BMC Evolutionary Biology

RESEARCH ARTICLE

Open Access

Molecular evolutionary and structural analysis of familial exudative vitreoretinopathy associated FZD4 gene

Suman Seemab[1†], Nashaiman Pervaiz[1†], Rabail Zehra[1], Saneela Anwar[1], Yiming Bao[2*] and Amir Ali Abbasi[1*]

Abstract

Background: Frizzled family members belong to G-protein coupled receptors and encode proteins accountable for cell signal transduction, cell proliferation and cell death. Members of Frizzled receptor family are considered to have critical roles in causing various forms of cancer, cardiac hypertrophy, familial exudative vitreoretinopathy (FEVR) and schizophrenia.

Results: This study investigates the evolutionary and structural aspects of Frizzled receptors, with particular focus on FEVR associated *FZD4* gene. The phylogenetic tree topology suggests the diversification of Frizzled receptors at the root of metazoans history. Moreover, comparative structural data reveals that FEVR associated missense mutations in *FZD4* effect the common protein region (amino acids 495–537) through a well-known phenomenon called epistasis. This critical protein region is present at the carboxyl-terminal domain and encompasses the K-T/S-XXX-W, a PDZ binding motif and S/T-X-V PDZ recognition motif.

Conclusion: Taken together these results demonstrate that during the course of evolution, *FZD4* has acquired new functions or epistasis via complex patter of gene duplications, sequence divergence and conformational remodeling. In particular, amino acids 495–537 at the C-terminus region of FZD4 protein might be crucial in its normal function and/or pathophysiology. This critical region of FZD4 protein may offer opportunities for the development of novel therapeutics approaches for human retinal vascular disease.

Keywords: Multigene family, G-protein coupled receptors, Frizzled, FZD4, Metazoans, Phylogenetic analysis, Structural analysis, Familial exudative vitreoretinopathy

Background

G-protein coupled receptors (GPCRs) are diverse group of membrane proteins, encoded by more than 800 genes in the human genome [1]. GPCRs are found within the plasma membrane and share a common architecture entailing seven-transmembrane (TM) domains [2]. The function of GPCRs is to detect a wide range of extracellular signals, mainly small organic molecules and whole proteins [3]. After detection, GPCRs bind to the ligand and undergo several conformational changes which result in the activation of multiplex signaling networks and a cellular response [4]. GPCRs play a vital role in various physiological processes, ranging from sight, smell and taste to enormous number of neurological, cardio-vascular, and reproductive functions [5]. The involvement of GPCR superfamily in such complex processes makes it a major target for various drug discovery approaches [6]. GPCRs are categorized into five main families namely, glutamate, rhodopsin, secretin, adhesion and Frizzled receptor family [4].

Frizzleds (FZDs) are 7- transmembrane (TM) receptors reside on plasma membrane. They play an important regulatory role in controlling cell polarity and cell proliferation during embryonic development by transmitting signals from glycoproteins, mainly Wnt proteins [7]. Because many Frizzleds have been defined in metazoans, the tissue specific expression of Frizzled genes is

* Correspondence: baoym@big.ac.cn; abbasiam@qau.edu.pk
†Suman Seemab and Nashaiman Pervaiz contributed equally to this work.
2BIG Data Center & CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China
1National Center for Bioinformatics, Program of Comparative and Evolutionary Genomics, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad 45320, Pakistan

a very intricate task [8]. Predominantly, Frizzled receptors are extensively and broadly expressed and almost every single cell expresses one or more Frizzled receptors [9]. The Frizzled family consists of 10 Frizzled receptors, *FZD1, FZD2, FZD3, FZD4, FZD5, FZD6, FZD7, FZD8, FZD9* and *FZD10*. Frizzled receptors vary in length ranging from 537 to 706 amino acids. Exploring differences and similarities among members of Frizzled Receptor family can be helpful in unraveling functional evolution of this family.

Among the members of Frizzled receptor family, *FZD4* is the most finely understood member in context of its biological function, its interaction with ligands and other proteins, and also with respect to its role in human disease phenotype. Several studies suggest that mutations in *FZD4* gene may affect its normal cellular function [10–12]. In humans, a large number of missense mutations in *FZD4* results in a variable amount of retinal hypovascularization, a condition called familial exudative vitreoretinopathy (FEVR) [13]. Mice null for frizzeled4 (*Fzd4*$^{-/-}$) exhibit premature intra-retinal vasculature and thus provide further evidence for a role of FZD4 in retinal angiogenesis [14].

Keeping in view the prominence of Frizzled Receptor family in several developmental processes, the detailed phylogenetic analysis is performed to infer the evolutionary origin and diversification of its members. To investigate the variations in domain topologies, a comparative analysis of known functional domains is conducted. Moreover, due to the substantial role of *FZD4* in FEVR, 40 missense mutations are targeted for comparative structural analysis to scrutinize their impact on overall protein conformation.

Results
Phylogenetic analysis
The NJ and MJ based phylogenetic tree topologies of the FZD family reveal two major clusters. Cluster-I is further divided into three sub-clusters; sub-cluster of *FZD10, FZD9* and *FZD4*, sub-cluster of *FZD5-FZD8*, and sub-cluster of *FZD7, FZD2* and *FZD1*. Cluster-II consists of *FZD3-FZD6* (Fig. 1; Additional file 1). The tree topology pattern further suggests that Frizzled receptor family is diversified by in total 9 duplication events (Fig. 1; Additional file 1). First duplication has occurred after the divergence of eumetazoa from metazoans separating the ancestral genes of Cluster-I from ancestor of Cluster-II (Fig. 1; Additional file 1). Second duplication event has occurred at least prior to the bilaterian-nonbilaterian split separating the ancestral gene of sub-cluster *FZD7, FZD2* and *FZD1* from the ancestral genes of sub-clusters *FZD10, FZD9, FZD4, FZD5* and *FZD8* (Fig. 1; Additional file 1). The third duplication event results in the split-up of the ancestral genes of *FZD5* and

FZD8 from the ancestors of *FZD10, FZD9* and *FZD4* (Fig. 1; Additional file 1). This duplication event occurred at least prior to the divergence of protostomes and deuterostomes (Fig. 1; Additional file 1). Fourth duplication event has transpired at the root of Nephrozoa approximately 650 million years ago and is responsible for the separation of *FZD4* and ancestral gene of *FZD9/FZD10* (Fig. 1; Additional file 1). The remaining five duplication events are specific to the vertebrate lineage (Fig. 1; Additional file 1). Furthermore, tree topology revealed that fifteen specie-specific duplications have occurred at different evolutionary time periods both in invertebrate and vertebrate FZD homologs (Fig. 1; Additional file 1).

Domain organization of frizzled receptor family
In order to gain insight into comparative domain organization of frizzled receptor family, previously reported domains and motif of this family are searched through extensive literature survey and mapped on putative human paralogs (Fig. 2).

Comparative domain investigation exhibits an amino-terminal membrane localizing signal peptide sequence rich in hydrophobic residues (Fig. 2). This signal peptide is followed by a 120 residues long conserved Cystein Rich Domain (CRD) (Fig. 2). Amino-terminal Cystein Rich Domain (CRD) comprehends YNXT motif which is highly conserved among all the putative paralogs of FZD family (Fig. 2). This motif is located exactly 5 residues after the second cysteine residue of CRD domain and considered to be a potential N-glycosylation site. This site might play a significant role in Wnt ligand binding [15]. CRD domain is necessary for WNT-FZD interaction and activation of Wnt pathways (Fig. 2) [16]. In addition, FZD1 contains an extra Glycosaminoglycan (GAG) binding site RxR residing on position 194–196 of CRD domain (Fig. 2). Previous investigation suggests that CRD domain of FZD1 requires glycosaminoglycans (GAGs) for regulation of signal transduction and cell proliferation [16].

The transmembrane (TMs) domain is conserved among all the putative paralogs of the family (Fig. 2). The occurrence of this domain in frizzled receptors indicates that this family belongs to the superfamily of G- Protein Coupled Receptors (GPCRs). It is the membrane spanning region of frizzled receptors which is predicted to contain seven transmembrane alpha helices [17]. The carboxyl-terminal region contains conserved K-T/S-XXX-W PDZ-binding motif that is located immediately after the hydrophobic transmembrane domain (Fig. 2) [18–20]. This motif is conserved among all the members of FZD family and is very essential for initiation of canonical Wnt pathway and for phosphorylation of Disheveled (Dvl) protein [21]. Interaction
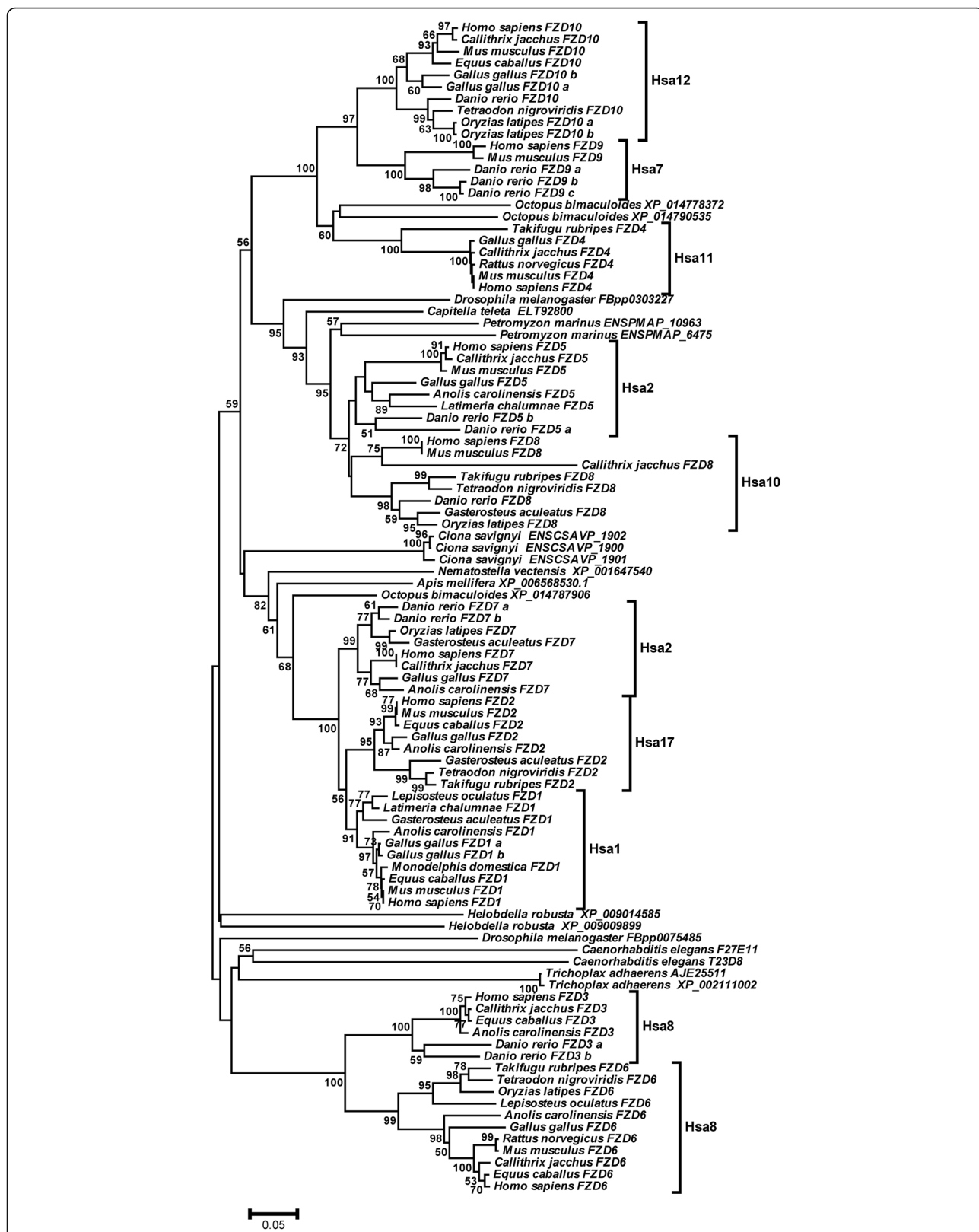
Fig. 1 Phylogenetic tree of Frizzled receptor gene family. Evolutionary history of Frizzled receptor gene family was inferred by neighbor joining method. Complete-deletion option was used to eradicate gaps and missing data. Numbers on branches represent bootstrap values (based on 1000 replications), only the values ≥50% are presented here. Scale bar depicts amino acid substitutions per site
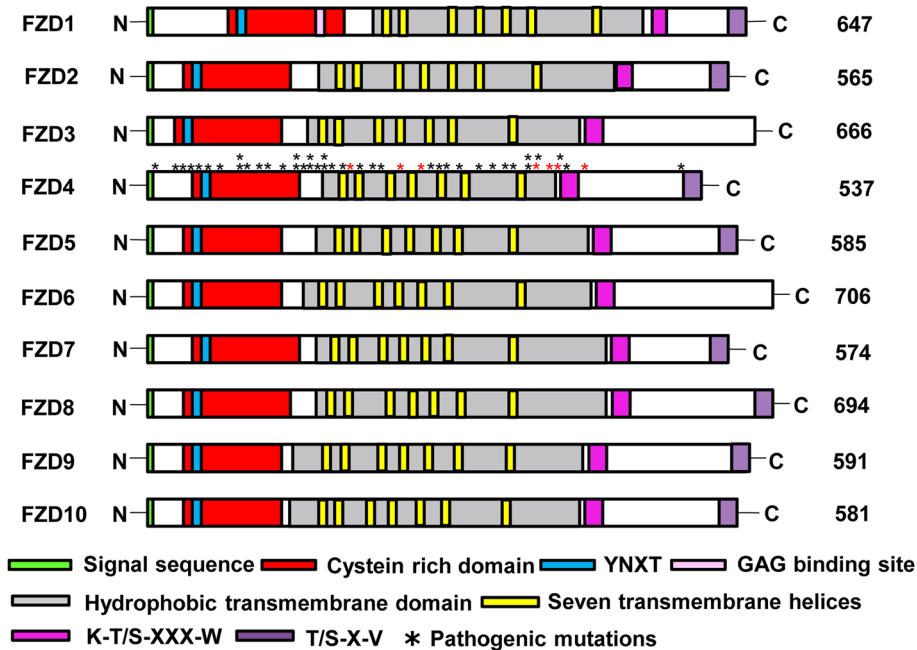
Fig. 2 Domain organization of human Frizzled receptors. Schematic illustration depicts comparative organization of key functional domains and motifs of Frizzled receptor across human paralogs. Protein lengths are drawn approximately to scale. Domains and motifs are color coded. Asterisk symbol indicates the positions of 47 different FEVR causing variants on human FZD4 protein (40 missense variants are indicated with the black asterisk symbol, whereas 7 nonsense variants are depicted by red asterisks)

between FZD-Dvl completely diminishes when this motif encounters any mutation [21].

The comparative domain investigation revealed another carboxy-terminal residing T/S-X-V PDZ binding motif, present in all the members of FZD family except FZD3 and FZD6 (Fig. 2) [21]. This motif resides exact 14 residues downstream of internal Dvl-PDZ binding motif in FZD1, FZD2 and FZD7, while this gap is greater than 29 residues in other FZD family members (Fig. 2).

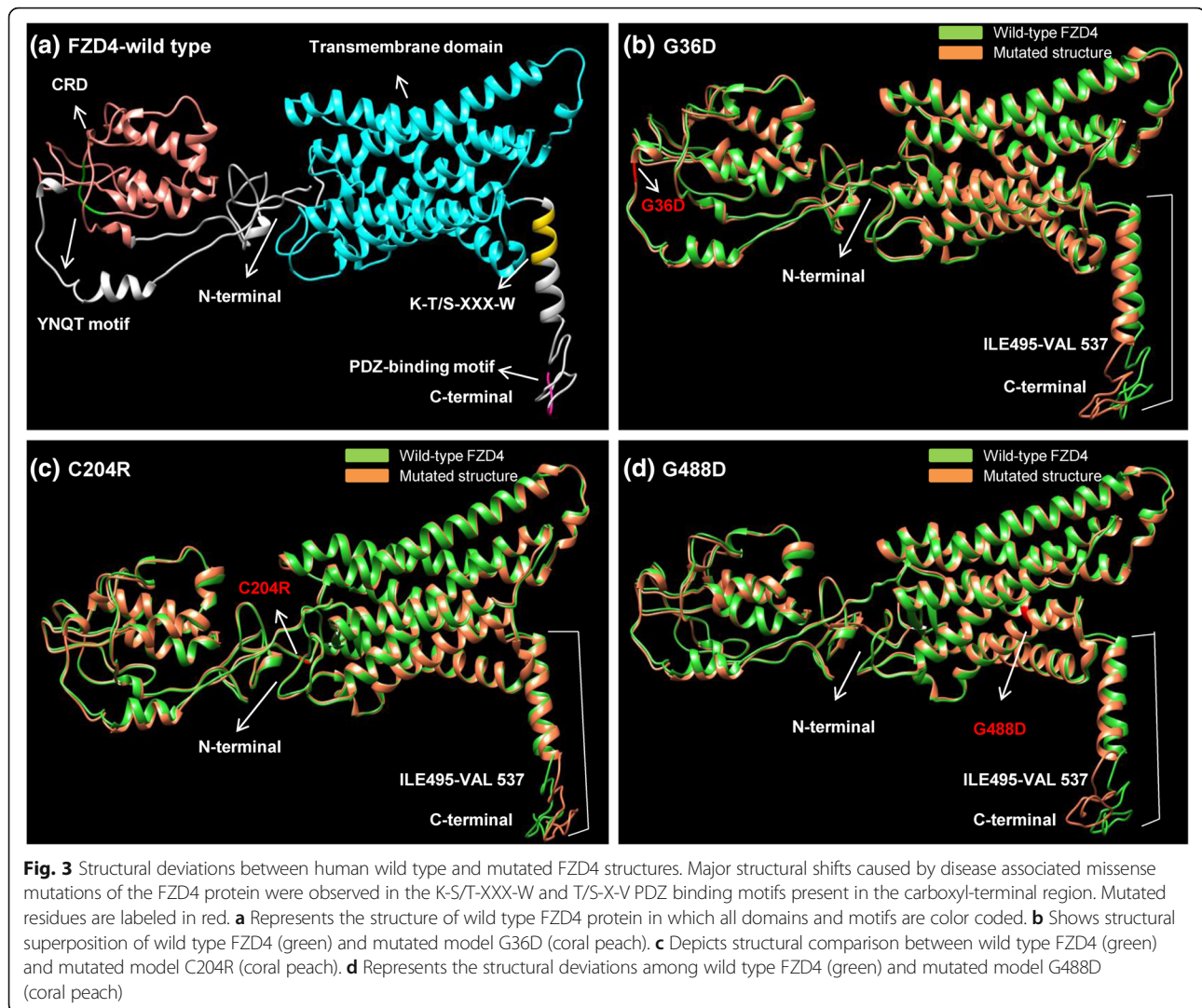### Structural analysis of wild-type and mutant FZD4
In order to study the impact of mutations causing FEVR on overall conformation of FZD4 protein, a comparative structural analysis was performed (Fig. 3; Additional file 2). The full length X-ray crystallography structure of FZD4 is not available, therefore, three dimensional structure of wild-type human FZD4 protein is predicted through the I-TASSER structure prediction server [22]. The predicted wild-type protein structure is taken as a reference and compared with 40 predicted mutated models (previously reported FEVR causing FZD4 missense mutants). The predicted structures of wild-type and mutated proteins show high scores when their quality is evaluated through RamPAGE and ERRAT (Additional file 3) [23]. Structural deviations are evaluated with RMSD values [24]. All of the superimposed models (wild-type and 40 mutated structures) expose a common deviated region (495–537) (Fig. 3; Table 1; Additional file 2). This region possesses

the K-T/S-XXX-W Dvl-PDZ binding motif and the T/S-X-V PDZ recognition motif at the carboxyl-terminal [9] (Table 1). Therefore, this particular region (amino acids; 495–537) of FZD4 protein is considered as critical with respect to its structure, function and involvement in FEVR. In addition to C-terminal deviated residues (495–537) shared among all mutants, structural deviations are also detected in highly conserved N-terminal extracellular cysteine-rich domain (CRD) and seven transmembrane domains of FZD4 (Fig. 3; Table 1; Additional file 2).

Discussion
Increasing availability of the genome sequence data and high throughput annotation of genes permitted the molecular evolutionary analysis of various human genes and gene families [25]. Frizzled receptor family encode proteins that constitute the key component of Wnt signaling pathway with numerous developmental roles including cell proliferation, cell differentiation, tissue hemostasis and cell apoptosis [4]. The present study is an attempt to relate the evolutionary history and comparative structural aspects of human Frizzled receptors with particular phenotypic trait and disease.

NJ and ML based phylogenies of the FZD family (Fig. 1; Additional file 1) supported by high bootstrap scores revealed the complex evolutionary relationship among members of Frizzled receptor family. The phylogenetic tree topology portrayed a very ancient evolutionary

Fig. 3 Structural deviations between human wild type and mutated FZD4 structures. Major structural shifts caused by disease associated missense mutations of the FZD4 protein were observed in the K-S/T-XXX-W and T/S-X-V PDZ binding motifs present in the carboxyl-terminal region. Mutated residues are labeled in red. **a** Represents the structure of wild type FZD4 protein in which all domains and motifs are color coded. **b** Shows structural superposition of wild type FZD4 (green) and mutated model G36D (coral peach). **c** Depicts structural comparison between wild type FZD4 (green) and mutated model C204R (coral peach). **d** Represents the structural deviations among wild type FZD4 (green) and mutated model G488D (coral peach)

history of Frizzled receptor family. The overall branching pattern revealed that FZD3 and FZD6 are the most ancient members of this family and have evolved prior to the divergence of placozoa (trichoplax) from eumetazoans, forming the most basal branch. The remaining members of this family are diversified in ParaHoxozoa history prior to tetrapod-teleost split. The close phylogenetic relationship among gene family members might depict their functional resemblance and similar cellular localization [26, 27]. For instance, both FZD3 and FZD6 are expressed in central nervous system (CNS) and are known to participate in the neural tube closure-related planer cell polarity (PCP) pathway and axonal growth during development [28]. In addition, FZD1, FZD2, and FZD7 falling in the same clade also share their functions in closure of the palate and ventricular septum [26].

Inspecting the protein domain topologies of the putative paralogs revealed conserved domain features; a cysteine rich domain and a seven helix transmembrane domain. However, some cluster specific differences in motifs were observed. For example, in the carboxyl-terminal PDZ binding site, a three residue motif crucial for the binding of PDZ domain of the DVL proteins is present in all putative members of the family except FZD3 and FZD6 [29]. The most divergent positioning of FZD3 and FZD6 in the phylogenetic tree and the absence of the C-terminal PDZ binding motif in these genes suggests that this short motif might have originated at the root of ParaHoxozoa after first duplication event approximately 680 million years ago (Figs. 1 and 2). Comparative domain analysis has further revealed that the amino-terminal region of this family is more preserved than the carboxyl-terminal region. This might reflect fundamental biological role and strong purifying selection on amino-terminal regions of Frizzled proteins. For instance, the N-terminal extracellular CRD domain determines binding specificity for Wnt ligands and transduces Frizzled-mediated Wnt/β-catenin signaling pathway [30].

Table 1 Structural deviation in the backbone torsion angles of the wild type and mutated human FZD4 proteins

Comparison of wild-type and mutated structures	Major change in backbone torsion angles (residue number)	Major shift in regions	Critical region
P33S	44,45,110,132-135,137,141-143,147–150	CRD domain	
	285,417–432	Seven TMs domain.	
	496–537	K-T-XXX-W, PDZ binding motif	
G36D	41–46,70,71,133-138,141–150	CRD domain	
	416–431,468	Seven TMs domain.	
	495–537	K-T-XXX-W, PDZ binding motif	
E40Q	135–144,156–158	CRD domain	
	182–186,199, 202	UCR	
	243–244,248,284-286,337-338,417-420,466–469	Seven TMs domain	
	505–537	PDZ binding motif	
Y58C	134–135,141-144,147–150	CRD domain	
	243–244,248,284-285,334-337,416-431,467-469,495–497	Seven TMs domain	
	508–537	PDZ binding motif	
H69Y	147–149	CRD domain	
	285,287,334-339,415–432	Seven TMs domain	
	504–537	K-T-XXX-W, PDZ binding motif	
M105 T	274–278,284-286,467,468	Seven TMs domain	
	506–537	PDZ binding motif	495–537
M105 V	107–114,132–135	CRD domain	
	284–286,332-338,416-430,467–469.	Seven TMs domain	
	505–537	PDZ binding motif	
I114T	15–44,49,62-81,110,118,134–156	Signal peptide and CRD domain	
	213–214,274–275	Seven TMs domain	
	495–537	K-T-XXX-W, PDZ binding motif	
C117R	49,70-71,73–74	CRD domain	
	274–275	Seven TMs domain	
	496–537	K-T-XXX-W, PDZ binding motif	
R127H	15–44,49,63-81,110-111,118,133–156	Signal peptide and CRD domain	
	173–179	UCR	
	213–214,274–275	Seven TMs domain	
	496–537	K-T-XXX-W, PDZ binding motif	
M157 V	133–135,142–151	CRD domain	
	284,285,337,338,417-423,428-431,467–470	Seven TMs domain	
	496–537	K-T-XXX-W, PDZ binding motif	
C181R	419–422	Seven TMs domain	
	503–537	PDZ binding motif	
C181Y	420–423	Seven TMs domain	
	508–537	PDZ binding motif	
C204R	69–72,134,135,144–150	CRD domain	
	337,338,418-420,430,431	Seven TMs domain	
	495–537	K-T-XXX-W, PDZ binding motif	
C204Y	66–67,70-71,102,141–150	CRD domain	
	243–244,419–420	Seven TMs domain	

Table 1 Structural deviation in the backbone torsion angles of the wild type and mutated human FZD4 proteins *(Continued)*

Comparison of wild-type and mutated structures	Major change in backbone torsion angles (residue number)	Major shift in regions	Critical region
	504–537	K-T-XXX-W, PDZ binding motif	
Y211H	496–537	K-T-XXX-W, PDZ binding motif	
M223K	504–537	K-T-XXX-W, PDZ binding motif	
T237R	214,274–275	Seven TMs domain	
	503–537	K-T-XXX-W, PDZ binding motif	
R253C	504–537	K-T-XXX-W, PDZ binding motif	
W335C	274–275	Seven TMs domain	
	496–537	K-T-XXX-W, PDZ binding motif	
A370G	46–52,68-80,141	CRD domain	
	504–537	K-T-XXX-W, PDZ binding motif	
R417Q	42–46,147-150,134-137,147–150	CRD domain	
	418–432	Seven TMs domain	
	507–537	PDZ binding motif	
G488D	41–46,109-114,133–151	CRD domain	
	416–432	Seven TMs domain	
	495–537	K-T-XXX-W, PDZ binding motif	
G488 V	38,70	CRD domain	
	505–537	PDZ binding motif	
S497F	70–71,73–74	CRD domain	
	503–537	K-T-XXX-W, PDZ binding motif	
K499E	147–148	CRD domain	
	284–286, 336–338,418-431, 468–470	Seven TMs domain	
	505–537	K-T-XXX-W, PDZ binding motif	

This table shows the impact of FEVR associated missense mutations in the FZD4 protein on its backbone torsion angles by comparing them with its normal structure. The critical region encompasses the K-T-XXX-W and T/S-X-VPDZ binding motifs. In the first column, amino acid residue on the left indicates the wild-type residue; the number shows the amino acid position of the residue in the protein sequence, while the residue on the right shows the mutated residue. The second column specifies the positions at which major structural deviations were observed

The third column depicts the deviated region/residues shared among all mutant proteins analyzed (critical region)

CRD cysteine rich domain, *TM* transmembrane, *UCR* uncharacterized region

The least preserved COOH-tail of Frizzled family members is known to be inherently unstructured [31]. Such disordered regions of a given protein are expected to evolve more rapidly and exhibit low level of sequence conservation than structured regions in the same protein because they are not constrained by structure [32].

FZD4 gene is well understood among all members of the family in terms of its biological role and involvement in disease, therefore we sought to infer the protein structural basis for familial exudative vitreoretinopathy (FEVR) phenotype caused by mutations in this gene. Comparative structural analysis of human wild-type FZD4 and its mutant structures revealed that even though the FEVR-causing missense changes scattered across the FZD4 protein structure, all of them appears to impact the confirmation of the C-terminal 495–537 region through a phenomenon called "epistasis" (Fig. 3; Table 1) [33]. The functional significance of this critical region is supported by previous data, as it harbors two

functionally crucial binding sites [29]. For instance, residues 499–505 of critical region exhibits the internal cytoplasmic K-T/S-XXX-W motif that directly interacts with the PDZ domain of Dvl (Dishevelled) [34], whereas residues 535–537 of critical region contain distal C-terminal S/T-X-V PDZ-recognition motif that also interacts with the PDZ domain bearing proteins such as post-synaptic density protein (PSD) and GRIP1 [35]. The C-terminal of FZD4 is an intrinsically disordered region and therefore possesses structural plasticity [31]. Such inherently disordered region of proteins are capable of adopting a variety of structurally unrelated conformations whose features/functions are largely dependent on the cellular environment and/or on the available interacting partners [32, 36]. The effect of structural deviations in response to FEVR causing missense mutation could be deleterious either because it destabilizes structure of FZD4 or alters its intracellular localization or disrupt ligand-binding and interaction with other proteins. For instance,

amongst the identified missense variants associated with FEVR; a phenylalanine-to-serine mutation at amino acid position 328 (F328S) of intracellular loop 2 (iloop2) yields a FZD4 with a reduced capability to activate the Tcf/Lef transcription in response to Norrin. Intriguingly, this FZD4 mutant was unable to employ Dvl2 to the cell membrane through C-terminus KTXXXW motif at the transmembrane domain 7 proximal segment of the cytosolic tail. Based on these biochemical data it can be suggested that iloop2 of FZD4 may modulate interactions between Frizzled and other signaling molecules through distantly located carboxy-terminus region [37]. Furthermore, a deletion of two nucleotides that led to a frameshift and created a stop codon at amino-acid 533 (L501fsX533) resulted in abnormal PDZ binding motif at C-terminus of FZD4 and hence leading to FEVR [31]. Biochemical investigations show that this particular frameshift mutation (L501fsX533) generates a disorder-to-order structural transition in the C-terminal cytosolic tail of FZD4. The resulting mutant FZD4 protein fails to reach the Plasma Membrane of the cells and accumulates in the Endoplasmic Reticulum [31]. These previously reported functional data supports the prediction that the critical region (amino acid positions 495-537at the C-terminus) of FZD4 protein identified in the present study is crucial for normal function and any conformational change in this region contributes to the disease pathology. Comparative structural analysis also revealed structural deviation in cysteine-rich domain (CRD) and seven transmembrane domains of subset of FZD4 mutant proteins analyzed. In particular, the N-terminal extracellular CRD domain is conserved among Frizzled family members and determines binding specificity for Wnt ligands [30]. For instance, the extracellular ligand Norrin binds specifically to the CRD domain of FZD4 but not the CRDs of other members of mammalian Frizzleds [38]. This FZD4-Norrin binding results in Wnt/β-catenin pathway activation in normal retinal vascular development [30, 38]. Missense mutations within and near the highly conserved CRD domain are known to affect correct protein folding of the CRD and disrupt Norrin binding, consequently, causing FEVR [30]. Therefore, the protein conformational changes in CRD domain of FZD4 mutants can potentially affect retinal development and contributes to disease pathogenesis.

Conclusion

Frizzled receptor family is an evolutionary conserved multi-gene family with a fundamental biological role in metazoans development. The phylogenetic data suggests the origin of Frizzled receptor family in early metazoans and subsequently diversified at different time points during bilaterians history. The evolutionary investigations reveal that Frizzled receptors have undergone extensive functional diversification through complex pattern of gene duplications

and sequence divergence. Comparative structural analysis of the FZD4 protein pinpoints a critical region (amino acids 495–537) with crucial implications in normal cellular function and disease pathogenesis (FEVR). These findings provide a base for future structural and evolutionary studies to elucidate further the role of Frizzled receptors in metazoans evolution, development and disease.

Methods

Members of the Frizzled receptor family and their human protein sequences were identified and extracted from Ensembl Genome Browser [39]. In total, 59 putative orthologous protein sequences were retrieved from Ensemble Genome Browser and National Center for Biotechnology Information (NCBI) by using BLASTp bidirectional searches [40, 41]. Further confirmation of the common ancestry of the putative orthologs was obtained by clustering homologous proteins within phylogenetic trees. Sequences whose position within a tree was in sharp conflict with the uncontested animal phylogeny were excluded from the analysis.

The species that were selected for sequence analysis includes *Homo sapiens* (human), *Callithrix jacchus* (marmoset), *Mus musculus* (mouse), *Rattus norvegicus* (rat), *Equus caballus* (horse), *Monodelphis domestica* (opossum), *Gallus gallus* (chicken), *Anolis carolinensis* (anole lizard), *Danio rerio* (zebrafish), *Takifugu rubripes* (fugu), *Tetraodon nigroviridis* (tetraodon), *Gasterosteus aculeatus* (stickleback), *Oryzias latipes* (medaka), *Lepisosteus oculatus* (spotted gar), *Petromyzon marinus* (lamprey), *Latimeria chalumnae* (coelacanth), *Ciona intestinalis* (C.intestinalis), *Ciona savignyi* (C.savignyi), *Drosophila melanogaster* (fruitfly), *Caenorhabditis elegans* (C. elegans), *Capitella teleta* (capitella), *Helobdella robusta* (Californian leech), *Octopus bimaculoides* (octopus), *Nematostella vectensis* (starlet sea anemone), *Trichoplax adhaerens* (trichoplax).

Sequence analysis

Phylogenetic analysis of Frizzled receptor family was performed using Neighbor-joining method (NJ) [42]. Uncorrected p-distance and Poisson corrected (PC) amino acid distance were used as amino acid substitution models [43]. Both models produced similar results, therefore, only NJ tree based on uncorrected p-distance is presented here (Fig. 1). Maximum likelihood (ML) tree based on Whelen and Golman (WAG) amino acid substitution model was also constructed (Additional file 1) [44, 45]. To test the reliability and correctness of NJ and ML trees, bootstrap method (at 1000 pseudo replicates) was used which produced the bootstrap value for each internal branch in trees [46].

Comparative protein domain analysis was conducted to assign domains, motifs and sub-motifs to each

member of Frizzled receptor family [20, 29, 47, 48]. The positions of domains and motifs on each putative paralogs were identified and mapped by employing a multiple sequence alignment tool, the Clustal Omega [49]. Approximate scaling and positioning of the respective domains and motifs was also carried out. Previously reported disease causing mutations in FZD family members were also mapped.

Structural analysis

To date, approximately 47 different FEVR pathogenic mutations have been associated with human *FZD4*, of which ~ 40 are missense variants dispersed across the protein structure [10, 11]. These 40 missense mutations are targeted for protein structural and conformation analyses (G22E, P33S, G36D, E40Q, C45Y, Y58C, H69Y, M105 T, M105 V, C106G, I114T, C117R, R127H, M157 V, M157K, E180K, C181R, C181Y, K203 N, C204R, C204Y, Y211H, M223K, T237R, R253C, I256V, W335C, A339T, M342 V, A370G, R417Q, T445P, R466W, D470N G488D, G488 V, G492R, S497F, K499E, G525R) [50–52]. For this purpose, the structure of wild-type FZD4 protein sequence predicted through I-TASSER and taken as a reference model to study the impact of selected subset of missense mutations on overall conformation of protein [22]. The mutant structures of FZD4 are predicted through MODELLER (version 9.14) [53]. Best structures are selected on the basis of Discrete Optimized Protein Energy (DOPE) score. Mutant models are superimposed over wild-type FZD4 by using Chimera and Root Mean Square Deviation (RMSD) values are evaluated [24, 54]. All forty mutant protein structures have RMSD values > 0.4 Å. The effect of these mutations on protein stability was also investigated through MuPro [55]. The overall model quality of the predicted proteins are analyzed through structure validation tools ERRAT and Ramachandran plot [23, 56].

Additional files

Additional file 1: Phylogenetic analysis of frizzled receptors gene family. (PDF 505 kb)

Additional file 2: Structural comparison of wild type and Familial Exudative Vitreoretinopathy mutated FZD4. (PDF 1958 kb)

Additional file 3: Evaluation of 3D models of mutant FZD4 proteins. (PDF 6199 kb)

Abbreviations
CRD: Cystein Rich Domain; DOPE: Discrete Optimized Protein Energy; Dvl: Disheveled; FEVR: Familial exudative vitreoretinopathy; FZDs: Frizzleds; GAG: Glycosaminoglycan; GPCRs: G-protein coupled receptors; ML: Maximum Likelihood; NCBI: National Center for Biotechnology Information; NJ: Neighbor-joining; PC: Poisson corrected; PCP: Planer cell polarity; RMSD: Root Mean Square Deviation; TM: Transmembrane; WAG: Whelan and Goldman

Availability of data and materials
The datasets analyzed during the current study are available in the Ensembl database (http://www.ensembl.org), NCBI database (https://www.ncbi.nlm.nih.gov/).

Authors' contributions
AAA conceived the project. AAA and YB designed the experiments. SS and NP performed the experiments. AAA, SS, YB, RZ, SA and NP analyzed the data. AAA, SS and YB wrote the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Venkatakrishnan A, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. Nature. 2013; 494(7436):185.
2. Millar RP, Newton CL. The year in G protein-coupled receptor research. Mol Endocrinol. 2010;24(1):261–74.
3. Terrillon S, Bouvier M. Roles of G-protein-coupled receptor dimerization: from ontogeny to signalling regulation. EMBO Rep. 2004;5(1):30–4.
4. Fredriksson R, Lagerström MC, Lundin L-G, Schiöth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol. 2003;63(6):1256–72.
5. Hanyaloglu AC, Zastrow M. Regulation of GPCRs by endocytic membrane trafficking and its potential implications. Annu Rev Pharmacol Toxicol. 2008;48: 537–68.
6. Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein–coupled receptor superfamily. Annu Rev Pharmacol Toxicol. 2013;53:531–56.
7. Peifer M. Signal transduction: neither straight nor narrow. Nature. 1999; 400(6741):213.
8. Pan C-L, Howell JE, Clark SG, Hilliard M, Cordes S, Bargmann CI, Garriga G. Multiple Wnts and frizzled receptors regulate anteriorly directed cell and growth cone migrations in Caenorhabditis elegans. Dev Cell. 2006;10(3):367–77.
9. Huang H-C, Klein PS. The Frizzled family: receptors for multiple signal transduction pathways. Genome Biol. 2004;5(7):234.
10. Robitaille JM, Zheng B, Wallace K, Beis MJ, Tatlidil C, Yang J, Sheidow TG, Siebert L, Levin AV, Lam W-C. The role of Frizzled-4 mutations in familial exudative vitreoretinopathy and Coats disease. Br J Ophthalmol. 2010;95: 574–9.
11. Toomes C, Bottomley HM, Scott S, Mackey DA, Craig JE, Appukuttan B, Stout JT, Flaxel CJ, Zhang K, Black GC. Spectrum and frequency of FZD4 mutations in familial exudative vitreoretinopathy. Invest Ophthalmol Vis Sci. 2004;45(7):2083–90.
12. Milhem RM, Ben-Salem S, Al-Gazali L, Ali BR. Identification of the cellular mechanisms that modulate trafficking of frizzled family receptor 4 (FZD4) missense mutants associated with familial exudative vitreoretinopathy. Invest Ophthalmol Vis Sci. 2014;55(6):3423–31.

13. Wang Y, Chang H, Rattner A, Nathans J. Frizzled receptors in development and disease. Curr Top Dev Biol. 2016;117:113–9.

14. Shastry BS. Genetics of familial exudative vitreoretinopathy and its implications for management. Expert Rev Ophthalmol. 2012;7(4):377–86.

15. MacDonald BT, He X. Frizzled and LRP5/6 receptors for Wnt/β-catenin signaling. Cold Spring Harb Perspect Biol. 2012;4(12):a007880.

16. Pei J, Grishin NV. Cysteine-rich domains related to Frizzled receptors and Hedgehog-interacting proteins. Protein Sci. 2012;21(8):1172–84.

17. Ramasarma T. Transmembrane domains participate in functions of integral membrane proteins. Indian J Biochem Biophys. 1996;33(1):20–9.

18. Gayen S, Li Q, Kim YM, Kang C. Structure of the C-terminal region of the Frizzled receptor 1 in detergent micelles. Molecules. 2013;18(7):8579–90.

19. Magdesian MH, Carvalho MM, Mendes FA, Saraiva LM, Juliano MA, Juliano L, Garcia-Abreu J, Ferreira ST. Amyloid-β binds to the extracellular cysteine-rich domain of Frizzled and inhibits Wnt/β-catenin signaling. J Biol Chem. 2008; 283(14):9359–68.

20. Rehn M, Pihlajaniemi T, Hofmann K, Bucher P. The frizzled motif: in how many different protein families does it occur? Trends Biochem Sci. 1998; 23(11):415–7.

21. Wong H-C, Bourdelas A, Krauss A, Lee H-J, Shao Y, Wu D, Mlodzik M, Shi D-L, Zheng J. Direct binding of the PDZ domain of Dishevelled to a conserved internal sequence in the C-terminal region of Frizzled. Mol Cell. 2003;12(5):1251–60.

22. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9(1):40.

23. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993;2(9):1511–9.

24. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12.

25. Abbasi AA. Molecular evolution of HR, a gene that regulates the postnatal cycle of the hair follicle. Sci Rep. 2011;1:32.

26. Yu H, Ye X, Guo N, Nathans J. Frizzled 2 and frizzled 7 function redundantly in convergent extension and closure of the ventricular septum and palate: evidence for a network of interacting genes. Development. 2012;139(23):4383–94.

27. Zhang R, Fang Y, Wu B, Chemban M, Laakhey N, Cai C, Shi O. Gene-gene interaction between VANGL1, FZD3, and FZD6 correlated with neural tube defects in Han population of Northern China. Genet Mol Res. 2016;15(3). https://doi.org/10.4238/gmr.15039010.

28. Stuebner S, Faus-Kessler T, Fischer T, Wurst W, Prakash N. Fzd3 and Fzd6 deficiency results in a severe midbrain morphogenesis defect. Dev Dyn. 2010;239(1):246–60.

29. Romero G, von Zastrow M, Friedman PA. Role of PDZ proteins in regulating trafficking, signaling, and function of GPCRs: means, motif, and opportunity. Adv Pharmacol. 2011;62:279.

30. Zhang K, Harada Y, Wei X, Shukla D, Rajendran A, Tawansy K, Bedell M, Lim S, Shaw PX, He X. An essential role of the cysteine-rich domain of FZD4 in Norrin/Wnt signaling and familial exudative vitreoretinopathy. J Biol Chem. 2011;286(12):10210–5.

31. Lemma V, D'agostino M, Caporaso MG, Mallardo M, Oliviero G, Stornaiuolo M, Bonatti S. A disorder-to-order structural transition in the COOH-tail of Fz4 determines misfolding of the L501fsX533-Fz4 mutant. Sci Rep. 2013;3:2659.

32. Dunker AK, Silman I, Uversky VN, Sussman JL. Function and structure of inherently disordered proteins. Curr Opin Struct Biol. 2008;18(6):756–64.

33. Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. Crystal structure of an ancient protein: evolution by conformational epistasis. Science. 2007; 317(5844):1544–8.

34. Chen W, ten Berge D, Brown J, Ahn S, Hu LA, Miller WE, Caron MG, Barak LS, Nusse R, Lefkowitz RJ. Dishevelled 2 recruits ß-arrestin 2 to mediate Wnt5A-stimulated endocytosis of Frizzled 4. Science. 2003;301(5638):1391–4.

35. Bian WJ, Miao WY, He SJ, Wan ZF, Luo ZG, Yu X. A novel Wnt5a-Frizzled4 signaling pathway mediates activity-independent dendrite morphogenesis via the distal PDZ motif of frizzled 4. Dev Neurobiol. 2015;75(8):805–22.

36. Uversky VN. A protein-chameleon: conformational plasticity of α-synuclein, a disordered protein involved in neurodegenerative disorders. J Biomol Struct Dyn. 2003;21(2):211–34.

37. Pau MS, Gao S, Malbon CC, Wang H-y, Bertalovitz AC. The intracellular loop 2 F328S Frizzled-4 mutation implicated in familial exudative vitreoretinopathy impairs Dishevelled recruitment. J Mol Signal. 2015;10:5.

38. Smallwood PM, Williams J, Xu Q, Leahy DJ, Nathans J. Mutational analysis of Norrin-Frizzled4 recognition. J Biol Chem. 2007;282(6):4057–68.

39. Fernández XM, Birney E. Ensembl Genome Browser. In: Vogel and Motulsky's Human Genetics. Berlin Heidelberg: Springer; 2010. p. 923–939.

40. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

41. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2007; 36(suppl_1):D13–21.

42. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4(4):406–25.

43. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A. 2004;101(30):11030–5.

44. Goldman N, Yang Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol. 1994;11(5):725–36.

45. Saitou N, Imanishi T. Relative efficiencies of the Fitch-Margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. Mol Biol Evol. 1989;6(5):514–25.

46. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985;39(4):783–91.

47. Hering H, Sheng M. Direct interaction of Frizzled-1,-2,-4, and -7 with PDZ domains of PSD-95. FEBS Lett. 2002;521(1–3):185–9.

48. Wheeler DS, Barrick SR, Grubisha MJ, Brufsky AM, Friedman PA, Romero G. Direct interaction between NHERF1 and frizzled regulates β-catenin signaling. Oncogene. 2011;30(1):32–42.

49. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7(1):539.

50. Kondo H, Hayashi H, Oshima K, Tahira T, Hayashi K. Frizzled 4 gene (FZD4) mutations in patients with familial exudative vitreoretinopathy with variable expressivity. Br J Ophthalmol. 2003;87(10):1291–5.

51. Robitaille JM, Wallace K, Zheng B, Beis MJ, Samuels M, Hoskin-Mott A, Guernsey DL. Phenotypic overlap of familial exudative vitreoretinopathy (FEVR) with persistent fetal vasculature (PFV) caused by FZD4 mutations in two distinct pedigrees. Ophthalmic Genet. 2009;30(1):23–30.

52. Kondo H. Complex genetics of familial exudative vitreoretinopathy and related pediatric retinal detachments. Taiwan J Ophthalmol. 2015;5(2):56–62.

53. Webb B, Sali A. Protein structure modeling with MODELLER. In: Kihara D, editor. Protein struct prediction. Methods in molecular biology (methods and protocols), vol. 1137. New York: Humana Press; 2014. p. 1–15.

54. Maiorov VN, Crippen GM. Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. J Mol Biol. 1994;235(2):625–34.

55. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins: structure, function. Bioinformatics. 2006;62(4):1125–32.

56. Hooft RW, Sander C, Vriend G. Objectively judging the quality of a protein structure from a Ramachandran plot. Bioinformatics. 1997;13(4):425–30.