GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Sequence variation and molecular evolution of *BMP4* genes

D.J. Zhang[1,5]*, J.H. Wu[2,3]*, G. Husile[4], H.L. Sun[2,3] and W.G. Zhang[4]

[1]College of Life Sciences, Inner Mongolia University, Hohhot, China
[2]Inner Mongolia Prataculture Research Center, Chinese Academy of Science, Hohhot, China
[3]Inner Mongolia Academy of Agricultural & Animal Husbandry Sciences, Hohhot, China
[4]College of Animal Science, Inner Mongolia Agricultural University, Hohhot, China
[5]College of Forestry, Agricultural University of Hebei, Hebei, China

*These authors contributed equally to this study.
Corresponding authors: W.G. Zhang
E-mail: zhangwg508@126.com

Genet. Mol. Res. 13 (4): 9196-9201 (2014)
Received June 11, 2013
Accepted October 25, 2013
Published November 7, 2014
DOI http://dx.doi.org/10.4238/2014.November.7.6

ABSTRACT. Bone morphogenetic protein 4 (*BMP4*) regulates skeletogenesis, osteoblastic differentiation, and the induction of hair follicles. Its protein-coding region contains a signal peptide, prodomain (which regulates post-translational synthesis), and a mature domain (which mediates gene function). Previous studies considered this gene to be conserved. By reanalyzing the coding region of *BMP4* in 16 mammalian species, we found that the mature domain is conserved in mammals. A comparison of the putative amino acid sequence demonstrates that BMP4 is relatively conserved. Two domains in BMP4 are connected by a random coil. The protein conformation differs between the Muridae family and other species, which might be associated with the body type of the former group.

Key words: *BMP4*; Mature domain region; Evolution

INTRODUCTION

Bone morphogenetic proteins (BMPs) are secreted signaling molecules that belong to the transforming growth factor-beta (TGF-b) superfamily, and exert their activity through interactions with specific BMP receptors (Derynck and Zhang, 2003; ten Dijke and Hill, 2004). In the extracellular space, BMP activity is modulated by BMP antagonists that regulate the magnitude, location, and timing of signaling through BMP receptors (Balemans and Van, 2002).

BMPs are multifunctional regulators of vertebrate development, controlling cell proliferation, differentiation, and apoptosis in various organs, including the skin (Botchkarev and Paus, 2003; Li et al., 2003; Mishina, 2003). In postnatal life, BMPs have significant function in normal tissue remodeling and homeostasis. *BMP4* is highly conserved among the pig, human, mouse, and rat.

In this study, partial exonic and deduced amino acid sequences of *BMP4* from the white Cashmere goat, Mongolian sheep, Merino sheep, and Huanghuai goat were analyzed and compared with those of other mammalian species in the GenBank (www.ncbi.nlm.nih.gov/genbank/). We also constructed a phylogenetic tree for all species using the partial *BMP4* sequences.

Our results provide a rationale for the further examination of the phylogenetic and evolutionary relationships between white Cashmere goat and other species, and increase our understanding about the molecular evolution of *BMP4* genes in mammals.

MATERIAL AND METHODS

Healthy adult white Cashmere goat, Mongolian sheep, and Merino sheep were selected as experimental animals. One milliliter EDTA (10 mg/mL) was added to a tube that contained 10 mL blood from each animal. The blood sample in the tube was mixed gently and stored at -20°C. Genomic DNA was isolated from blood by phenol-chloroform extraction and an ethanol precipitation step.

A primer pair was designed using the bovine *BMP4* sequence (GenBank accession number: NM_001045877) in Premier 5.0 and Oligo 6.0 and synthesized by Dalian TaKaRa Company Limited [forward primer (PF): 5'-AGCGCAGTCATCCCGGATTACAT-3'; reverse primer (PR): 5'-TTCAGAACCACCTTGTCATACTCAT CC-3']. The size of the expected amplicon, which should contain a partial sequence of an exon in *BMP4*, was 989 bp.

The PCR products were separated on 2% agarose gels, stained with ethidium bromide, and purified using a gel extraction kit following manufacturer protocols. The purity was checked on 2% agarose gels. Once the desired purity was achieved, the products were sequenced using the forward and reverse primers.

The partial exonic sequences of *BMP4* from white Cashmere goat, Mongolian sheep, and Merino sheep and the deduced amino acid sequences were compared with those of the pig, human, mouse, and rat from the GenBank using Clustal W 1.83 (Thompson et al., 1997), BioEdit 7.0.0 (http://www.mbio.ncsu.edu/bioedit/bioedit.html), DNAMAN 4.0 (http://dnaman.software.informer.com/), and DNAstar (http://www.dnastar.com/).

A phylogenetic tree was constructed, based on the partial BMP4 sequences in all species using MEGA 4 (Tamura et al., 2007). A likelihood method was used to examine nucleotide substitutions by selective pressure using Codeml in the Phylogenetic Analysis by Maximum Likelihood (PAML) software package (Yang, 2007). The models assumed no molecular

clock (clock = 0) and a single omega for all tree branches (model = 0), and we used likelihood ratio tests to compare the improvement in likelihood for models that incorporated 1, 2, or 3 site classes (NS sites = 0 1 2). Each analysis ran until convergence (Small_Diff = 0.5e-6); the control file is available on request (Yang et al., 2000).

To visualize the variation in x along each gene, we performed a sliding window analysis using SWAAP1.0.2 (Pride, 2000). We set the window size to 30 bp (10 codons) and the step size to 12 bp (4 codons). Values of ω were estimated based on Nei and Gojobori (Nei and Gojobori, 1986).

RESULTS AND DISCUSSION

The partial exon of *BMP4* in white Cashmere goat, Mongolian sheep, and Merino sheep was amplified, sequenced, and compared with the corresponding sequences in Huanghuai goat [EF632080], cow [NM_001045877], deer [S79174], pig [EU334835], dog [XM_547817], horse [XM_001494942], mouse [NM_007554], rat [NM_012827], meriones [AB201310], human [NM_130851], rhesus [XM_001084801], chimpanzee [XM_509954], and chicken [NM_205237] from the GenBank (Figure 1), of which the chicken sequence served as an outgroup. The homology and genetic distance between species were estimated.

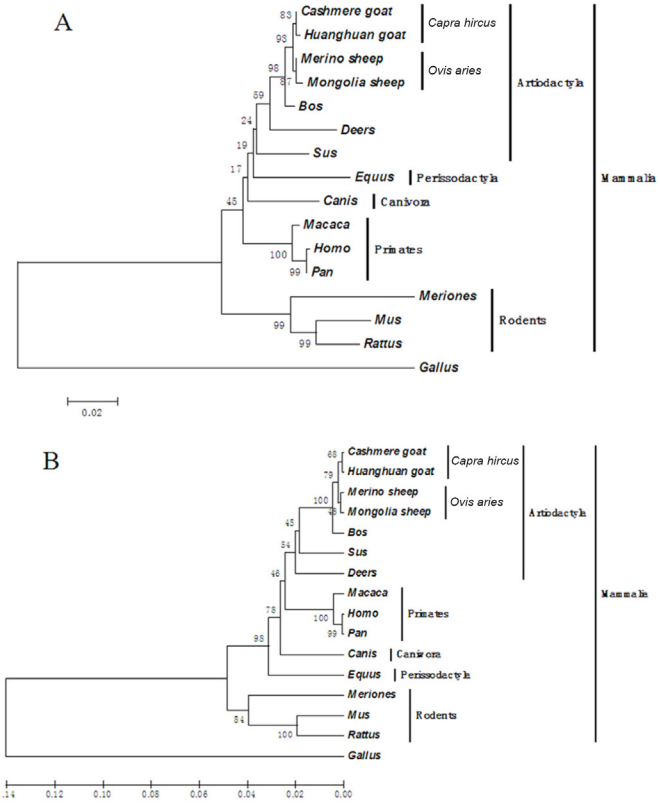

Figure 1. A. Neighbor-joining phylogenetic tree and **B.** UPGMA tree constructed using the partial exonic sequences of *BMP4* in 16 species.

The nucleotide sequences were relatively conserved between white Cashmere goat and other species, with homologies ranging from 78.6 to 99.9%. There were 6 insertions-deletions in mammals and chicken, 1 deletion in deer, and 1 deletion in the mouse and rat. Most single-nucleotide variations occurred at the same sites; more frequently at the 3rd nucleotide of the codon compared to the other 2 positions. This finding is consistent with the neutralism hypothesis of molecular evolution.

Neighbor-joining and maximum parsimony phylogenetic trees of these species were constructed, based on the sequences, and the confidence probabilities of every branch were analyzed using bootstrap values of 10,000 replications (Figure 1). The phylogenetic trees shared nearly the same topology; specifically, the white Cashmere goat was clustered with the Huanghuai goat first, followed by Merino sheep and Mongolian sheep, with 95% bootstrap values at the branches. Deer and Bovidae groups were clustered, with low bootstrap values of 63 and 51%, respectively. The mouse and rat formed a separate group, with 95 and 97% bootstrap values at the branches. Primates formed a separate group, with 100 and 98% bootstrapping values at the branches. Cow and Caprinae species were clustered, with 98 and 89% bootstrapping values.

The white Cashmere goat was the most closely related to the Huanghuai goat, followed by the Merino sheep and Mongolian sheep. At the molecular level, *BMP4* in the white Cashmere goat was the closest to the Huanghuai goat *BMP4*, and relatively distant to Mongolian sheep, despite the same region of distribution. This analysis also showed the presence of inter-species is the reproductive isolation during evolution.

BMP4 showed purifying selection between all pairs of species. By site-specific selection analysis using PAML, we noted strong support of a model with 2 site classes (M1 *vs* M0: $\chi^2 = 31.5$, df = 2, $P < 1e^{-7}$), of which 1 was under strong purifying selection (mean dN/dS = 0.055). Although certain positions showed signs of positive selection, there was no significant support for a positive selection model (M2a *vs* M1: $\chi^2 = 0$, df = 2, P = 1).

Sliding window analysis was used to determine where rates of nonsynonymous substitutions (dN) exceeded those of synonymous substitutions (dS) in specific functional regions, and hence identify positive selection. This analysis was also used to compare the variability in dN, dS, and ω ratios between clades and functional regions. None of the ω ratios exceeded 1 in *BMP4* for any species (Figure 2).

Putative amino acid sequence comparison confirmed that BMP4 is relatively conserved. The variation at the amino acid level was notably smaller compared to the DNA level, because most nucleotide substitutions are synonymous.

We submitted these amino acid sequences to 3D-PSSM (http://www.sbg.bio.ic.ac.uk/servers/3dpssm) (Kelley et al., 2000). The amino acid sequence was scanned against an up-to-date sequence library, to detect homologs and generate a 3-D model for the protein (Figure 3). 3D-PSSM identifies the best structural match to a queried region. Often, 3D-PSSM finds a confident hit, but only to 1 portion of a query. The region that we queried spanned amino acid 120 to 245 (about 126 residues), which had little difference in the queried sequence lengths among these species.

Protein conformation differed between Muridae and other species (Figure 3). We noted 2 domains that were connected by a random coil. *BMP4* is a major regulator of bone development (Thomas et al., 1996); thus, we speculate that these transitions in Muridae induce a smaller body type compared to other species.

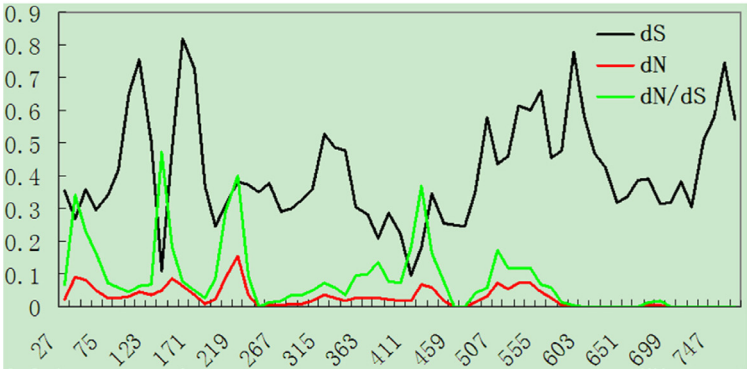

Figure 2. dN/dS rate ratio (ω, the green line) and dN (the red line) and dS (the black line) variation in the coding region of *BMP4* for all species (16 species listed in Figure 1).

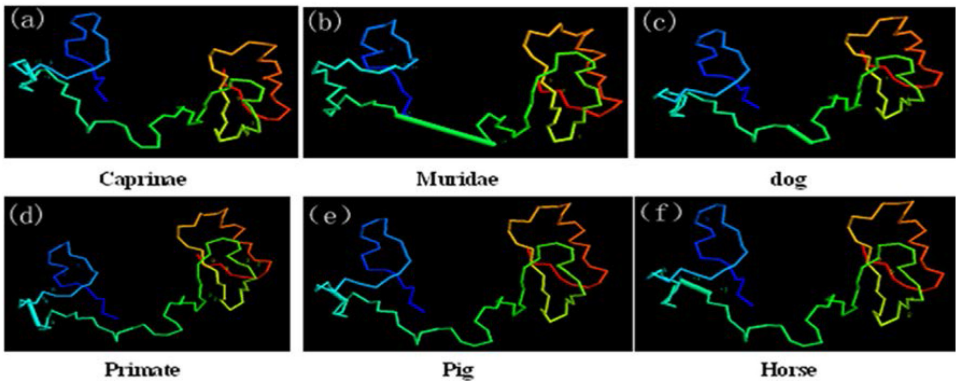

Figure 3. 3-D model of BMP4 protein in all species.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of China (#31260538, #31260545), the Key Project of National Science and Technology Pillar Program of China (#2011BAD28B05), and the Natural Science Foundation of Inner Mongolian (#2013MS0414, #2011MS0401). The authors thank the Inner Mongolia Cashmere goat stud farm staff for provisions and help with sample collection.

REFERENCES

Balemans W and Van Hul W (2002). Extracellular regulation of BMP signaling in vertebrates: a cocktail of modulators. *Dev. Biol.* 250: 231-250.

Botchkarev VA and Paus R (2003). Molecular biology of hair morphogenesis: development and cycling. *J Exp Zool* 298: 164-180.

Derynck R and Zhang YE (2003). Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature* 425: 577-584.

Kelley LA, MacCallum RM and Sternberg MJE (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* 299(2), 499-520.

Li AG, Koster MI and Wang XJ (2003). Roles of TGF-beta signaling in epidermal/appendage development. *Cytokine Growth Factor Rev.* 14:99-111.

Mishina Y (2003). Function of bone morphogenetic protein signaling during mouse development. *Front. Biosci.* 8:855-869.

Nei M and Gojobori T (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3(5):418-426.

Pride DT (2000). SWAAP Version 1.0.0-sliding windows alignment analysis program: a tool for analyzing patterns of substitutions and similarity in multiple alignments. Available at [http://www.bacteriamuseum.org/SWAAP/SwaapPage.htm]. Accessed June 1, 2010.

Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.

ten Dijke P and Hill C (2004). New insights into TGF-beta-Smad signalling. *Trends Biochem. Sci.* 29:265-273.

Thomas JT, Lin K, Nandedkar M, Camargo M, et al. (1996). A human chondrodysplasia due to a mutation in a TGF-beta superfamily member. *Nat. Genet*. 12(3): 315-317.

Thompson JD, Gibson TJ, Plewniak F, et al. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882.

Yang Z (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24: 1586-1591.

Yang Z, Nielsen R, Goldman N and Pedersen AM (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155: 431-449.

