Trends in Bioinformatics

ISSN 1994-7941

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS Trends in Bioinformatics

ISSN 1994-7941
DOI: 10.3923/tb.2016.70.80

Short Communication
Conservation Pattern, Homology Modeling and Molecular Phylogenetic Study of BMP Ligands

[1]Sudarshan Chatterjee, [1]Madhuri Haque, [2]Md. Shaifur Rahman, [2]Hossen M. Jamil, [2]Naznin Akhtar, [3]S.M. Abdul-Awal, [1]Md. Shahedur Rahman and [2]Sikder M. Asaduzzaman

[1]Department of Biotechnology and Genetic Engineering, Jahangirnagar University, 1342 Dhaka, Bangladesh
[2]Tissue Banking and Biomaterial Research Unit, Atomic Energy Research Establishment, 1349 Dhaka, Bangladesh
[3]Biotechnology and Genetic Engineering Discipline, Khulna University, 9208 Khulna, Bangladesh

Abstract

Background: Bone Morphogenic Proteins (BMPs) play many important roles in embryogenesis and metamorphosis throughout the life of vertebrates and other animals. The BMPs are multifunctional and some BMPs do similar functions which lead us to believe that it may have existed before the evolution and they may evolve from a single BMP. Additionally, less is known regarding phylogeny and conservation based conserved molecular mechanism of BMPs. Hence, evolutionary relationship of 16 BMP ligands (phylogenetic as well as protein sequence conversed patterns) were done. **Materials and Methods:** For this study, protein sequences were retrieved from UniProtKB, homology modeling was executed by Swiss-model using 3rjr.1A and 2qcq.1A as template, followed by MSA and phylogeny. Later conserved regions of BMP ligand were compared. Protein subfamily determination was done in Zebra and supported by the phylogenetic data. **Results:** Remarkably, similar region of conserved area were observed and different disulfide linkage pattern had been identified. It was found notable patterns in C1-C4, C2-C5 and C3-C6 that all but BMP3 and BMP15 do not contain 7th Cys. Phylogeny study indicate, according to evolutionary clock, GDF11 and BMP15 were more distantly diverged taxa, BMP3A and BMP3B had same point origin and GDF5, 6 and 7 are homologous. These seven proteins, as per this study indicates that, those are evolved at simultaneously during evolution, whereas other nine forms monophyletic taxa. **Conclusion:** However, more studies needs to explore on this. As per homology modelling studied, BMP ligands shares common evolutionary origin, some of the members are highly diverged indicating ancient evolutionary history of these protein. This study will be useful for wet lab once the evolutionary relationship gets established.

Key words: Bone morphogenic protein, bioinformatical analysis, conservation pattern, homology modeling, molecular phylogenetic

Received: April 25, 2016 Accepted: May 17, 2016 Published: September 15, 2016

Citation: Sudarshan Chatterjee, Madhuri Haque, Md. Shaifur Rahman, Hossen M. Jamil, Naznin Akhtar, S.M. Abdul-Awal, Md. Shahedur Rahman and Sikder M. Asaduzzaman, 2016. Conservation pattern, homology modeling and molecular phylogenetic study of BMP ligands. Trends Bioinform., 9: 70-80.

Corresponding Author: Md. Shaifur Rahman, Tissue Banking and Biomaterial Research Unit, Atomic Energy Research Establishment, 1349 Dhaka, Bangladesh

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

The Bone Morphogenic Protein (BMP), Growth and Differentiation Factor (GDF) belongs to large TGF-β superfamily[1]. Based on sequence similarities new TGF-β family members is added to increase the number and revealed the newer functions[2]. It is beyond the scope of the article to present the biology of TG factors in any detail. However, BMP family comprises the largest members and most of them are multifunctional and few individual BMP shares many functions[3,4]. For instances, BMP2, 4, 6, 7 and 9 are capable of inducing bone formation in ectopic sites while BMP4 is also a critical determinant of the non-osteogenic embryological patterning of mammals[5,6]. Moreover, some of BMP family members shares common pathways and followed conserved molecular mechanism[7,8]. These multiple functions and sharing feature of BMPs family members are suggesting by the fact that all BMP-like molecules may have an evolutionary relation and same origin point. For example, chromosomal assignments of TGF family members indicate that these genes have become widely dispersed during their evolution[9]. Many investigators have isolated, sequenced and showed crystallographic structure of BMPs[10,11]. It is widely accepted that the degree and extent of sequence homology can suggest a phylogenetic relationships within the any protein superfamily, for instance BMPs family[12-14]. For distantly related genes the definition of homologous (orthologous) or duplicated (paralogous) genes is often difficult, a phylogenetic study can often solve this problem[15,16]. Evolutionary information from sequence data may therefore allow a prediction of the possible biochemical properties of these otherwise uncharacterized protein.

Evolutionary histories of proteins study have been intensified by the growing genomic data and large uncharacterized protein sequence records[17]. More so, proteins that are a part of a large family are being conveniently identified, characterized and classified with the help of their comparative or homologous features and structural figure since all protein families exhibit some level of similarity and relationship. Functional characterization of an unknown protein can be analysis by its predicted 3D structure through computational methods and later proved by wet lab experimental search[18]. But, the prediction of protein model have some short comings such as the determination techniques are time consuming and complicated nature[19]. Fortunately, different computational methods have been employed to bridge the gap between the number of known sequences and 3D models[20]. Template-based protein structure modeling techniques is one of very interesting methods that rely on principles of natural proteins from the theory of evolution viewpoint[21,22]. In this study, 16 major BMP ligands (Table S1) were analyzed to understand their protein sequence conversed patterns and evaluate phylogenetic relationship among these protein family members. Overall phylogeny unfolds the evolutionary relationship; however more studies needs to explore on this. Model evaluation were also performed to check the reliability of the models. This study will be useful for wet lab once the evolutionary relationship gets established.

Table S1: Human BMP ligands with their alternative name, function their Swiss-prot accession number, gene name, amino acid residue and corresponding RCSB PDB ID were analyzed during this study

Protein name	Alternative name	Functions	Accession No.	Gene name	Amino acid residues	PDB Id
BMP4	BMP2B	Induces cartilage and bone formation	P12644	BMP4	408	N/A*
BMP2	BMP2A	Induces cartilage and bone formation	P12643	BMP2	396	3bmp
BMP10	N/A	Cardiac development	O95393	BMP10	424	N/A*
BMP15	GDF9b	Follicle-stimulating hormone activity	O95972	BMP15	392	N/A*
GDF2	BMP9	Potent circulating inhibitor of angiogenesis	Q9UK05	GDF2	429	4mpl
BMP5		Chondrogenesis	P22003	BMP5	454	N/A*
BMP6	Vrg1	Osteoblast differentiation, chondrogenesis	P22004	BMP6	513	2r52
BMP7	OP1	Osteoinductive, development of kidney and eye	P18075	BMP7	431	1bmp
BMP8A	OP2	Bone morphogenesis, maintenance of spermatogenesis	Q7Z5Y6	BMP8A	402	N/A*
BMP8B	OP3	Bone induction, calcium regulation and bone homeostasis	P34820	BMP8B	402	N/A*
BMP3A	Osteogenin	Negative regulator of bone morphogenesis	P12645	BMP3A	472	2qcq
BMP3B	GDF10	Endochordial bone formation and bone morphogenesis inhibitor	P55107	GDF10	478	N/A*
GDF5	BMP12	Enhances tendon healing and bone formation	P43026	GDF5	501	1waq
GDF6	BMP13	Chondrogenesis and hypertrophy	Q6KF10	GDF6	455	N/A*
GDF7	BMP14	Tendon and ligament formation, repair and development of sensory neurons	Q7Z4P5	GDF7	450	N/A*
GDF11	BMP11	Development of spinal cord during embryogenesis, eye, pancreas development, kidney formation and skeleton patterning	O95390	GDF11	407	N/A*

*N/A: No experimental structure available

MATERIALS AND METHODS

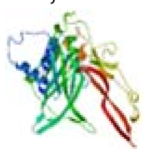
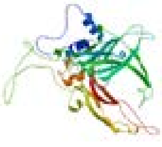
Sequence and structure retrieval: Sequence data of BMP2, 3, 3B, 4, 5, 6, 7, 8, 8B, 10, 15 and GDF2, 5, 6, 7, 11 were retrieved from UniProtKB (http://www.uniprot.org/) as FASTA format[23]. Crystallographic structures for BMP2, 3, 6, 7 and GDF2, 5 were collected from RCSB-PDB (http://www.rcsb.org/pdb)[24]. Rest of the protein tertiary structures those are not available in RCSB-PDB, protein structural modeling by Swiss-model (http://swissmodel.expasy.org/) were executed to derive predictive structure[25,26].

Sequence alignment and analysis of conserved position: The retrieved sequences were analyzed using ClustalW-MSA for obtaining pairwise distance[27]. Calculation was done in MEGA6.06[28]. Web based aligner T-Coffee (http://www.tcoffee.org/Projects/tcoffee/) was used to generate a reliable MSA[29] and this data were saved in Phylip format those aim to be incorporated in Phylogenetic Web Repeater (POWER: http://power.nhri.org.tw/power/home.htm)[30,31]. Multiple Sequence Comparison by Log-Expectation (MUSCLE): http://www.drive5.com/) was used to analyze the conserved and consensus positions of the C-terminal and results were observed in JAVA runtime environment viewer JALVIEW[32,33].

Protein subfamily predictions and phylogenetic tree construction: In Zebra web-server (http://biokinet.belozersky.msu.ru/zebra), MSA files were given in FASTA format with ".fa" extension was used to predict the subfamilies within a BMPs family[34]. This web interface identified amino acid residues responsible for functional discrimination and selected hotspots for directed evolution[34]. The POWER web interface (http://power.nhri.org.tw/) was applied for phylogenetic tree construction and analysis which based on the Dayhoff-PAM and PROTDIST method[31]. The KITSCH (http://caps.ncbs.res.in/iws/ phylip_files/kitsch.html) program (Fitch-Margoliash least squares method) was used for distance analysis of phylogenetic tree[35]. The reliability of the estimated trees was evaluated by Bootstrap method with 1000 and 500 replications.

Homology modeling of BMP ligands and conservation pattern determination: To infer structural characteristics of BMP ligands, homology modeling was performed and predicted 3D structures using the Swiss-model (http://swiss model.expasy.org)[26]. Ten proteins (Table 1) were subjected to model. Figure 1-6 were generated with RASWIN program[36,37]. Sequence identity above 30% is a relatively good predictor of the expected accuracy of a model[38]. Model quality is assessed

Table 1: Modeled BMP ligands with corresponding templates (with chain identifier), figures and model evaluation data

Protein name	BMP4	GDF10/BMP3B	BMP10	BMP8A	BMP8B	GDF6	BMP5	GDF11	GDF7	BMP15
Template	3rjr.1.A	2qcq.1.A	3rjr.1.A	3rjr.1.A	3rjr.1.A	3rjr.1.A	3rjr.1.A	3rjr.1.A	3rjr.1.A	3rjr.1.A
Figure of models										
Model evaluation (Swiss-model)										
GMQE	0.51	0.16	0.45	0.50	0.50	0.42	0.45	0.51	0.38	0.45
QMEAN4	-6.86	-0.00	-11.28	-8.67	-9.22	-9.06	-9.71	-8.11	-12.26	-11.83
Model template rang	49-408	372-478	50-424	30-402	30-402	58-455	42-454	59-407	85-450	46-392
Template evaluation (PSI blast, NCBI)										
Query coverage (%)	51	23	87	59	59	26	45	81	22	54
E-value	1e-11	2e-64	2e-12	6e-11	9e-11	9e-10	7e-14	3e-19	1e-09	2e-04
Sequence identity (%)	29	82	25	26	26	33	27	27	37	24

with the local composite scoring function QMEAN. The weights of QMEAN have been specifically retrained for Swiss-model, leading to more accurate local quality predictions. In addition, global QMEAN scores are calculated as indicators for the overall model quality. Further, a combined quality estimate is provided, which combines the QMEAN estimate with the GMQE obtained from the target-template alignment. The resulting GMQE is expressed as a number between zero and one, where higher numbers indicate higher reliability. Later, PDB files of BMPs structures from both RCSB-PDB and homology modeled with corresponding chain identifier were uploaded in the Consurf server to perform MSA using Multiple Alignment using Fast Fourier Transform (MAFFT). Consurf server was run in "Conseq mode" which retrieved us conserved amino acid residues with their specific position[39]. The homologues were collected from UNIREF90 with search algorithm PSI-BLAST. The E-value of PSI-BLAST E-value was settled at 0.0001 and the numbers of iterations were 3. The chosen sequences by the server were manually set between the ranges of maximal identity of 95% to minimal identity of 35%. Neighbor joining tree were constructed with maximum likelihood distance. Conservation scores were calculated by Bayesian method and model of substitution for proteins[40,41]. For functional protein interaction network, STRING database was used to identify the known and predicted protein interactions partners of BMPs both direct (physical) and indirect (functional) associations[42,43].

RESULTS AND DISCUSSION

Pairwise distances from Multiple Sequence Alignment (MSA) and conservation pattern of BMP: For understanding a protein family it is good to know the protein sequences, structures and conservation patterns. The BMPs were classified into subfamily according to structural similarity, previously reported. However, amino acid sequence and position specific information were used for classification. Sequence similarity was found to vary among the studied 16 BMP proteins. The MEGA pairwise distances score between the BMP sequences was estimated from 1.744 (maximum) to 0.011 (minimum) with an average of 1.208. Pairwise distances of all members are given in four different charts in Fig. 1.

The results indicated that BMPs sequences shared different level of distances between themselves and this may be a hint of their evolutionary disperses. Six cysteine residues were identified as highly conserved at 501, 620, 624, 659, 692 and 694 positions of multiply aligned consensus sequence. The C-terminal region of BMP ligands were found very conserved (Fig. 2).

Phylogenetic tree reconstruction and analysis: The evolutionary history inferred from phylogenetic analysis is usually depicted as branching, tree like diagrams that represent an estimated pedigree of the inherited relationships among molecules[44]. The phylogenetic tree analysis made possible to explore the evolutionary history and divergence of BMP protein. In this study, two different studies are used to generate unrooted trees[45], one relying on protein distance matrix (Protdist) and another is maximum parsimonious tree relying on Fitch-Margoliash least squares method. In the unrooted tree resulted from protdist method, GDF11 was most diverged and BMP15 was second most diverged taxon (Fig. 3a). The BMP3A and 3B formed a monophyletic group and GDF5, 6 and 7 formed another monophyletic group. Ligands in subfamily one formed a single monophyletic taxon which later divided and formed into four different monophyletic taxon. In the maximum parsimonious tree topology were established with 500 bootstrap replicates (Fig. 3b). All leaf nodes and internal nodes were same for both trees except for BMP15 and GDF11.

Protein subfamily determination: The BMP shares common evolutionary origin but some of the members are highly diverged which indicated an ancient evolutionary history of these proteins. Their clustering pattern in tree demanded categorizing into three subfamily which was done in Zebra and supported by the phylogenetic data. Previously, BMP family members were classified into four categories: BMP-2/4, BMP-5/6/7/8a/8b, BMP-9/10 and BMP-12/13/14[46]. Selected 16 BMP proteins from this study were found to be clustered into three subfamilies. Nine ligands BMP2, 4, 5, 6, 7, 8a, 8b, 10 and GDF2 were categorized into subfamily-1. The rest of the members were classified into subfamily-2 comprising BMP15, GDF5, 6, 7 and subfamily-3 comprising BMP3A, 3B, GDF11.

Protein homology model data: Protein homology modeling executed in this study was that effort to generate the structural information that has not been experimentally obtained yet. The structures of homologous proteins are generally better conserved than their sequences. This phenomenon is demonstrated by the prevalence of Structurally Conserved Regions (SCRs) even in highly divergent protein families[47] Template based protein modeling were performed for 10 BMP ligands that have no experimental 3D structural data available (Table 1). All the templates that have been used are with more than 20% sequence identity and some are above 30%. All the modeled protein in this study contains a good GMQE (around 50%) which supports the reliability of these modeled proteins.

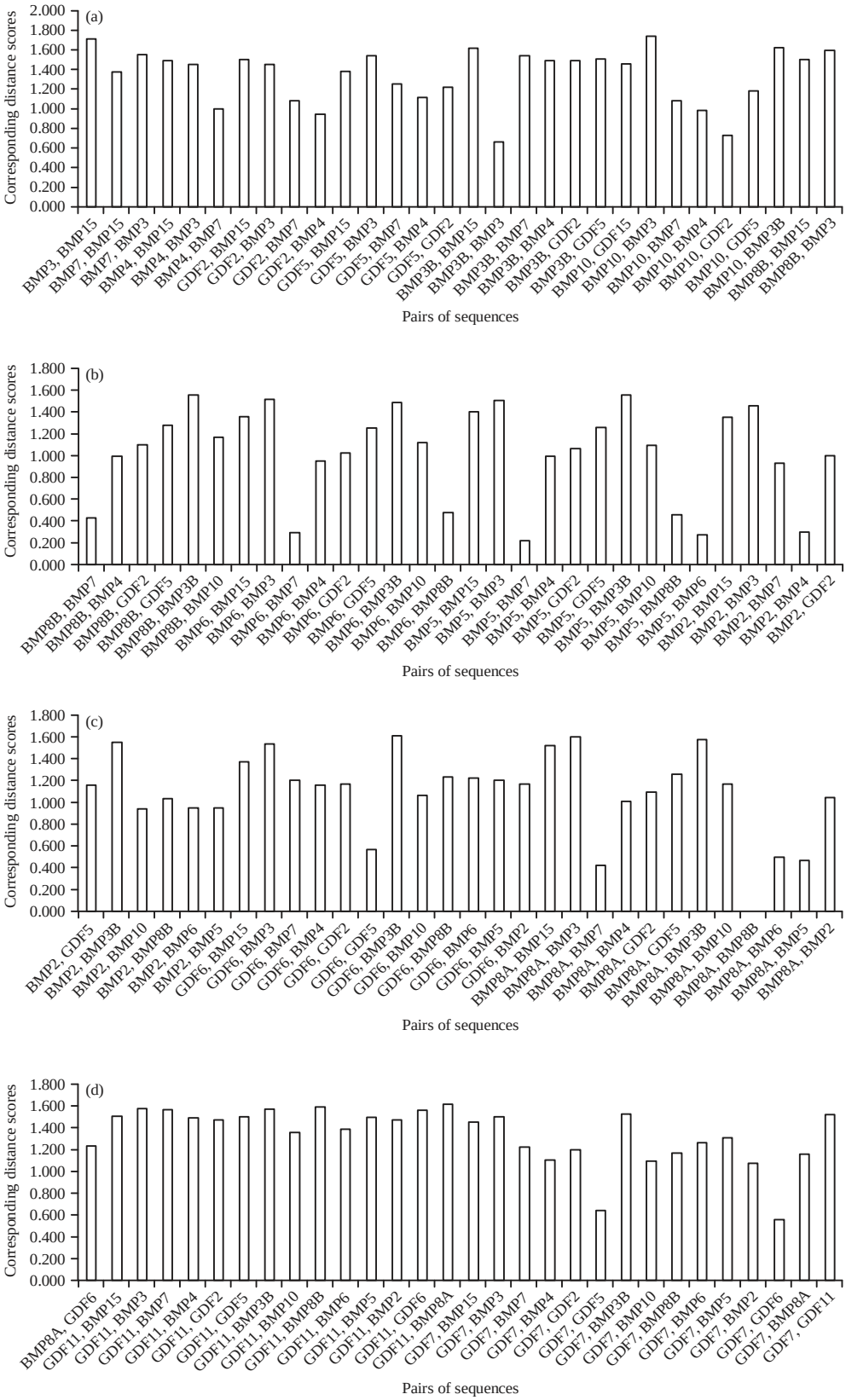


Fig. 1(a-d): MSA scores of protein sequences of 16 human BMP ligands

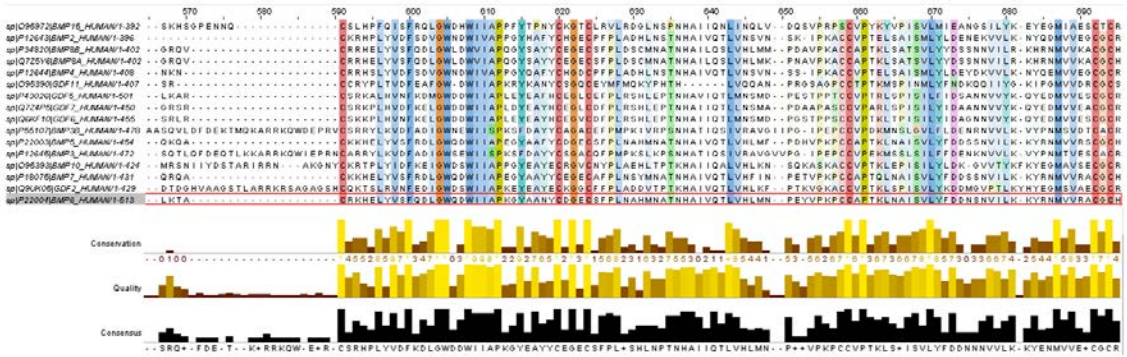


Fig. 2: Conservation pattern of BMP in C-terminal region from their MSA file, upper panel is aligned sequences and lower panel is conservation score, quality and consensus sequence

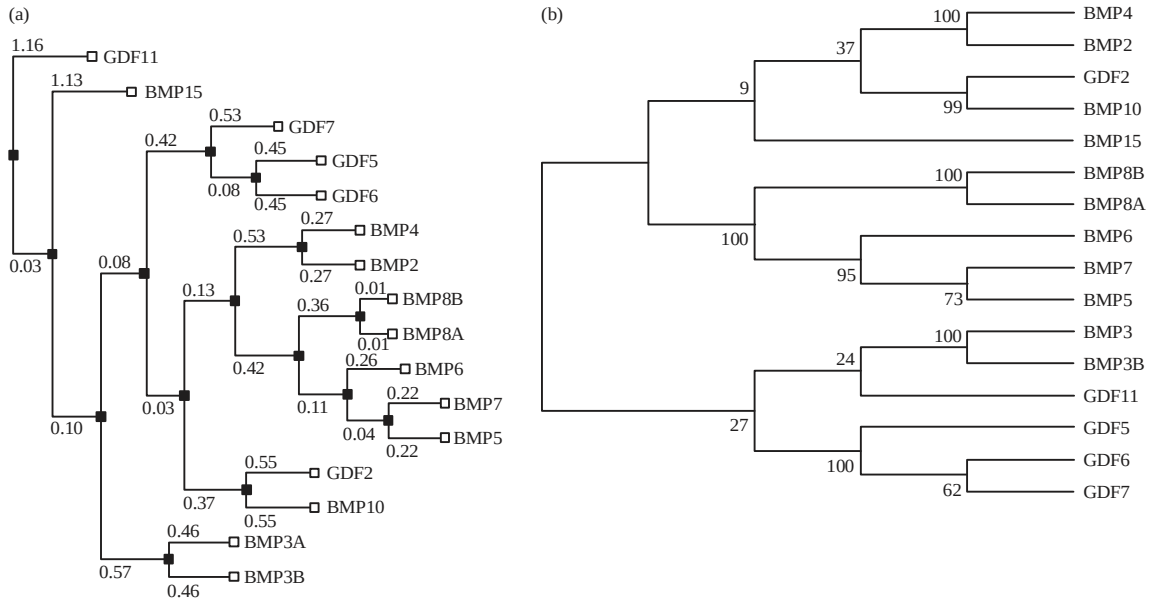


Fig. 3(a-b): (a) Unrooted phylogenetic tree of human BMP family based on Protdist method, phylogenetic tree was reconstructed with evolutionary clock and corresponding branch length indicating state of divergence and (b) Cladogram of human BMP family based on Fitch-Margoliash least squares method, maximum parsimonious tree with 500 times bootstrap replication, values above a branch is the parsimony bootstrap value for the branching defining each monophylectic clade of BMPs, the bootstrap value at the top of each branch suggests confidence of a branch to be formed, both cladogram and phylogram indicated similar phylogenetic history

Conserved pattern of BMP ligands: However, BMPs are highly conservative in structures[1]. But, residue swapping selectively alters BMP orienting ability and structure lead to distinct function[48]. The structural data comprising the conserved amino acid residues in the BMP is represented in Fig. 4. Cysteine plays very important role in stabilization of protein structure at higher level because disulfide bridges formed by cysteine residues are permanent component of protein primary structure, it can also change secondary structure by steric constraints. Cysteine related disulfide bridges are permanent element for stabilization of the tertiary structure and in most cases; interchain S-S bridges are absolute condition for quaternary structure to exist. In this study, the entire cysteine residue with their position and found a remarkable pattern of C1-C4, C2-C5 and C3-C6. All but BMP3 and BMP15 do not contain the 7th cysteine residue[49].

N-terminal region is rich in variable amino acid residues which indicate these regions are flexible to mutations.

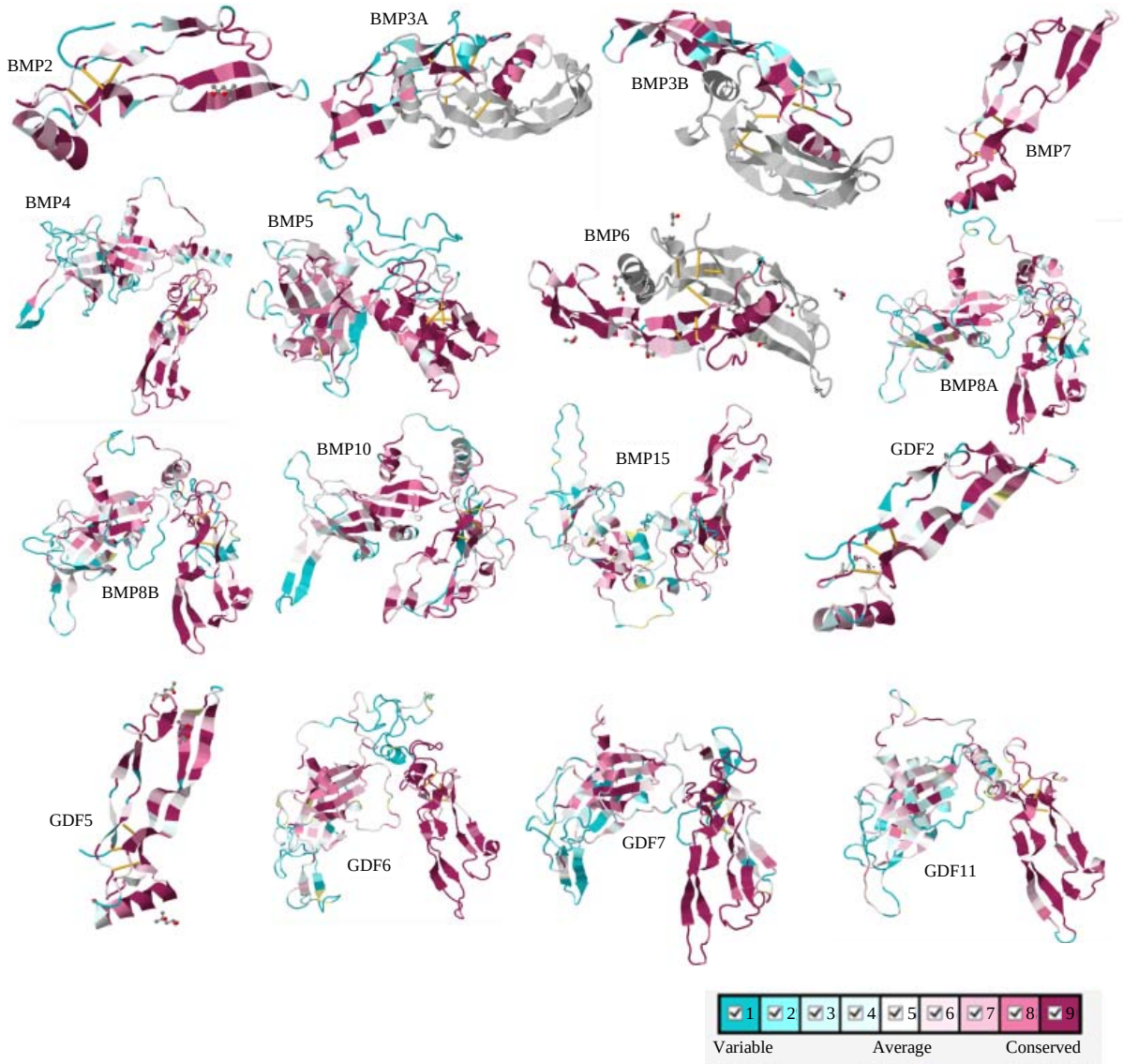


Fig. 4: Conservation pattern of BMP tertiary structures shown in cartoon view and conservation scores are categorized into 9 levels

First 200 amino acids are exposed to outer environment and are rarely functionally or structurally involved. On the contrary, C-terminal region is highly rich in conserved amino acid residues and these conserved amino acids are of functional and structural significance. Cysteine residues were specially found to be conserved and their disulfide linking pattern is C1-C4, C2-C5 and C3-C6. The conserved cysteine residues are given in Table 2 and found to be following three distinct patterns which are shown in Fig. 5.

One pattern consist three disulphide bond and variable number of free cysteine residues as in BMP2 and most other BMPs. Another pattern consist four disulphide bonds and five free cysteine residues as in GDF11. The last pattern is for BMP proteins forming homodimers like BMP3A, BMP3B and BMP6. They contain an A-chain and B-chain which are identical molecules forming dimer and they are bridged by the same amino acids from each chains.

Protein-protein interaction among BMP ligands: From the interaction string image, BMP receptors are seems to be the central functional partner to be interacting with BMP ligands. Effectors like SMADs, receptors like ACVR and other proteins like noggin and chordin are seen to be interacting with BMP (Fig. 6). From this network that scores of all the nodes ranges from 0.998-0.999, therefore it is deducible that all the proteins are strongly interconnected finally, protein interaction image has given their signaling and interaction information briefly.

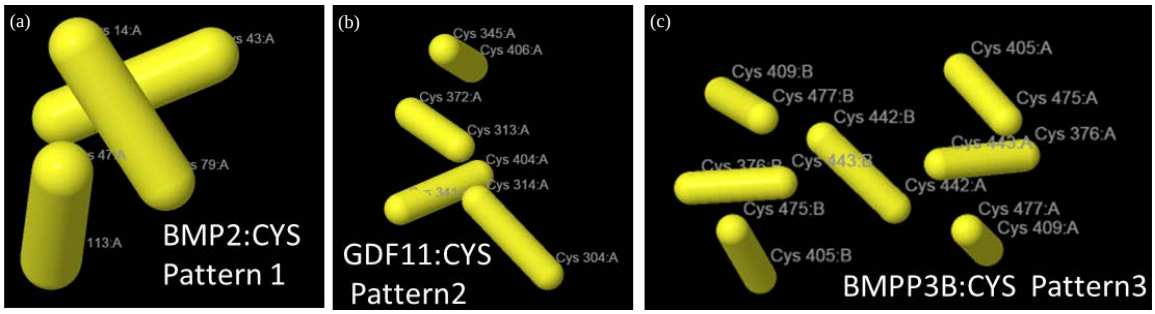
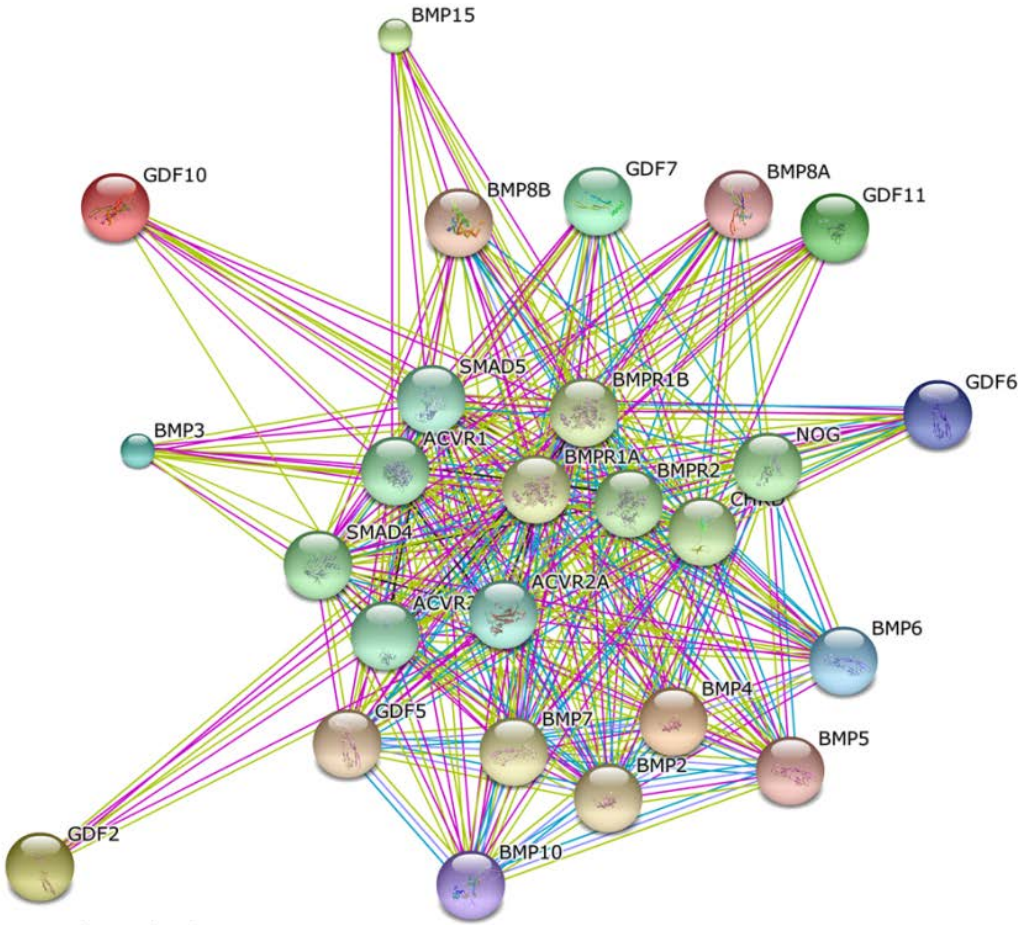


Fig. 5(a-c): Disulphide linking pattern, (a) Pattern 1: There are three disulphide linkage as they are assumed to be monomer, (b) Pattern 2: Consists four linkage, it is a monomer too and (c) Pattern 3: Consists of seven disulfide linkage three from each monomer, a common disulphide bridge between two monomer



Predicted functional partners		Score
BMPR1A:	Bone morphogenetic protein receptor, type IA on ligand binding, forms a receptor complex consi […] (532 aa)	0.999
BMPR1B:	Bone morphogenetic protein receptor, type IB on ligand binding, forms a receptor complex consi […] (502 aa)	0.999
CHRD:	Chordin (955 aa)	0.999
BMPR2:	Bone morphogenetic protein receptor, type II (serine/threonine kinase) on ligand binding, forms […] (1038 aa)	0.999
NOG:	Noggin; Essential for cartilage morphogenesis and joint formation. Inhibitor of bone morphogene […] (232 aa)	0.999
SMAD4:	SMAD family member 4 common SMAD (co-SMAD) is the coactivator and mediator of signal transduct […] (552 aa)	0.999
ACVR1:	Activin A receptor, type I on ligand binding, forms a receptor complex consisting of two type […] (509 aa)	0.999
ACVR2B:	Activin A receptor, type IIB transmembrane serine/threonine kinase activin type-2 receptor for […] (512 aa)	0.998
SMADS:	SMAD family member 5 transcriptional modulator activated by BMP (bone morphogenetic proteins […] (464 aa)	0.998
ACVR2A:	Activin A receptor, type IIA on ligand binding, forms a receptor complex consisting of two type […] (513 aa)	0.998

Fig. 6: Protein-protein network between the proteins of BMP ligands and STRING image shows the proteins and their functional partners in colored circles and the interaction with colored straight lines

Table 2: Conserved cysteine residues with their position and disulphide bonding information

Protein name	BMP2	BMP4	BMP5	BMP7	BMP8A	BMP8B	BMP10	GDF2	GDF5	GDF7	GDF6	GDF11	BMP15	BMP3A	BMP3B	BMP6
Free cysteine residue (s)	C78	C372	C272, C418	C103	C31, C205, C341, C366	C31, C205, C341, C366	C388	C73	C84	C414, C185, C231	C191, C230, C419	C62, C65, C161, C162, C371	C150, C209	-	-	-
Cysteine residue in disulfide bonds	C14-C79, C43-C111, C47-C113	C308-C373, C337-C405, C341-C407	C353-C419, C382-C451, C386-C453	C38-C104, C67-C136, C71-C118	C301-C367, C330-C399, C334-C401	C301-C367, C330-C399, C334-C401	C323-C389, C352-C421, C356-C423	C8-C74, C37-C107, C41-C109	C19-C85, C48-C117, C52-C119	C349-C415, C378-C447, C382-C449	C335-C420, C383-C452, C387-C454	C304-C314, C313-C372, C341-C404, C345-C406	C291-C357, C320-C389, C324-C391	C8-C75*, C37-C107*, C41-C109*, C74(A)-C74(B)	C376-C443*, C405-C475*, C409-C477*, C442(A)-C442(B)	C31-C97*, C60-C129*, C64-C131*, C96(A)-C96(B)
Total cysteine residues	7	7	8	7	10	10	7	7	7	9	9	13	8	14 (Homodimer)	14 (Homodimer)	14 (Homodimer)

CONCLUSION

The BMPs are a complex family of extracellular proteins. The results reported in this study, along with the accumulating data and models from other experimental study are brought together to present the evolutionary history of BMPs. Phylogenetic analyses of this study suggest that human BMPs are a differentiated group of proteins rendering similar function and belongs to a monophylatic taxon although few of them are highly diverged. The C-terminal fold of BMPs and conserved nature of amino acid residues seems to have the potential for many different uses, depending on the protein expression and interactions. Hypotheses of BMPs function can now be placed in a phylogenetic context and it can be anticipated that the phylogenetic framework presented here will provide important insights into the history of changing functions of these apparently simple but versatile proteins.

SIGNIFICANCE STATEMENTS

Bone Morphogenetic Proteins (BMPs) family are being increasingly studied due to their contribution in embryonic development and physiological functions in human body throughout life. Practically, recombinant BMP2 and 7 are able to induce bone formation on scaffold and ectopic site. It has already been used in different clinical applications such as open fracture of long bones, non-unions and vertebral arthrodesis and lumbar fusion. This study reported about the template-based protein 3D structure modeling; molecular phylogenetics, conserved pattern analysis of 16 BMP ligands and interaction partner identification which could provide a broader understanding of this protein family. From functional standpoint, this study will be useful for wet lab to study molecular mechanisms of action and significance in dental, cartilage and bone-related diseases once the structure and evolutionary relationship gets established.

ACKNOWLEDGMENT

The authors would like to thank TBBRU, Atomic Energy Research Establishment, 1349 Dhaka for the support of software during the bioinformatical study.

REFERENCES

1. Yang, J., P. Shi, M. Tu, Y. Wang, M. Liu, F. Fan and M. Du, 2014. Bone morphogenetic proteins: Relationship between molecular structure and their osteogenic activity. Food Sci. Human Well., 3: 127-135.

2. Nana, A.W., P.M. Yang and H.Y. Lin, 2015. Overview of transforming growth factor β superfamily involvement in glioblastoma initiation and progression. Asian Pac. J. Cancer Prev., 16: 6813-6823.
3. Rahman, M.S., N. Akhtar, H.M. Jamil, R.S. Banik and S.M. Asaduzzaman, 2015. TGF-β/BMP signaling and other molecular events: Regulation of osteoblastogenesis and bone formation. Bone Res., Vol. 3. 10.1038/boneres.2015.5
4. Poniatowski, L.A., P. Wojdasiewicz, R. Gasik and D. Szukiewicz, 2015. Transforming growth factor beta family: Insight into the role of growth factors in regulation of fracture healing biology and potential clinical applications. Mediators Inflammation. 10.1155/2015/137823
5. Kingsley, D.M., 1994. The TGF-beta superfamily: New members, new receptors and new genetic tests of function in different organisms. Genes. Dev., 8: 133-146.
6. Granjeiro, J.M., R.C. Oliveira, J.C. Bustos-Valenzuela, M.C. Sogayar and R. Taga, 2005. Bone morphogenetic proteins: From structure to clinical use. Braz. J. Med. Biol. Res., 38: 1463-1473.
7. Schmierer, B. and C.S. Hill, 2007. TGFβ-SMAD signal transduction: Molecular specificity and functional flexibility. Nat. Rev. Mol. Cell Biol., 8: 970-982.
8. Massague, J., J. Seoane and D. Wotton, 2005. Smad transcription factors. Genes, 19: 2783-2810.
9. Burt, D.W. and A.S. Law, 1994. Evolution of the transforming growth factor-beta superfamily. Prog. Growth Factor Res., 5: 109-107-118.
10. Brown, M.A., Q. Zhao, K.A. Baker, C. Naik and C. Chen *et al*., 2005. Crystal structure of BMP-9 and functional interactions with pro-region and receptors. J. Biol. Chem., 280: 25111-25118.
11. Scheufler, C., W. Sebald and M. Hulsmeyer, 1999. Crystal structure of human bone morphogenetic protein-2 at 2.7 A resolution. J. Mol. Biol., 287: 103-115.
12. Moulton, J.K. and B.M. Wiegmann, 2007. The phylogenetic relationships of flies in the superfamily Empidoidea (Insecta: Diptera). Mol. Phylogenet. Evol., 43: 701-713.
13. Wharton, K.A., G.H. Thomsen and W.M. Gelbart, 1991. Drosophila 60A gene, another transforming growth factor beta family member, is closely related to human bone morphogenetic proteins. PNAS-Develop. Biol., 88: 9214-9218.
14. Hadjebi, O., E. Casas-Terradellas, F.R. Garcia-Gonzalo and J.L. Rosa, 2008. The RCC1 superfamily: From genes, to function, to disease. Biochimica Biophysica Acta (BBA)-Mol. Cell Res., 1783: 1467-1479.
15. Koonin, E.V. and M.Y. Galperin, 2003. Sequence-Evolution-Function: Computational Approaches in Comparative Genomics. Kluwer Academic, Boston, ISBN-13: 9781475737837, Pages: 462.
16. Tekaia, F., 2016. Inferring orthologs: Open questions and perspectives. Genomics Insights, 9: 17-28.
17. Zhou, J., D.K. Thompson and J.M. Tiedje, 2004. Genomics: Toward A Genome-Level Understanding of the Structure, Functions and Evolution of Biological Systems. In: Microbial Functional Genomics, Zhou, J., D.K. Thompson, Y. Xu and J.M. Tiedje (Eds.). Chapter 1, John Wiley & Sons Inc., Canada, ISBN-13: 9780471071907, pp: 1-20.
18. Fernandez-Fuentes, N., J. Zhai and A. Fiser, 2006. ArchPRED: A template based loop structure prediction server. Nucl. Acids Res., 34: W173-W176.
19. Fiser, A., 2010. Template-based protein structure modeling. Methods Mol. Biol., 673: 73-94.
20. Zahiri, J., J.H. Bozorgmehr and A. Masoudi-Nejad, 2013. Computational prediction of protein-protein interaction networks: Algorithms and resources. Curr. Genomics, 14: 397-414.
21. Bordoli, L. and T. Schwede, 2012. Automated protein structure modeling with SWISS-MODEL workspace and the protein model portal. Meth. Mol. Biol., 857: 107-136.
22. Lee, H.S. and W. Im, 2012. Identification of ligand templates using local structure alignment for structure-based drug design. J. Chem. Inform. Model, 52: 2784-2795.
23. UniProt Consortium, 2014. UniProt: A hub for protein information. Nucl. Acids Res., 43: D204-D212.
24. Berman, M.H., J. Westbrook, Z. Feng, G. Gilliland and T.N. Bhat *et al*., 2000. The protein data bank. Nucleic Acids Res., 28: 235-242.
25. Biasini, M., S. Bienert, A. Waterhouse, K. Arnold and G. Studer *et al*., 2014. SWISS-Model: Modelling protein tertiary and quaternary structure using evolutionary information. Nucl. Acids Res., 42: W1-W252.
26. Arnold, K., L. Bordoli, J. Kopp and T. Schwede, 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22: 195-201.
27. Fox, G., F. Sievers and D.G. Higgins, 2016. Using *de novo* protein structure predictions to measure the quality of very large multiple sequence alignments. Bioinformatics, 32: 814-820.
28. Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar, 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol., 30: 2725-2729.
29. Di Tommaso, P., S. Moretti, I. Xenarios, M. Orobitg and A. Montanyola *et al*., 2011. T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucl. Acids Res., 39: W13-W17.
30. Notredame, C., D.G. Higgins and J. Heringa, 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol., 302: 205-217.

31. Lin, C.Y., F.K. Lin, C.H. Lin, L.W. Lai, H.J. Hsu, S.H. Chen and C.A. Hsiung, 2005. POWER: PhylOgenetic WEb Repeater-an integrated and user-optimized framework for biomolecular phylogenetic analysis. Nucl. Acids Res., 33: W553-W556.
32. Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res., 32: 1792-1797.
33. Waterhouse, A.M., J.B. Procter, D.M.A. Martin, M. Clamp and G.J. Barton, 2009. Jalview version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics, 25: 1189-1191.
34. Suplatov, D., D. Shalaeva, E. Kirilin, V. Arzhanik and V. Svedas, 2014. Bioinformatic analysis of protein families for identification of variable amino acid residues responsible for functional diversity. J. Biomol. Struct. Dyn., 32: 75-87.
35. Felsenstein, J., 1989. PHYLIP-Phylogeny Inference Package. Cladistics, 5: 164-166.
36. Marcey, D., 1995. A manual for molecular viewing using RASWIN (RASMOL). http://biology.kenyon.edu/BMB/RASWIMAN.HTM.
37. Bernstein, H.J., 2000. Recent changes to RasMol, recombining the variants. Trends Biochem. Sci., 25: 453-455.
38. Forrest, L.R., C.L. Tang and B. Honig, 2006. On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. Biophy. J., 91: 508-517.
39. Berezin, C., F. Glaser, J. Rosenberg, I. Paz and T. Pupko, 2004. ConSeq: The Identification of functionally and structurally important residues in protein sequences. Bioinformatics, 20: 1322-1324.
40. Ashkenazy, H., E. Erez, E. Martz, T. Pupko and N. Ben-Tal, 2010. ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucl. Acids Res., 38: W529-W533.
41. Landau, M., I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko and N. Ben-Tal, 2005. ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res., 33: W299-W302.
42. Von Mering, C., L.J. Jensen, B. Snel, S.D. Hooper and M. Krupp *et al*., 2005. STRING: Known and predicted protein-protein associations, integrated and transferred across organisms. Nucl. Acids Res., 33: D433-D437.
43. Jensen, L.J., M. Kuhn, M. Stark, S. Chaffron and C. Creevey *et al*., 2009. STRING 8-a global view on proteins and their functional interactions in 630 organisms. Nucl. Acids Res., 37: D412-D416.
44. Brinkman, F.S.L. and D.D. Leipe, 2001. Phylogenetic Analysis. In: Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Baxevanis, A.D. and B.F.F. Ouellette (Eds.). 2nd Edn., John Wiley and Sons, New York, USA., pp: 323-358.
45. Ganfornina, M.D., G. Gutierrez, M. Bastiani and D. Sanchez, 2000. A phylogenetic analysis of the lipocalin protein family. Mol. Biol. Evol., 17: 114-126.
46. Miyazono, K., S. Maeda and T. Imamura, 2005. BMP receptor signaling: transcriptional targets, regulation of signals and signaling cross-talk. Cytokine Growth Factor Rev., 16: 251-263.
47. Huang, I.K., J. Pei and N.V. Grishin, 2013. Defining and predicting structurally conserved regions in protein superfamilies. Bioinformatics, 29: 175-181.
48. Perron, J.C. and J. Dodd, 2012. Structural distinctions in BMPs underlie divergent signaling in spinal neurons. Neural Dev., Vol. 7. 10.1186/1749-8104-7-16
49. Little, S.C. and M.C. Mullins, 2009. Bone morphogenetic protein heterodimers assemble heteromeric type I receptor complexes to pattern the dorsoventral axis. Nature Cell Biol., 11: 637-643.