Identification of a Novel Human Zinc Finger Gene, *ZNF438*, with Transcription Inhibition Activity

Zhaomin Zhong[1,#], Bo Wan[1,#], Yun Qiu[2], Jun Ni[1], Wenwen Tang[1], Xinya Chen[1], Yun Yang[1],
Suqin Shen[1], Ying Wang[1], Meirong Bai[1], Qingyu Lang[1] and Long Yu[1,*]

[1]State Key Laboratory of Genetic Engineering, School of Life Science, Fudan University, Shanghai, 200433, People's Republic of China
[2]Anesthesiology Department, Suzhou Yongding Hospital, Wujiang, 2152000, People's Republic of China

There were many different families of zinc finger proteins that contained multiple cysteine and/or histidine residues and used zinc to stabilize their folds. The classical C2H2 zinc finger proteins were the founding members of this superfamily and were among the most abundant proteins in eukaryotic genomes. C2H2 proteins typically contained several C2H2 fingers that made tandem contacts along the DNA. Here we reported a novel C2H2 type zinc finger gene, *ZNF438*, which encoded 828 amino acids that formed five zinc finger domains. Bioinformatics analysis revealed that the *ZNF438* was mapped to human chromosome 10p11.2 and shared 62% identity with rat and mouse homologues. RT-PCR analysis indicated that it was ubiquitously expressed in 18 human adult tissues. With immunofluorescence assay, it was shown that the exogenous Flag-tagged ZNF438 was located in nucleus of COS-7 cells. To further explore the function of ZNF438, we examined the transcriptional activity of ZNF438 protein by transfecting recombinant pM-ZNF438 into mammalian cells. The subsequent analysis based on the duel luciferase assay system showed that ZNF438 was a transcriptional repressor.

Keywords: C2H2, Transcription inhibition, Transcriptional repressor, Zinc finger, *ZNF438*,

[#]These authors contributed equally to this work.

*To whom correspondence should be addressed.
Tel: 86-21-65643954; Fax: 86-21-65643404
E-mail: longyu@fudan.edu.cn

Introduction

The zinc finger domain was first recognized as a repeated zinc-binding domain that contained conserved cysteine and histidine residues in *Xenopus laevis* transcription factor IIIA (TFIIIA) (Brown *et al.*, 1985; Miller *et al.*, 1985). Ever since then, numerous other zinc-binding domains had been identified and designated as zinc fingers. Zinc fingers constituted one of the most common structural domains in eukaryotes. According to their structural diversities, they were classified into eight groups: C2H2-like finger, Gag knuckle, treble clef finger, zinc ribbon, Zn2/Cys6-like finger, TAZ2 domain-like, short zinc binding loops and metallothioneins (Krishna *et al.*, 2003). Zinc fingers were not only structurally diverse, but they also had distinct functions. They were involved in nucleic acid and lipid binding as well as in protein-protein interactions (Grishin, 2001). By performing these functions, zinc fingers played an important role in many fundamental cellular processes, such as replication and repair, transcription and translation, metabolism and signaling, cell proliferation and apoptosis.

The C2H2 zinc finger family belonged to the C2H2-like finger group according to Krishna's classifying method, and represented the classical type of zinc fingers. All C2H2 zinc fingers typically contained a repeated 30 amino acid sequence, which folded in the presence of zinc to form a compact βßα domain (Wolfe *et al.*, 2000). The zinc was coordinated between two cysteines at one end of the β-sheet and two histidines in the C-terminal portion of the α-helix.

Proteins containing C2H2 zinc fingers were quite common in eukaryotic organisms. In the human genome, approximately 700 C2H2 type zinc finger genes were found, making it the second most frequently occurring domain in the human proteome (Venter *et al.*, 2001). These genes played important roles during differentiation and development through their involvement in DNA binding, causing the specific activation or repression of gene expression, The DNA binding activity

of C2H2 zinc finger proteins had been studied in considerable detail. DNA recognition usually required 2 to 4 tandemly arranged zinc fingers; if only one or two finger were present, other secondary structure elements were generally used to augment DNA recognition. Some zinc finger proteins functioned independently as the major regulator of a set of genes, such as *TFIIIA* (Shastry, 1996) and *NRSF/REST* (Schoenherr and Anderson, 1995), while others worked cooperatively with other DNA binding proteins, such as the members of SP1 family (Lania *et al.*, 1997). However, the zinc finger domain was used not only for protein-DNA but also for protein-RNA and protein-protein interactions (Lania, 1996; Mackay and Crossley, 1998).

In this report, we described the isolation of a novel human C2H2 zinc finger gene *ZNF438*. Genomic structure analysis indicated that this gene contained 7 exons and was localized to human chromosome 10p11.2. By RT-PCR it was found that *ZNF438* was ubiquitously expressed in all human adult tissues. Subcellular localization revealed that ZNF438 was localized in nucleus when overexpressed in COS-7 cells. Our study also presented that ZNF438 protein possessed transcription inhibition activity and might play an important role as a transcription factor.

Materials and Methods

Cloning and sequencing of *ZNF438* cDNA. According to the sequences flanking the predicted open reading frame (ORF), two pairs of primers ZNF-clone-F1/ZNF-clone-R1 and ZNF-clone-F2/ ZNF-clone-R2 (Table 1) were designed to amplify *ZNF438* from the brain cDNA library. The PCR reaction was performed as follows: initial 5 min at 94°C and then 35 cycles of 60 s at 94°C; 60 s at 53°C, 120 s at 72°C, and a final 5 min at 72°C. Two amplified fragments, F1R1 and F2R2, were obtained and subsequently integrated into a cDNA fragment with another PCR.

The final PCR product was then subcloned into the pMD-18-T vector (Takara) and was sequenced. The generated sequences had been deposited to GenBank Data Library under Accession No. AF428258.

Bioinformatics analyses. DNA and protein sequence comparisons were carried out by BLAST at NCBI Web server (http://www.ncbi. nlm.nih.gov/BLAST/). Multiple sequence alignment was performed using the CLUSTALX and GENDOC program. Chromosomal localization and genomic organization prediction were carried out using the UCSC Genome Browser (http://www.genome.ucsc.edu/ cgi-bin/hgBlat). Motif analysis was performed at SMART server (http://smart.embl-heidelberg.de/). SUMOplot™ (http://www.abgent. com/doc/sumoplot) was used to predict the probability for the SUMO consensus sequence to be engaged in SUMO attachment.

Plasmid construction. All inserted fragments were obtained by PCR amplification using the primers listed in Table 1. Full length *ZNF438* was subcloned into the eukaryotic expression vector pFlag-CMV at *Hind*III and *Xba*I sites using primers pFlag-CMV-ZNF438-F and pFlag-CMV-ZNF438-R. The full-length of *ZNF438* was subcloned into eukaryotic expression vector pM at *Bam*HI and *Xba*I sites using primers pM-ZNF438-F and pM-ZNF438-R. After confirmation by sequencing, all eukaryotic expression plasmids were extracted and purified for transfection using Endofree Plasmid Maxi Kit (Qiagen).

RT-PCR. The human multiple tissue cDNA (MTC) panels (Clontech) were used as PCR templates according to the manufacture's protocol. The specific RT-PCR primers of *ZNF438* and *β2-MG* were listed in Table 1. 32 cycles for *ZNF438* and 26 cycles for *β2-MG* of amplification (30 s at 94°C; 30 s at 54°C, 40 s at 72°C) were performed using DNA polymerase. The PCR products of *ZNF438* and *β2-MG* were then electrophoresed on 2% aggrose gel.

Cell culture, transfection and western blotting. COS-7 and H1299 cells were grown in RPMI 1640 medium (GIBCO BRL)

Table 1. Nucleotide sequence of oligonucleotides used for cloning of ZNF438 and construction of expression vectors

Primers	Sequence
cDNA cloning primers	
ZNF-clone-F1	5'-TGCAAGAACAGTAAAAGGAAAG-3'
ZNF-clone-R1	5'-TGTATTGGAGAAGCCAGGGAG-3'
ZNF-clone-F2	5'-AAAGAGAAAAGGAAGAAAACGG-3'
ZNF-clone-R1	5'-TGGGGTCTCATTTCTCAGCTTC-3'
RT-PCR primers	
ZNF438-RT-F	5'-TAAGTCAAATCACTTCAGGAAGC-3'
ZNF438-RT-R	5'-AGGGATGTTTGATTCACCTTCATC-3'
β2-MG-RT-F	5'-ATGAGTATGCCTGCCGTGTGAAC-3'
β2-MG-RT-R	5'-TGTGGAGCAACCTGCTCAGATAC-3'
Gene expression primers	
pFlag-CMV-ZNF438-F	5'-TTCAAGCTTATGCAGAATTCTGTATCAGTAC-3'
pFlag-CMV-ZNF438-R	5'-AACTCTAGATCATTTCTCAGCTTCACTGG-3'
pM-ZNF438-F	5'-AAAGGATCCTGCAGAATTCTGTATCAGTAC-3'
pM-ZNF438-R	5'-AAGTCTAGATCATTTCTCAGCTTCACTGG-3'

with 10% fetal calf serum (GIBCO BRL); HEK293T cells were grown in DMEM medium (GIBCO BRL) supplemented with 10% fetal calf serum. The cultured HEK293T or COS-7 cells were seeded. After overnight growth the cells were transfected with expression vectors using Lipofectamine™ (Invitrogen) according to the manufacturer's protocol. Protein samples separated by SDS-PAGE were electro-transferred onto nitrocellulose membrane. The membrane was blocked at room temperature for 1 h with PBS containing 5% (w/v) BSA, and then washed with PBS containing 0.2% Tween-20 (Sigma; PBST). After blocking, the membrane was then incubated overnight at 4°C with primary antibody diluted with PBS containing 3% (w/v) BSA. The membrane was washed 3 times with PBST then incubated with HRP-conjugated secondary antibody at room temperature for 1 h. The membrane was washed 3 times again with PBST and then developed with ECL system (Santa Cruz Biotechnology).

Transcriptional function assay of ZNF438. The Gal4 fusion plasmid pM-ZNF438 was co-transfected with 100 ng of reporter plasmid pGal45tkLUC into 1×10^5 HEK293T or H1299 cells on 24-well plates (Falcon), and luciferase activity was measured with Lumat LB9507 (Berthold). Moreover, 20 ng of plasmid pRL-SV40 (Promega) encoding *Renilla* luciferase was used as an internal control in each transfection. The plasmids pM3-VP16 and pM-ZNF191 were used as activation and repression controls. The assay system of dual-luciferase reporter (Promega) was used to assay the luciferase activity according to the manufacturer's protocol. Transfection of each construct was repeated four times in one assay and each assay was repeated at least three times. The relative luciferase activity represented the value of the firefly luciferase versus that of *Renilla* luciferase, and all data were analyzed by statistics method (t-test).

Immunofluorescence analysis of ZNF438 subcellular localization. COS-7 cells were plated on cover slips and transfected with 900 ng of respective expression vectors pFlag-CMV-ZNF438 with Lipofectamine (Invitrogen) according to the manufacturer's protocol. 36 h after transfection, the cells were fixed in 4% paraformaldehyde for 10 min. The cells were washed once with TBS [50 mM Tris-HCl (pH7.4), 150 mM NaCl], then permeabilized with 0.2% Triton X-100 for 5 min; washed three times in TBS followed by incubated in fresh 0.1% sodium borohydride in TBS for 5 min to quench endogenous fluorescence. After blocked in blocking buffer (1% horse serum, 1% BSA, 1 × PBS) for 1 h, the cells were incubated with Flag-antibody overnight at 4°C then washed three times in TBS and incubate with goat anti-mouse IgG Fluorescein-conjugate antibody (Upstate Biotechnology) for 50 min in dark. After washed two times in TBS, the slips were incubated with DAPI (4, 6-diamidino-2-phenylindole) for 15 min and wash another two times in TBS and drained. Microscopy observation was carried out under fluorescent light at relevant wavelengths.

Results and Discussion

Molecular cloning and characterization of human *ZNF438*. When searching human EST databases for sequences with

homology to the conserved Cys2/His2 zinc finger motif, we obtained a series of human ESTs and assembled them into a contig of 3151 nucleotides. This contig contained an open reading frame (ORF) of 2487 nucleotides. To verify the contig, PCR primers (ZNF-clone-F1/ZNF-clone-R1, ZNF-clone-F2/ZNF-clone-R2) (Table 1) were designed to perform PCR in a human brain cDNA library. Two amplified fragments, F1R1 and F2R2, were obtained and subsequently integrated into a cDNA fragment. The final PCR product was then subcloned into the pMD-18-T vector (Takara) and was sequenced. The sequencing result verified the contig sequence, and this sequence was subsequently deposited in GenBank with the GenBank Accession No. AF428258 and was later named ZNF438 by the Human Gene Nomenclature Committee. As shown in Fig. 1, the cDNA sequence had an intact ORF extending from nucleotide 331 to 2817, and it encoded a putative protein of 828 amino acids. An in-frame stop codon upstream of the initiator, and a typical polyadenylation signal, AATAA, was found at two distinct sites: nt 2974-2979 and 3068-3073.

The molecular weight of the putative protein was predicted to be 91.8 kDa. After searching the SMART database (http://smart.embl-heidelberg.de/), we found that the protein contained five C2H2 zinc finger domains at the C-terminus (Fig. 1). The first two domains belonged to the Krüppel type; the other three were non-Krüppel type. ZNF438 lacked some N-terminal domains that were present in other zinc finger proteins such as POZ, KRAB, SCAN, FAX, FAR and ZiN (Knochel *et al.*, 1989; Bellefroid *et al.*, 1993; Numoto *et al.*, 1993; Bardwell and Treisman, 1994; Zollman *et al.*, 1994; Williams *et al.*, 1999). These domains seemed to be involved in transcriptional regulation and protein-protein interactions. However, some zinc finger proteins, such as the GC box binding transcription factor Sp3 protein, lacked those conserved domain, but still displayed transcriptional activity (Hagen *et al.*, 1994; De Luca *et al.*, 1996; Liang *et al.*, 1996; Kumar and Butler, 1997; Noti, 1997; Suske, 1999; Ross *et al.*, 2002).

The multiple sequence alignments indicated that the putative human ZNF438 protein exhibited 98.4% identity to chimpanzee (XP_001138402), 96.4% identity to orangutan (CAH90705), and 62% identity to the mouse (NP_848837) and rat (XP_001056328) homologous protein, respectively. And the alignments showed us that human ZNF438 protein had the highest similarity to chimpanzee and orangutan homologous protein, while the mouse homologous protein was closest to the rat homologous protein (Fig. 2A). We analyzed subsequently the evolutionary relationship among different zinc finger proteins with cladogenesis tree analysis (Fig. 2B), which showed the relationship among the ZNF438 homologous proteins more clearly. Furthermore, the N-terminal sequence of ZNF438 was more conserved than the middle sequence when compared to the mouse and rat homologues. This suggested that there might be another conserved domain at the N-terminal that remained to be identified.
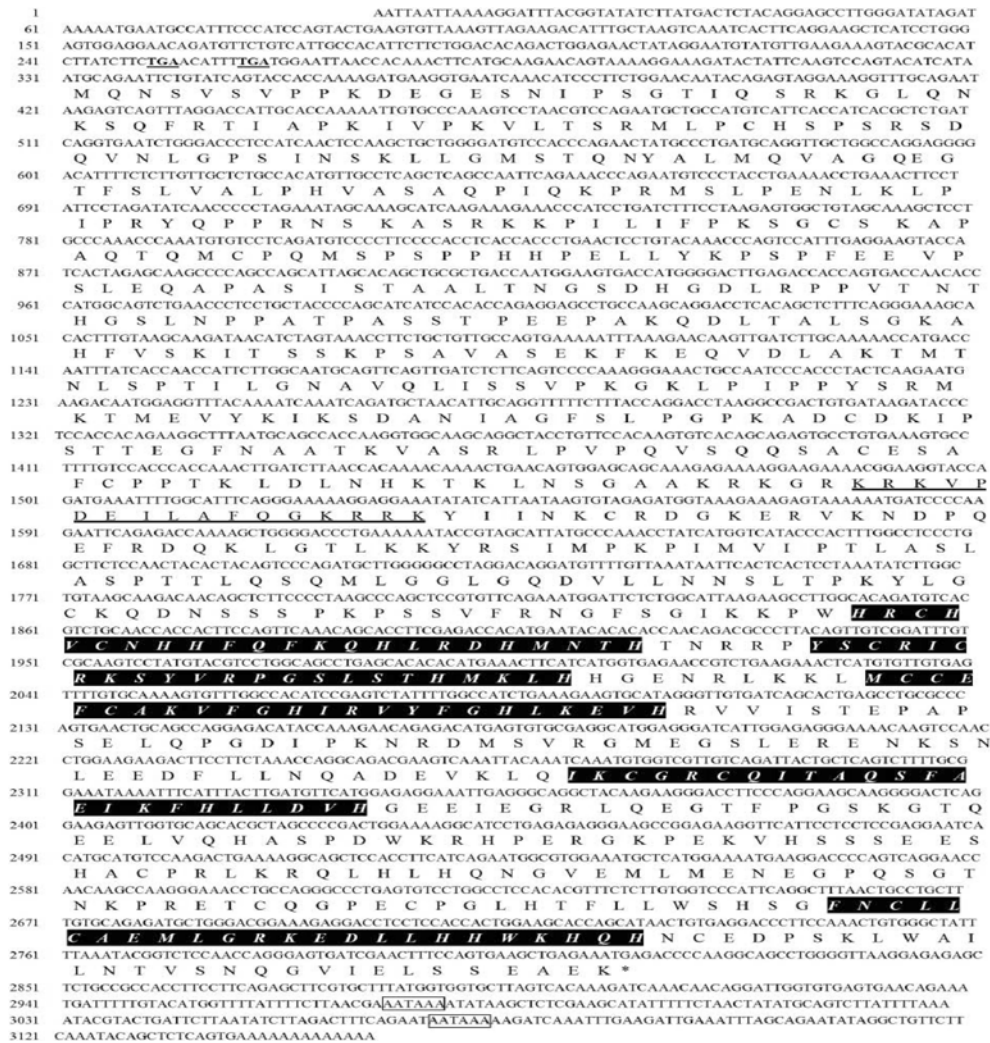
```
   1                                                    AATTAATTAAAAGGATTTACGGTATATCTTATGACTCTACAGGGAGCCTTGGGATATAGAT
  61  AAAAATGAATGCCATTTCCCATCCAGTACTGAAGTGTTAAAGTTAGAAGACATTTGCTAAGTCAAATCACTTCAGGAAGCTCATCCTGGG
 151  AGTGGAGGAACAGATGTTCTGTCATTGCCACATTCTTCTGGACACAGACTGGAGGAACTATAGGAATGTATGTTGAAGAAGTACGCACAT
 241  CTTATCTTCTGAACATTTTGATGGAATTAACCACAAACTTCATGCAAGAAGATACTATTCAAGTCCAGTACATCATA
 331  ATGCAGAATCTCTGTATCAGTACCACCAAAAGATGAAGGTGAATCAAACATCCCTTCTGGAACATACAGAGTAGGAAAGGTTTGCAGAAT
       M  Q  N  S  V  S  V  P  P  K  D  E  G  E  S  N  I  P  S  G  T  I  Q  S  R  K  G  L  Q  N
 421  AAGAGTCAGTTTAGGACCATTGCACCAAAAATTGTGCCCAAAGTCCTAACGTCCAGAATGCTGCCATGTCATTCACCATCACGCTCTGAT
       K  S  Q  F  R  T  I  A  P  K  I  V  P  K  V  L  T  S  R  M  L  P  C  H  S  P  S  R  S  D
 511  CAGGTGAATCTGGGACCCTCCATCAACTCCAAGCTGCTGGGGATGTCCACCCAGAACTATGCCCTACCTGAAAACCTGAAACTTCCT
       Q  V  N  L  G  P  S  I  N  S  K  L  L  G  M  S  T  Q  N  Y  A  L  M  Q  V  A  G  Q  E  G
 601  ACATTTTCTCTTGTTGCTCTGCCACATGTTGCCTCAGCTCAGCCAATCCATCCTGATCTTTCCTAAGGATGGCTGTAGCAAAGCTCCT
       T  F  S  L  V  A  L  P  H  V  A  S  A  Q  P  I  Q  K  P  R  M  S  L  P  E  N  L  K  L  P
 691  ATTCCTAGATATCAACCCCCTAGAAATAGCAAAGCATCAAGAAAGAAAACCATCCTGATCTTTCCTAAGGATGGCTGTAGCAAAGCTCCT
       I  P  R  Y  Q  P  P  R  N  S  K  A  S  R  K  K  P  I  L  I  F  P  K  S  G  C  S  K  A  P
 781  GCCCAAACCCAAATGTGTCCTCAGATGTCCCCTTCCCCACCTCACCACCCTGAACTCCTGTACAAACCCAGTCCATTGAGGAAGTACCA
       A  Q  T  Q  M  C  P  Q  M  S  P  S  P  P  H  H  P  E  L  L  Y  K  P  S  P  F  E  E  V  P
 871  TCACTAGAGCAAGCCCCAGCCAGCATTAGCACAGCTGCGCTGACCAATGGAAGTGACCATGGGGACTTGAGACCACCAGTGACCAACACC
       S  L  E  Q  A  P  A  S  I  S  T  A  A  L  T  N  G  S  D  H  G  D  L  R  P  P  V  T  N  T
 961  CATGGCAGTCTGAACCCTCCTGCTACCCCAGCATCATCACACCAGAGGAGCCTGCCAAGCAGGACCTCACAGCTCTTTCAGGGAAAGCA
       H  G  S  L  N  P  P  A  T  P  A  S  S  T  P  E  E  P  A  K  Q  D  L  T  A  L  S  G  K  A
1051  CACTTTGTAAGCAAGATAACATCTAGTAAAACCTTCTGCTGTTGCCGATGAAAAATTTAAAGAACAAGTTGATCTTGCAAAAACCATGACC
       H  F  V  S  K  I  T  S  S  K  P  S  A  V  A  S  E  K  F  K  E  Q  V  D  L  A  K  T  M  T
1141  AATTTATCACCAACCATTCTTGGCAATGCAGTTCAGTTGATCTCTTCAGTCCCCAAAGGGAAACTGCCAATCCCACCCTACTCAAGAATG
       N  L  S  P  T  I  L  G  N  A  V  Q  L  I  S  S  V  P  K  G  K  L  P  I  P  P  Y  S  R  M
1231  AAGACAATGGAGGTTTACAAAATCAAATCAGATGCTAACATTGCAGGATTTTCTTTACCAGGGACCTAAGGCCTGTGATAAGATACCC
       K  T  M  E  V  Y  K  I  K  S  D  A  N  I  A  G  F  S  L  P  G  P  K  A  D  C  D  K  I  P
1321  TCCACCACAGAAGGCTTTAATGCAGCCACCAAGGTGGCAAGCAGGCTACCTGTTCCACAAGTGTCACAGCAGAGTGCCTGTGAAAGTGCC
       S  T  T  E  G  F  N  A  A  T  K  V  A  S  R  L  P  V  P  V  S  Q  Q  S  A  C  E  S  A
1411  TTTTGTCCACCCACCAAACTTGATCTTAACCACAAAACAAAACTGAACAGTGGAGCAGCAAAGAGAAAAGGAAGAAAACGGAAGGTACCA
       F  C  P  P  T  K  L  D  L  N  H  K  T  K  L  N  S  G  A  A  K  R  K  G  R  *K  R  K  V  P*
1501  GATGAAAATTTTGGCATTTCAGGGAAAAAAGGAGGGAAATATATCATTAATAAGTGTAGAGATGGTAAAGAAAGAGTAAAAAATGATCCCCAA
       *D  I  L  A  F  Q  G  K  R  R  K* Y  I  I  N  K  C  R  D  G  K  E  R  V  K  N  D  P  Q
1591  GAATTCAGAGACCAAAAGCTGGGGACCCTGAAAAAATACCGTAGCATTATGCCCAAACCTATCATGGTCATACCCACTTTGGCCTCCCTG
       E  F  R  D  Q  K  L  G  T  L  K  K  Y  R  S  I  M  P  K  P  I  M  V  I  P  T  L  A  S  L
1681  GCTTCTCCAACTACACTACAGTCCCAGATGCTTGGGGGCCTAGGACAGGATGTTTTGTTAAATAATTCACTCACTCCTAAGTATCTTGGA
       A  S  P  T  T  L  Q  S  Q  M  L  G  G  L  G  Q  D  V  L  L  N  N  S  L  T  P  K  Y  L  G
1771  TGTAAGCAAGACAACAGCTCTTCCCCTAAGCCCAGTTCCGTGTTCAGGAATGGATTCTCTGGCATTAAGAAGCCTTGGCACAGATGTCAC
       C  K  Q  D  N  S  S  P  K  P  S  S  V  F  R  N  G  F  S  G  I  K  K  P  W  *H  R  C  H*
1861  GTCTGCAACCACCACTTCCAGTTCAAACAGCACCTTCGAGACCACATGAATACACCAACAGACGCCCTTACAGTTGTCGGATTTGT
       *V  C  N  H  H  F  Q  F  K  Q  H  L  R  D  H  M  N  T  H*  T  N  R  R  P  *Y  S  C  R  I  C*
1951  CGCAAGTCCTATGTACGTCCTGGCAGCCTGAGCACACACATGAAACTTCATCATGGTGAGAACCGTCTGAAGAAGCTTGTTGTTGTGAG
       *R  K  S  Y  V  R  P  G  S  L  S  T  H  M  K  L  H* H  G  E  N  R  L  K  K  L  *M  C  C*
2041  TTTTGTGCAAAAGTGTTTGGCCACATCCGAGTCTATTTGGCCATCTGAAAGAAGTGCATAGGGTTGTGATCAGCACAGATGTCGCGCCC
       *F  C  A  K  V  F  G  H  I  R  V  Y  F  G  H  L  K  E  V  H*  R  V  V  I  S  T  E  P  A  P
2131  AGTGAACTGCAGCCAGGAGACATACCAAAGAACAGAGACATGAGTGTGCGAGGCATGGAGGGATCATTGGAGAGGGAAAACAAGTCCAAC
       S  E  L  Q  P  G  D  I  P  K  N  R  D  M  S  V  R  G  M  E  G  S  L  E  R  E  N  K  S  N
2221  CTGGAAGAACTTCCTTCTAAACCAGGCAGACGAAGTCAAATTACAAATCAAATCTGGTCGTTGTCAGATTACCTCAGTCTTTTGCG
       L  E  E  D  F  L  L  N  Q  A  D  E  V  K  L  Q  *I  K  C  G  R  C  Q  I  T  A  Q  S  F  A*
2311  GAAATAAAATTTCATTTACTTGATGTTCATGGAGAAGAAATAGAAGGACGTTTCCCAGGAAAGCAAGGGGACTCAG
       *E  I  K  F  H  L  L  D  V  H*  G  E  E  I  E  G  R  L  Q  E  G  T  F  P  G  S  K  G  T  Q
2401  GAAGAGTTGGTGCAGCACGCCCCGACTGGAAGCCGCCATCCTGAGAGAGGGAAGCCGGAGAAAGTTCATTCCTCCTCCGAGGAATCA
       E  E  L  V  Q  H  A  S  P  D  W  K  R  H  P  E  R  G  K  P  E  K  V  H  S  S  E  E  S
2491  CATGCATGTCCAAGACTGAAAAGGCAGCTCCACCTTCATCAGAATGGCGTGGAAATGCTCATGGAAAATGAAGGACCCCAGTCAGGAACC
       H  A  C  P  R  L  K  R  Q  L  H  L  H  Q  N  G  V  E  M  L  M  E  N  E  G  P  G  S  F
2581  AACAAGCCAAGGGAAACCTGCCAGGGCCCTGAGTGTCCTGGCCTCCACACGTTTCTTTGTGGTCCCATTCAGGCTTTAACTGCCTGCTT
       N  K  P  R  E  T  C  Q  G  P  E  C  P  G  L  H  T  F  L  L  W  S  H  S  G  *F  N  C  L  H*
2671  TGTGCAGAGATGCTGGGACGGAAAGAGGACCTCCTCCACCACTGGAAGCACCAGCATAACTGTGAGGACCCTTCCAAACTGTGGGCTATT
       *C  A  E  M  L  G  R  K  E  D  L  L  H  H  W  K  H  Q  H*  N  C  E  D  P  S  K  L  W  A  I
2761  TTAAATACGGTCTCCAGCCAAGGGGATGATGATGTCCGAACTTTCCAGTAGACAGAGCCAAAGATCAAACAACAGGATTGGGTGAGTGAACAGAAA
       L  N  T  V  S  N  Q  G  V  I  E  L  S  S  E  A  E  K  *
2851  TCTGCGACGAATTCCTTCAGAGCTTCGTGCTTTATGGTGGTGGTCCTTAGTCACAAAGATCAAACAACAGGATTGGGTGAGTGAACAGAAA
2941  TGATTTTGTACATGGTTTTATTTTCTTAACGA[AATAAA]ATATAAGCTCTCGAAGCATATTTTTCTAACTATATGCAGTCTTATTTTAAA
3031  ATACGTACTGATTCTTAATATCTTAGACTTTCAGAAT[AATAAA]AAGATCAAATTTGAAGATTGAAATTTAGCAGAATATAGGCTGTTCTT
3121  CAAATACAGCTCTCAGTGAAAAAAAAAAAAAA
```

Fig. 1. Nucleotide and deduced amino acid sequences of human *ZNF438* gene with its 5' flanking region. The nucleotide of 3151bp cDNA was shown in the top lines and its deduced amino acid sequence was shown below. The ORF extends from nucleotide 331 to 2817, and encoded a protein of 828 amino acids. The NLS sequence and two stop codons (TGA) upstream of the ORF at the same reading frame were underlined. The poly adenylation signal (AATAAA) was boxed. The sequence of Znf domains were indicated by white, italic letters. Asterisk represents the stop codon.

Further sequence analysis indicated that the putative human ZNF438 had seven highly probable SUMO binding sites (Table 2). SUMO was a small ubiquitin-related modifier that got attached to lysine residues in target proteins via an isopeptide linkage in a process that is analogous to ubiquitination (Melchior, 2000). SUMO modification was a dynamic, reversible reaction, and distinct enzymes were responsible for adding or removing it from target proteins. Several transcription factors, such as the androgen receptor, c-jun, p53, NF-κB, and Sp3, had been found to be modified by sumoylation (Gostissa *et al.*, 1999; Poukka *et al.*, 1999; Rodriguez *et al.*, 1999; Muller *et al.*, 2000; Poukka *et al.*, 2000; Rodriguez *et al.*, 2001; Ross *et al.*, 2002), which causes the repression of their transcriptional activation. We speculated that ZNF438 would also undergo such post-translational modification based on the presence of the SUMO sites.

Chromosomal localization and genomic organization of human *ZNF438*. With the completion of the draft sequence of the human genome, it was possible to localize a gene with known sequence to a specific site on a chromosome. By searching the UCSC genomic database (http://genome.ucsc.edu), the human *ZNF438* was localized to 10p11.2. Numbers of well-known gene, such as *PAPD1, MAP3K8, LYZL2* and *TCF8,* were localized near *ZNF438.* Genomic structure analysis of this gene revealed that it contained 7 exons, and the exon/intron boundaries were all consistent with the AG/GT rule (Fig. 3).

Tissue distribution of human *ZNF438*. With using the human MTC (multiple-tissue cDNA) panels (Clontech), the tissue distribution of *ZNF438* transcription was examined by cycle-limited reversed transcription polymerase chain reaction
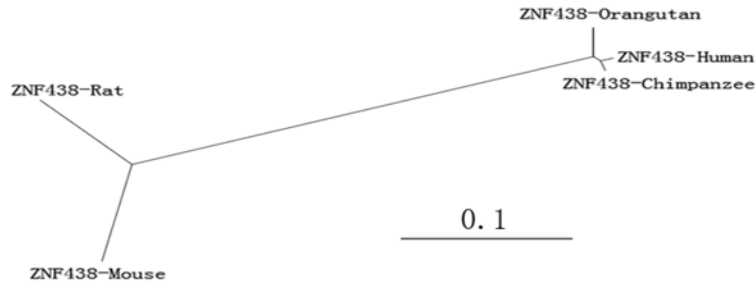
Fig. 2. The alignment of the deduced amino acid sequences of human ZNF438 (AAP97298), chimpanzee homologue (XP_001138402), orangutan homologue (CAH90705), mouse homologue (NP_848837) and rat homologue (XP_001056328). (A) Identity was indicated by black shading and the similarity was indicated by gray shading. Dashes and spaces introduced to optimize the alignment; dots omitted amino acid residues outside of the domain. The sequences were aligned by ClustalX software and viewed by GENEDOC software. (B) Unrooted phylogenetic tree of the five members were constructed by comparing the whole sequences using the ClustalW software.

(RT-PCR). As shown in Fig. 4, *ZNF438* was expressed ubiquitously in the eighteen tissues examined. However, it

Table 2. Putative SUMO acceptor sites of ZNF438

No.	Position	SUMO acceptor sites
1	K309	MEVYK IKSD ANIAG
2	K416	DGKER VKND PQEFR
3	K503	NGFSG IKKP WHRCH
4	K118	SLPEN LKLP IPRYQ
5	K230	TPEEP AKQD LTALS
6	K648	EVKLQ IKCG RCQIT
7	K709	RHPER GKPE KVHSS

was strongly expressed in adult uterus, bladder, colon, ovary, testis, skeletal muscle, heart and brain, while weakly expressed in bone marrow, kidney, liver and stomach.

Transcriptional function analysis of human ZNF438. Several kinds of protein motifs that are important to DNA recognition and binding have been identified in transcription factors, such as zinc finger, helix-loop-helix, helix-turn-helix, and leucine zipper (Busch and Sassone-Corsi, 1990).

Multiple papers showed that the zinc finger motifs existed in numerous transcription factors. Lots of zinc finger transcription factors, such as ZNF191 (Han *et al*., 1999), ZNF396 and ZNF397 (Wu *et al*., 2003), FOXP3 (Lopes *et al.,*
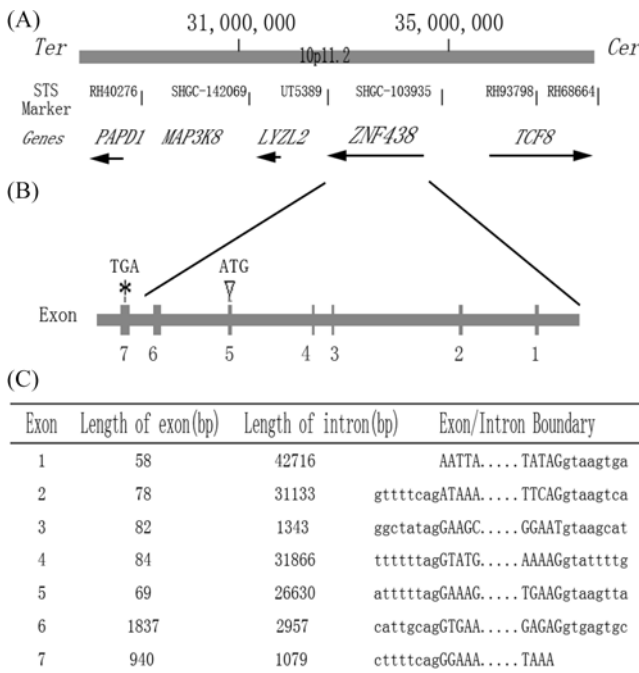
(A)

(B)

(C)

Exon	Length of exon(bp)	Length of intron(bp)	Exon/Intron Boundary
1	58	42716	AATTA.....TATAGgtaagtga
2	78	31133	gttttcagATAAA.....TTCAGgtaagtca
3	82	1343	ggctagGAAGC.....GGAATgtaagcat
4	84	31866	tttttagGTATG.....AAAAGgtattttg
5	69	26630	atttttagGAAAG.....TGAAGgtaagtta
6	1837	2957	cattgcagGTGAA.....GAGAGgtgagtgc
7	940	1079	cttttcagGGAAA.....TAAA

Fig. 3. The chromosomal localization and genomic organization of human *ZNF438*. (A) The chromosomal localization of human *ZNF438*. Near this gene, a number of well-known genes and the STS markers were indicated. (B) *ZNF438* consisted of 7 exons. They were drawn to scale, numbered, and represented by rectangles. The triangle and asterisk indicated the positions of the start and stop codon, respectively. (C) The sequences flanking the exon/intron boundaries were given; intron and exon nucleotide sequences were shown in lower-case and upper-case letter, respectively.

2006), functioned as transcriptional repressors with the putative zinc finger domain providing DNA-binding specificity.

Accordingly, it was reasonable to postulate that the ZNF438 might possess transcription-regulating activity because of its zinc finger motif. With the experiment based on dual luciferase assay system to assay the transcription function, the ORF of *ZNF438* was inserted into mammalian cell expression plasmid pM-vector, and recombinant plasmid pM-ZNF438 encoding the fusion protein Gal4-ZNF438 was cotransfected into different cell lines with reporter plasmid pGal45tkLUC encoding firefly luciferase which was correlated with the transcription-regulating activity of ZNF438, and the plasmid pRL-SV40 encoding *Renilla* luciferase which provided an internal control to serve as the baseline response, respectively.

One negative control, mock pM-vector, and two positive control, pM3-VP16 and pM-ZNF191, were introduced in this system. Among them, VP16 served as transcriptional activator and ZNF191 served as transcriptional repressor (Han *et al.*, 1999). The luciferase activities of samples were determined by Dual-luciferase reporter assay system (Promega, Madison, USA). In the Dual-luciferase reporter assay system, the activities of firefly and *Renilla* luciferase were measured sequentially from a single sample. The relative luciferase activity that normalized the activity of firefly luciferase to the activity of *Renilla* luciferase was to minimize experimental variability caused by differences in cell viability or transfection efficiency, and to eliminate effectively other sources of variability, such as differences in pipetting volumes, cell lysis efficiency and assay efficiency. As shown in Fig. 5, the pM-ZNF438 could significantly inhibit the expression of the firefly luciferase reporter. Compared with that of the negative control, the firefly luciferase activity were down-regulated about 3 and 2 folds in HEK293T and H1299 cells, respectively ($p < 0.001$, n = 4).

Furthermore, we explored whether transcriptional repression of the reporter gene was proportioned to the used amount of ZNF438. In this experiment, we applied a fixed amount of pGal45tkLuc reporter gene with varying amounts of ZNF438 to transfect HEK293T cells. By t-test statistical analysis, it was shown that ZNF438 had dose-dependent expression inhibition of the reporter gene (Fig. 5D). Therefore, it seemed that ZNF438 might play its role as transcriptional repressor in mammalian cells.

Subcellular localization of human ZNF438. After we indicated that the ZNF438 had a transcription repression
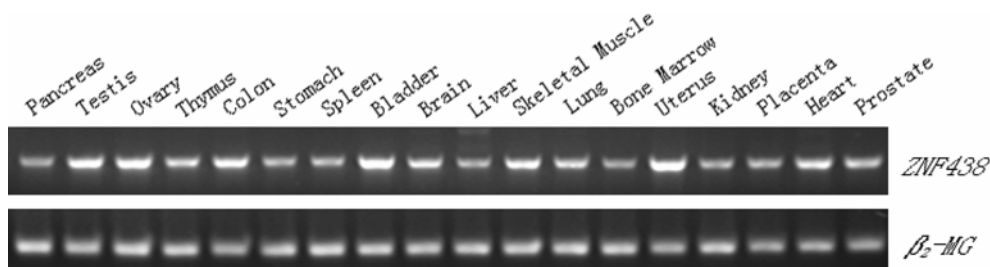

Fig. 4. Expression analysis of *ZNF438* in 18 human tissues by RT-PCR. The PCR amplification was performed using the condition as 94°C for 5 min followed by 32 cycles of 94°C for 30 s, 54°C for 30 s and 72°C for 40 s. As control, the *β₂-MG* gene was also amplified with the conditions as 94°C for 5 min followed by 26 cycles of 94°C for 30 s, 54°C for 30 s and 72°C for 40 s. The PCR products of ZNF438 and β₂-MG were then electrophoresed on 2% aggrose gel. *ZNF438* was found to express ubiquitously in the eighteen tissues examined. However, it was strongly expressed in adult uterus, bladder, colon, ovary, testis, skeletal muscle, heart and brain, while weakly expressed in bone marrow, kidney, liver and stomach.
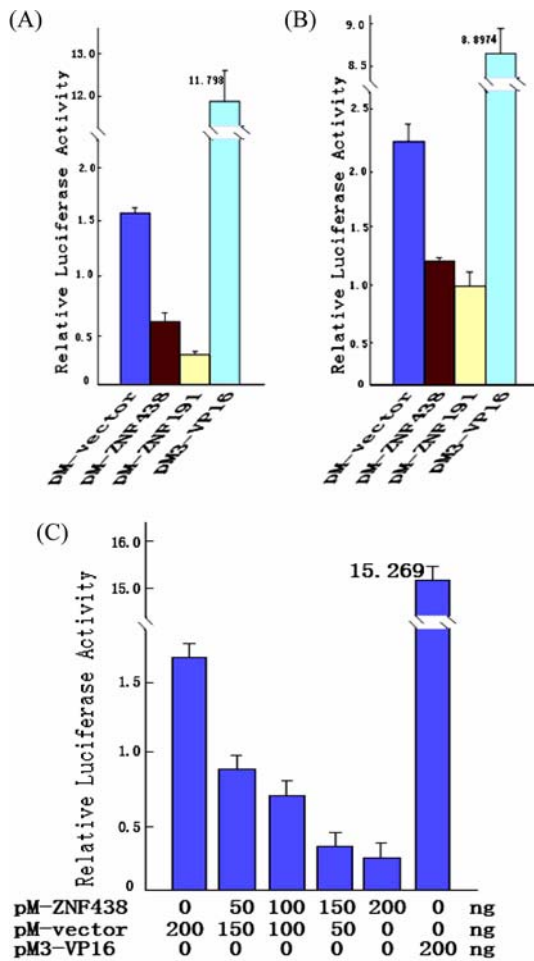
Fig. 5. Transcriptional function analysis of human ZNF438 protein. (A) HEK293T cells were transiently transfected with the pGAL45tkLUC reporter along with the recombinant plasmid pM-ZNF438. (B) H1299 cells were transiently transfected with the pGAL45tkLUC reporter along with the recombinant plasmid pM-ZNF438. (C) Dose-dependent inhibition of pM-ZNF438 to the activity of reporter gene pGAL45tkLUC in HEK293T cells. A fixed amount of reporter gene with varying amounts of ZNF438 was used. In this experiment, relative luciferase activity was normalized by co-transfection with pRL-SV40. Mock pM was used as a negative control, while pM-ZNF191 was a transcription repressor control and pM3-VP16 was a transcription activator control. Each value in this figure represented the mean of the enzyme activities determined in four assays. The *error bars* indicated standard deviation from means.

Fig. 6. Subcelluar localization of ZNF438 protein and western blot. (A) Cells expressing Flag-tagged ZNF438 were stained with monoclonal antibody against Flag epitope. The localization of ZNF438 was shown by the green fluorescence of fluorescein. (B) The nucleus of cells transfected with pFlag-CMV-ZNF438 was stained with DAPI. The nuclear was shown by the blue fluorescence. (C) The combined image of panels (A) and (B) showed that there was a large overlap between the green ZNF428 signal and the blue nuclear signal. (D) Western blot analysis of Flag-tagged ZNF438 expressed in COS-7 cells. The band was approximately 95 kDa, corresponding to the predicted size of ZNF438. Mock pFlag-CMV was used as a control.

function, we further explore the subcellular localization of ZNF438. We hypothesize that, if ZNF438 was a transcription factor, it might be localized in nucleus. In order to test our assumption, indirect immunofluorescence for Flag-ZNF438 was used to explore it subcellular localization. COS-7 cells were transfected with the pFlag-CMV-ZNF438 constructs. Indirect immunofluorescence assay with anti-Flag antibody showed that the Flag-ZNF438 was localized in nucleus (Fig.
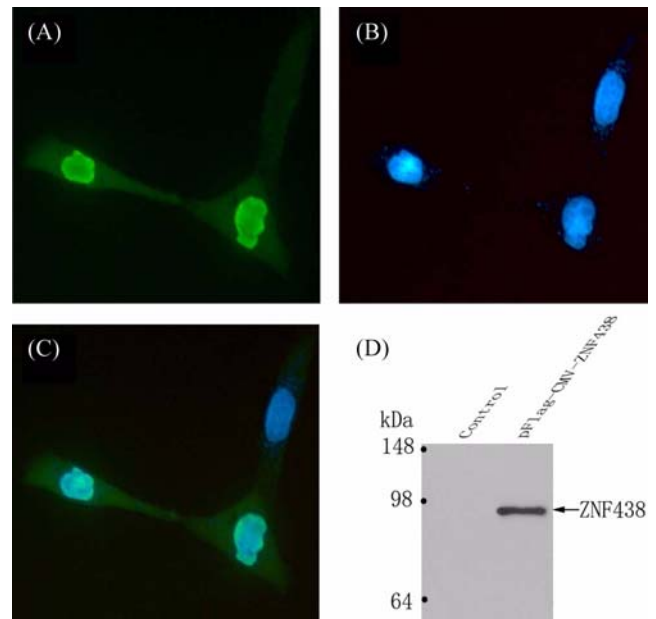
6). Thus, the nuclear accumulation experiments of ZNF438 proved our hypothesis.

In conclusion, we had cloned, characterized, and mapped a novel human zinc finger gene *ZNF438* which was widely expressed ubiquitously in 18 human tissues. ZNF438 was localized primarily in the nuclei. Overexpression of pM-ZNF438 in cultured cells could significantly inhibit the expression of the reporter gene in a dose-dependent manner, which suggested ZNF438 might serve a role as transcriptional repressor. In addition, we also performed Dual-luciferase reporter assay system to investigate the potential role of ZNF438 on signaling pathway, including AP1, CRE, E2F, GRE, Myc, NFκB, Rb and SRE pathway. However, it showed that ZNF438 failed to have any effect on the eight signaling pathways (data no shown). Further study would focus on whether ZNF438 would be modified by sumoylation and whether its activity was regulated by such modification.

References

Bardwell, V. J. and Treisman, R. (1994) The POZ domain: a conserved protein-protein interaction motif. *Genes Dev.* **8**, 1664-1677.

Bellefroid, E. J., Marine, J. C., Ried, T., Lecocq, P. J., Riviere, M., Amemiya, C., Poncelet, D. A., Coulie, P. G., de Jong, P., Szpirer, C., Ward, D. C. and Martial, J. A. (1993) Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. *EMBO J.* **12**, 1363-1374.

Brown, R. S., Sander, C. and Argos, P. (1985) The primary structure of transcription factor TFIIIA has 12 consecutive repeats. *FEBS Lett.* **186**, 271-276.

Busch, S. J. and Sassone-Corsi, P. (1990) Dimers, leucine zippers and DNA-binding domains. *Trends Genet.* **6**, 36-40.

De Luca, P., Majello, B. and Lania, L. (1996) Sp3 represses transcription when tethered to promoter DNA or targeted to promoter proximal RNA. *J. Biol. Chem.* **271**, 8533-8536.

Gostissa, M., Hengstermann, A., Fogal, V., Sandy, P., Schwarz, S. E., Scheffner, M. and Del Sal, G. (1999) Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1. *EMBO J.* **18**, 6462-6471.

Grishin, N. V. (2001) Treble clef finger-a functionally diverse zinc-binding structural motif. *Nucleic Acids Res.* **29**, 1703-1714.

Hagen, G., Muller, S., Beato, M. and Suske, G. (1994) Sp1-mediated transcriptional activation is repressed by Sp3. *EMBO J.* **13**, 3843-3851

Han, Z. G., Zhang, Q. H., Ye, M., Kan, L. X., Gu, B. W., He, K. L., Shi, S. L., Zhou, J., Fu, G., Mao, M., Chen, S. J., Yu, L. and Chen, Z. (1999) Molecular cloning of six novel Kruppel-like zinc finger genes from hematopoietic cells and identification of a novel transregulatory domain KRNB. *J. Biol. Chem.* **274**, 35741-35748.

Knochel, W., Poting, A., Koster, M., el Baradi, T., Nietfeld, W., Bouwmeester, T. and Pieler, T. (1989) Evolutionary conserved modules associated with zinc fingers in Xenopus laevis. *Proc. Natl. Acad. Sci. USA* **86**, 6097-6100.

Krishna, S. S., Majumdar, I. and Grishin, N. V. (2003) Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res.* **31**, 532-550.

Kumar, A. P. and Butler, A. P. (1997) Transcription factor Sp3 antagonizes activation of the ornithine decarboxylase promoter by Sp1. *Nucleic Acids Res.* **25**, 2012-2019.

Lania, L., Majello, B. and De Luca, P. (1997) Transcriptional regulation by the Sp family proteins. *Int. J. Biochem. Cell. Biol.* **29**, 1313-1323.

Liang, Y., Robinson, D.F., Dennig, J., Suske, G. and Fahl, W. E. (1996) Transcriptional regulation of the SIS/PDGF-B gene in human osteosarcoma cells by the Sp family of transcription factors. *J. Biol. Chem.* **271**, 11792-11797.

Lopes, J. E., Torgerson, T. R., Schubert, L. A., Anover, S. D., Ocheltree, E. L., Ochs, H. D. and Ziegler, S. F. (2006) Analysis of FOXP3 reveals multiple domains required for its function as a transcriptional repressor. *J. Immunol.* **177**, 3133-3142.

Mackay, J. P. and Crossley, M. (1998) Zinc fingers are sticking together. *Trends Biochem. Sci.* **23**, 1-4.

Melchior, F. (2000) SUMO--nonclassical ubiquitin. *Annu. Rev. Cell Dev. Biol.* **16**, 591-626.

Miller, J., Mclachlan, A. D. and Klug, A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *EMBO J.* **4**, 1609-1614.

Muller, S., Berger, M., Lehembre, F., Seeler, J. S., Haupt, Y. and Dejean, A. (2000) c-Jun and p53 activity is modulated by SUMO-1 modification. *J. Biol. Chem.* **275**, 13321-13329.

Noti, J. D. (1997) Sp3 mediates transcriptional activation of the leukocyte integrin genes CD11C and CD11B and cooperates with c-Jun to activate CD11C. *J. Biol. Chem.* **272**, 24038-24045.

Numoto, M., Niwa, O., Kaplan, J., Wong, K. K., Merrell, K., Kamiya, K., Yanagihara, K. and Calame, K. (1993) Transcriptional repressor ZF5 identifies a new conserved domain in zinc finger proteins. *Nucleic Acids Res.* **21**, 3767-3775.

Poukka, H., Aarnisalo, P., Karvonen, U., Palvimo, J. J. and Janne, O. A. (1999) Ubc9 interacts with the androgen receptor and activates receptor-dependent transcription. *J. Biol. Chem.* **274**, 19441-19446.

Poukka, H., Karvonen, U., Janne, O. A. and Palvimo, J. J. (2000) Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1). *Proc. Natl. Acad. Sci. USA* **97**, 14145-14150.

Rodriguez, M. S., Dargemont, C. and Hay, R. T. (2001) SUMO-1 conjugation *in vivo* requires both a consensus modification motif and nuclear targeting. *J. Biol. Chem.* **276**, 12654-12659.

Rodriguez, M. S., Desterro, J. M., Lain, S., Midgley, C. A., Lane, D. P. and Hay, R. T. (1999) SUMO-1 modification activates the transcriptional response of p53. *EMBO J.* **18**, 6455-6461.

Ross, S., Best, J. L., Zon, L. I. and Gill, G. (2002) SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization. *Mol. Cell* **10**, 831-842.

Schoenherr, C. J. and Anderson, D. J. (1995) Silencing is golden: negative regulation in the control of neuronal gene transcription. *Curr. Opin. Neurobiol.* **5**, 566-571.

Shastry, B. S. (1996) Transcription factor IIIA (TFIIIA) in the second decade. *J. Cell Sci.* **109**, 535-539.

Suske, G. (1999) The Sp-family of transcription factors. *Gene* **238**, 291-300.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A. and *et al.* (2001) The Sequence of the Human Genome. *Science* **291**, 1304-1351.

Williams, A. J., Blacklow, S. C. and Collins, T. (1999) The zinc finger-associated SCAN box is a conserved oligomerization domain. *Mol. Cell. Biol.* **19**, 8526-8535.

Wolfe, S. A., Nekludova, L. and Pabo, C. O. (2000) DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 183-212.

Wu, Y., Yu, L., Bi, G., Luo, K., Zhou, G. and Zhao, S. (2003) Identification and characterization of two novel human SCAN domain-containing zinc finger genes ZNF396 and ZNF397. *Gene* **310**, 193-201.

Zollman, S., Godt, D., Prive, G. G., Couderc, J. L. and Laski, F. A. (1994) The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila. *Proc. Natl. Acad. Sci. USA* **91**, 10717-10721.