beermann A, Aranda M, Schroder R: **The Sp8 zinc-finger transcription factor is involved in allometric growth of the limbs in the beetle Tribolium castaneum.** *Development* 2004, **131:**733-742.

18. Narayan VA, Kriwacki RW, Caradonna JP: **Structures of zinc finger domains from transcription factor Sp1. Insights into sequence-specific protein-DNA recognition.** *J Biol Chem* 1997, **272:**7801-7809.

19. Shea CM, Edgar CM, Einhorn TA, Gerstenfeld LC: **BMP treatment of C3H10T1/2 mesenchymal stem cells induces both chondrogenesis and osteogenesis.** *J Cell Biochem* 2003, **90:**1112-1127.

20. Katagiri T, Yamaguchi A, Ikeda T, Yoshiki S, Wozney JM, Rosen V, Wang EA, Tanaka H, Omura S, Suda T: **The non-osteogenic mouse pluripotent cell line, C3H10T1/2, is induced to differentiate into osteoblastic cells by recombinant human bone morphogenetic protein-2.** *Biochem Biophys Res Commun* 1990, **172:**295-299.

21. Wang EA, Israel DI, Kelly S, Luxenberg DP: **Bone morphogenetic protein-2 causes commitment and differentiation in C3H10T1/2 and 3T3 cells.** *Growth Factors* 1993, **9:**57-71.

22. Schock F, Purnell BA, Wimmer EA, Jackle H: **Common and diverged functions of the Drosophila gene pair D-Sp1 and buttonhead.** *Mech Dev* 1999, **89:**125-132.

23. Subramaniam M, Harris SA, Oursler MJ, Rasmussen K, Riggs BL, Spelsberg TC: **Identification of a novel TGF-beta-regulated gene encoding a putative zinc finger protein in human osteoblasts.** *Nucleic Acids Res* 1995, **23:**4907-4912.

24. Blok LJ, Kumar MV, Tindall DJ: **Isolation of cDNAs that are differentially expressed between androgen-dependent and androgen-independent prostate carcinoma cells using differential display PCR.** *Prostate* 1995, **26:**213-224.

25. Tau KR, Hefferan TE, Waters KM, Robinson JA, Subramaniam M, Riggs BL, Spelsberg TC: **Estrogen regulation of a transforming growth factor-beta inducible early gene that inhibits deoxyribonucleic acid synthesis in human osteoblasts.** *Endocrinology* 1998, **139:**1346-1353.

26. Choi KY, Lee SW, Park MH, Bae YC, Shin HI, Nam S, Kim YJ, Kim HJ, Ryoo HM: **Spatio-temporal expression patterns of Runx2 isoforms in early skeletogenesis.** *Exp Mol Med* 2002, **34:**426-433.

27. Banerjee C, Javed A, Choi JY, Green J, Rosen V, van Wijnen AJ, Stein JL, Lian JB, Stein GS: **Differential regulation of the two principal Runx2/Cbfa1 n-terminal isoforms in response to bone morphogenetic protein-2 during development of the osteoblast phenotype.** *Endocrinology* 2001, **142:**4026-4039.

28. Ducy P, Zhang R, Geoffroy V, Ridall AL, Karsenty G: **Osf2/Cbfa1: a transcriptional activator of osteoblast differentiation.** *Cell* 1997, **89:**747-754.

29. Hollnagel A, Oehlmann V, Heymer J, Ruther U, Nordheim A: **Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells.** *J Biol Chem* 1999, **274:**19838-19845.

30. Edgar AJ, Dover SL, Lodrick MN, McKay IJ, Hughes FJ, Turner W: **Bone Morphogenetic Protein-2 Induces Expression Of Murine Zinc Finger Transcription Factor ZNF450.** *J Cell Biochem* in press.

31. Prince V: **The Hox Paradox: More complex(es) than imagined.** *Dev Biol* 2002, **249:**1-15.

32. Xynos ID, Edgar AJ, Buttery LD, Hench LL, Polak JM: **Ionic products of bioactive glass dissolution increase proliferation of human osteoblasts and induce insulin-like growth factor II mRNA expression and protein synthesis.** *Biochem Biophys Res Commun* 2000, **276:**461-465.

33. Gough JE, Christian P, Scotchford CA, Rudd CD, Jones IA: **Synthesis, degradation, and in vitro cell responses of sodium phosphate glasses for craniofacial bone repair.** *J Biomed Mater Res* 2002, **59:**481-489.

34. Harris SA, Enger RJ, Riggs BL, Spelsberg TC: **Development and characterization of a conditionally immortalized human fetal osteoblastic cell line.** *J Bone Miner Res* 1995, **10:**178-186.

35. Murphy CL, Sambanis A: **Effect of oxygen tension and alginate encapsulation on restoration of the differentiated phenotype of passaged chondrocytes.** *Tissue Eng* 2001, **7:**791-803.

36. Edgar AJ, Polak JM: **Cloning and tissue distribution of three murine alpha/beta hydrolase fold protein cDNAs.** *Biochem Biophys Res Commun* 2002, **292:**617-625.