

---

# Comparative Genomics: Frizzled Receptors

---

William Mwine  
r0993335  
Master of Bioinformatics  
Academic year 2024-2025

# 1 Introduction

Frizzled proteins are transmembrane proteins that play a key role in the Wnt signaling pathway as well as other signaling pathways. They bind to Wnt proteins and pass the signal to dishevelled proteins during the signaling process. At their extracellular amino terminus, frizzled proteins have a cysteine-rich domain that binds Wnt ligands, while the intracellular carboxyl terminus contains a KTXXXW motif essential for Wnt signaling. Huang and Klein, 2004. These features are highly conserved across species highlighting their importance. In this analysis, I will focus on the Frizzled 7(FZD7) protein, investigating its conservation across species. This study aims to provide insights into its evolutionary history and the preservation or diversification of its function over time.

## 2 Methods

### 2.1 Sequences

The sequences used for the downstream analyses were acquired by searching for homologs of the human FZD7 protein. These homologous sequences were obtained using the NCBI BLASTP tool with default settings, except for the `max_target_seqs` parameter, which was set to 250 to retrieve up to 250 sequences. The results were then filtered to remove synthetic constructs, low-quality sequences, and predicted sequences. From the remaining sequences, 100 were randomly sampled and used for downstream analyses.

### 2.2 Multiple sequence alignment

The multiple sequence alignment (MSA) was constructed using ClustalW2 with the slow and accurate parameter.

### 2.3 Phylogenetic Tree construction

The phylogenetic tree of the sequences from subsection 2.1 was constructed using IQ-TREE multicore version 2.3.6. IQ-TREE determined the best model for the MSA to be `Q.mammal+I+G4` based on the Bayesian Information Criterion (BIC). This model uses the `Q.mammal` amino acid exchange rate matrix Minh et al., 2021 along with an invariable site plus discrete Gamma model to account for rate heterogeneity across sites.

To annotate this tree with evolutionary events like duplication and speciation, a species tree of the species of the FZD7 protein sequences from subsection 2.1 was constructed. As recommended by Tobe et al., 2010, the mitochondrial *cytochrome b* (*cyt b*) gene was used to construct the species tree due to its sufficient intraspecific and interspecific variation, which enables effective species differentiation. Whole mitochondrial genome sequences were obtained from NCBI. The *cyt b* gene was then extracted from these sequences and used to construct an MSA using clustaw. Similar to subsection 2.3, IQ-TREE was used to select the best model for our MSA based on the Bayesian Information Criterion (BIC). The `GTR+F+I+R5` model, with 1000 bootstraps, was selected. This model uses the General Time Reversible (GTR) substitution model with unequal rates and unequal base frequencies, along with empirical base frequencies, and accounts for invariable sites plus a FreeRate model. This model was then used to construct the species tree. The FZD7 protein tree was then annotated using Notung-2.9 by Stolzer et al., n.d. and is based on algorithms Stolzer et al., 2012, Darby et al., 2016, Durand et al., 2006, Lai et al., 2017, Vernot et al., 2008

## 2.4 Conserved regions

Jalview Procter et al., 2024 was used to visualize the FZD7 sequence alignment and to calculate Thompson conservation scores Thompson et al., 1999 for all columns.

## 3 Results

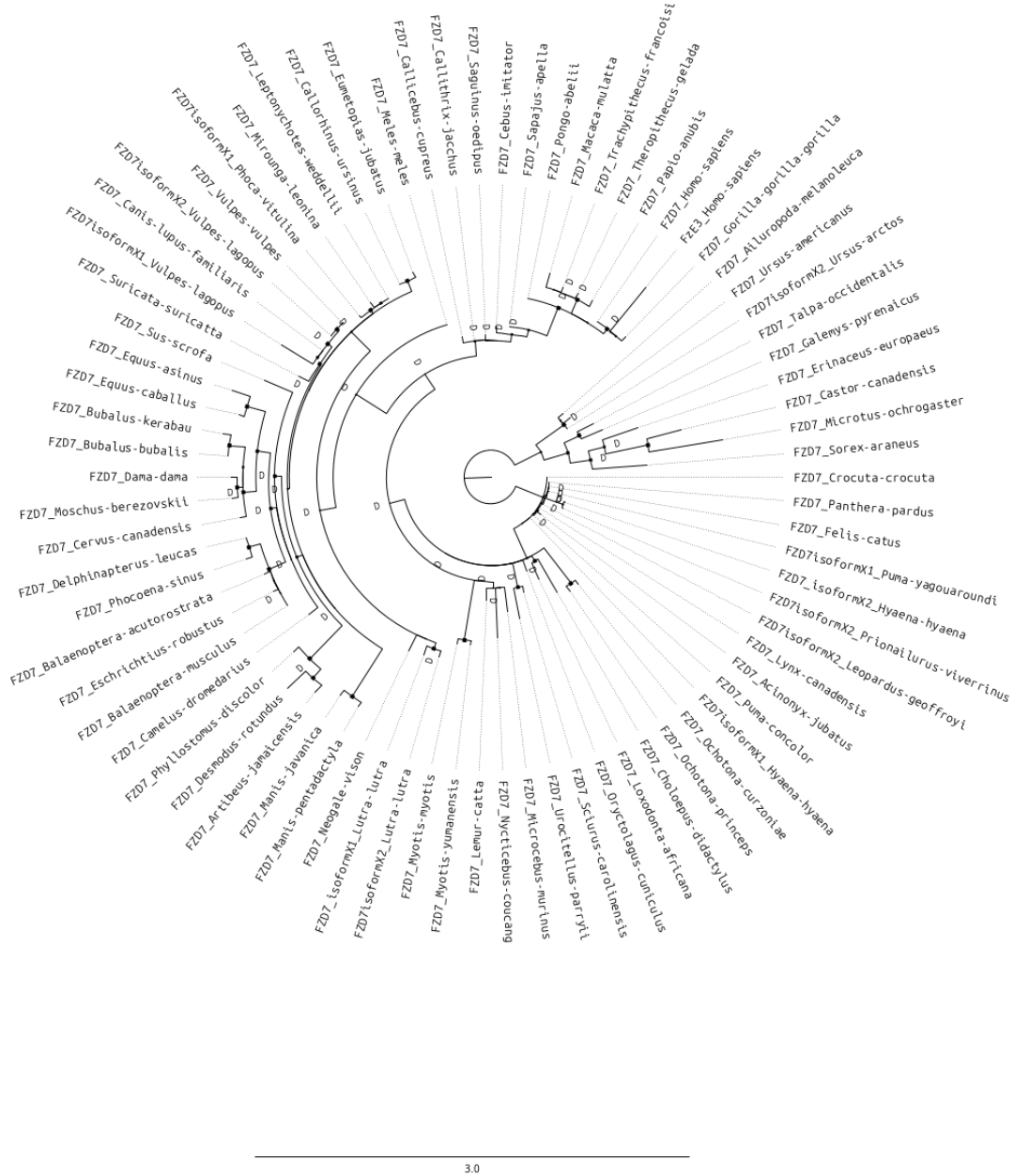


Figure 1: Phylogenetic tree of Frizzled 7 protein sequences. Nodes labeled with "D" represent duplication events, while unlabelled nodes indicate speciation events. The shape of the nodes reflects bootstrap values: nodes with clear shapes have high bootstrap values (greater than 90), while unclear or absent node shapes correspond to bootstrap values below 50. This tree demonstrates the evolutionary relationships and gene duplication events in FZD7 receptors.



Region	Position	Consensus Logo	Conservation score	GO:Terms	Function
region1	47-164	Figure 3	0.98	GO:0017147, GO:0042813	Wnt protein binding, Wnt receptor activity
region2	206-549	Figure 4	0.99	GO:0004888, GO:0042813	transmembrane signaling receptor activity, Wnt receptor activity
region3	552-557	Figure 5	1.0		activation of the Wnt- $\beta$ -catenin pathway Huang and Klein, 2004

Table 1: Conserved Regions Analysis. This table lists key conserved regions in the FZD7 protein, their positions, average Thompson conservation scores, and associated functional roles. The consensus logos (Figures 3-5) illustrate sequence conservation in these regions.

## 4 Discussion

From the analysis of FZD7 receptors, an evolutionary history characterized by both duplication and speciation events can be seen. Duplication events have led to the emergence of paralogs, which subsequently underwent processes such as neofunctionalization, specialization, subfunctionalization, or other evolutionary fates. As depicted in Figure 1, the FZD7 protein in *Homo sapiens* is a paralog to FZD7 proteins from primate species like *Gorilla gorilla gorilla* (western lowland gorilla) and *Pongo abelii* (Sumatran orangutan), as well as species like *Canis lupus familiaris* (domesticated dog). Isoforms are however also annotated as paralogs in the phylogenetic tree for example FZD7 isoformX1 and FZD7 isoformX2 from the Eurasian otter (*Lutra lutra*). However, this classification is inconsistent with the evolutionary definition of paralogs. Isoforms are distinct proteins produced from the same gene through alternative splicing of pre-mRNA while paralogs arise from duplicated genes. Orthologous relationships are observed with FZD7 proteins from species such as *Sapajus apella* and *Trachypithecus francoisi*.

The conservation of FZD7 function across diverse species highlights the critical role of frizzled receptors. As mentioned in section 1, frizzled receptors are integral to Wnt signaling, which regulates the expression of genes involved in cell proliferation and differentiation Liu et al., 2022. The conserved regions of FZD7 are shown in Table 1, along with their respective functions. Region 1 represents the cysteine-rich domain, which is highly conserved in frizzled receptors across species. This domain is responsible for the binding of Wnt ligands Huang and Klein, 2004. Region 2 refers to the seven linked transmembrane hydrophobic domains, and just after it is Region 3, which denotes the KTXXXW motif. Region 3 plays a role in the activation of the Wnt- $\beta$ -catenin pathway Huang and Klein, 2004.

In conclusion, this analysis reveals a rich evolutionary history of FZD7 receptors marked with duplication and speciation events. Despite these evolutionary events, FZD7 has preserved its function across species with key conserved regions such as the cysteine rich domain and KTXXXW motif which play crucial roles in Wnt signaling.

## References

- Darby, C. A., Stolzer, M., Ropp, P. J., Barker, D., & Durand, D. (2016). Xenolog classification. *Bioinformatics*, 33(5), 640–649.
- Durand, D., Halldorsson, B. V., & Vernot, B. (2006). A hybrid micro-macroevolutionary approach to gene tree reconstruction. *Journal of Computational Biology*, 13(2), 320–335.

- Huang, H.-C., & Klein, P. S. (2004). The frizzled family: Receptors for multiple signal transduction pathways. *Genome Biol.*, 5(7), 234.
- Lai, H., Stolzer, M., & Durand, D. (2017). Fast heuristics for resolving weakly supported branches using duplication, transfers, and losses. *Proceedings for RECOMB International Workshop on Comparative Genomics*, 298–320.
- Liu, J., Xiao, Q., Xiao, J., Niu, C., Li, Y., Zhang, X., Zhou, Z., Shu, G., & Yin, G. (2022). Wnt/ $\beta$ -catenin signalling: Function, biological mechanisms, and therapeutic opportunities. *Signal Transduct. Target. Ther.*, 7(1), 3.
- Minh, B. Q., Dang, C. C., Vinh, L. S., & Lanfear, R. (2021). QMaker: Fast and accurate method to estimate empirical models of protein evolution. *Syst. Biol.*, 70(5), 1046–1060.
- Procter, J., Carstairs, M., Soares, B., Mourao, K., Ofoegbu, T. ', Waterhouse, A., Engelhardt, J., Lui, L., Menard, A., Barton, D., Sherstnev, N., Roldan-Martinez, D., Clamp, M., Cuff, J., Searle, S., Martin, D., & Barton, G. (2024). *Jalview* (Version 2.11.4.1) [Built: 21:35:40 29 October 2024 from Source git-commit:6d1fc3490c [releases/Release\_2.11.4.Branch]]. <https://www.jalview.org/>
- Stolzer, M., Lai, H., Xu, M., Sathaye, D., Vernot, B., & Durand, D. (2012). Inferring duplications, losses, transfers, and incomplete lineage sorting with non-binary species trees. *Bioinformatics*, 28, i409–i415.
- Stolzer, M., Lai, H., Vernot, B., Darby, C., Xu, M., Goldman, A., Sathaye, D., & Durand, D. (n.d.). Notung-2.9. <http://www.cs.cmu.edu/~durand/Notung/index.html>
- Thompson, J. D., Plewniak, F., & Poch, O. (1999). A comprehensive comparison of multiple sequence alignment programs. *Nucleic Acids Research*, 27(13), 2682–2690. <https://doi.org/10.1093/nar/27.13.2682>
- Tobe, S. S., Kitchener, A. C., & Linacre, A. M. T. (2010). Reconstructing mammalian phylogenies: A detailed comparison of the cytochrome B and cytochrome oxidase subunit I mitochondrial genes. *PLoS One*, 5(11), e14156.
- Vernot, B., Stolzer, M., Goldman, A., & Durand, D. (2008). Reconciliation with non-binary species trees [Also appeared in Computational Systems Bioinformatics: CSB2007 Conference Proceedings, Imperial College Press: 441–452]. *Journal of Computational Biology*, 15(8), 981–1006.

## A Appendix

### Code

```
# Alignment
clustaw2
# Tree Generation
## Species tree
iqtree2 -s Cytb_seq.aln --seqtype DNA -m MFP -bb 1000 -wbt -alrt 1000 -nt
↪ AUTO
## FZD7 phylogenetic tree
iqtree2 -s FZDZ_seq.aln --seqtype AA -m MFP -bb 1000 -wbt -alrt 1000 -nt AUTO
```

### Extra Figures



Figure 3: Cysteine rich domain consensus logo. Generated by [WebLogo](#)



Figure 4: Hydrophobic domains consensus logo. Generated by [Skyline](#)

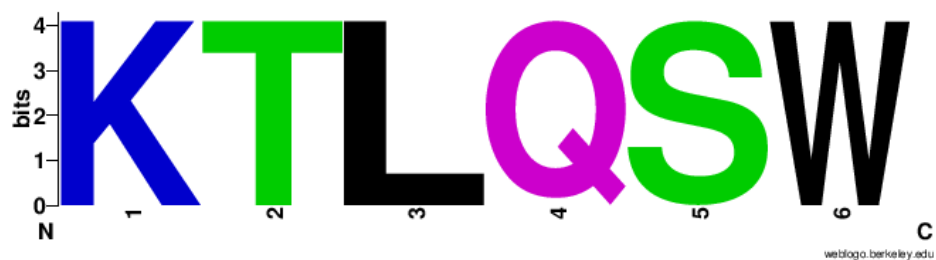


Figure 5: KTXXXW motif consensus logo. Generated by [WebLogo](#)