

# **AMPA Receptor Auxiliary Subunits Emerged During Vertebrate Evolution by Neo/Subfunctionalization of Unrelated Proteins**

David Ramos-Vicente<sup>1,2</sup> and Àlex Bayés<sup>1,2\*</sup>

1. Molecular Physiology of the Synapse Laboratory, Biomedical Research Institute Sant Pau, Barcelona, Spain.
2. Universitat Autònoma de Barcelona, Barcelona, Spain.

\* Corresponding author

Àlex Bayés, [abayesp@santpau.cat](mailto:abayesp@santpau.cat)

Molecular Physiology of the Synapse Laboratory

IIB Sant Pau

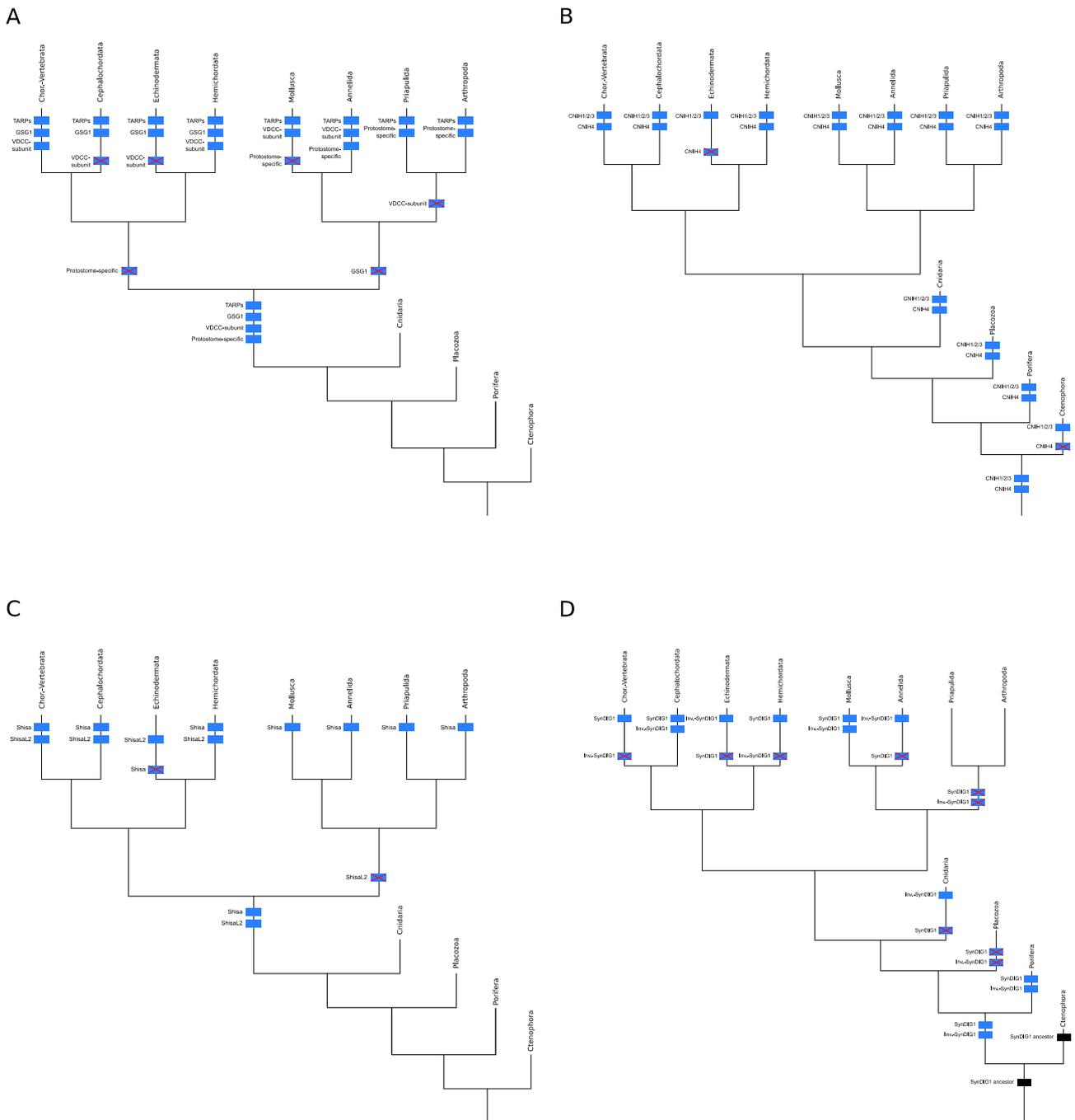
C/Sant Quintí, 77-79

08041 Barcelona, Spain



**Supplementary Figure 1. Maximum likelihood phylogenetic tree of CACNG-GSG1 protein family.**

The subfamilies in which the CACNG-GSG1 protein family is divided are highlighted by dashed line boxes. The name of each subfamily is presented at the right of the box. Bootstrap values are shown at tree nodes and protein names at the end of each branch. Tree branches are coloured based on phylum, as indicated in the legend. The closest relatives to vertebrate CACNG-GSG1 in the Claudin superfamily were used as outgroup. If known, the function of vertebrate GSG1 subfamily sequences was marked on the right of sequence name. Scale bar denotes number of amino acid substitutions per site. The amino-acid substitution model used was Vt+G+F, branch support was obtained after 1000 iterations of ultrafast bootstrapping.

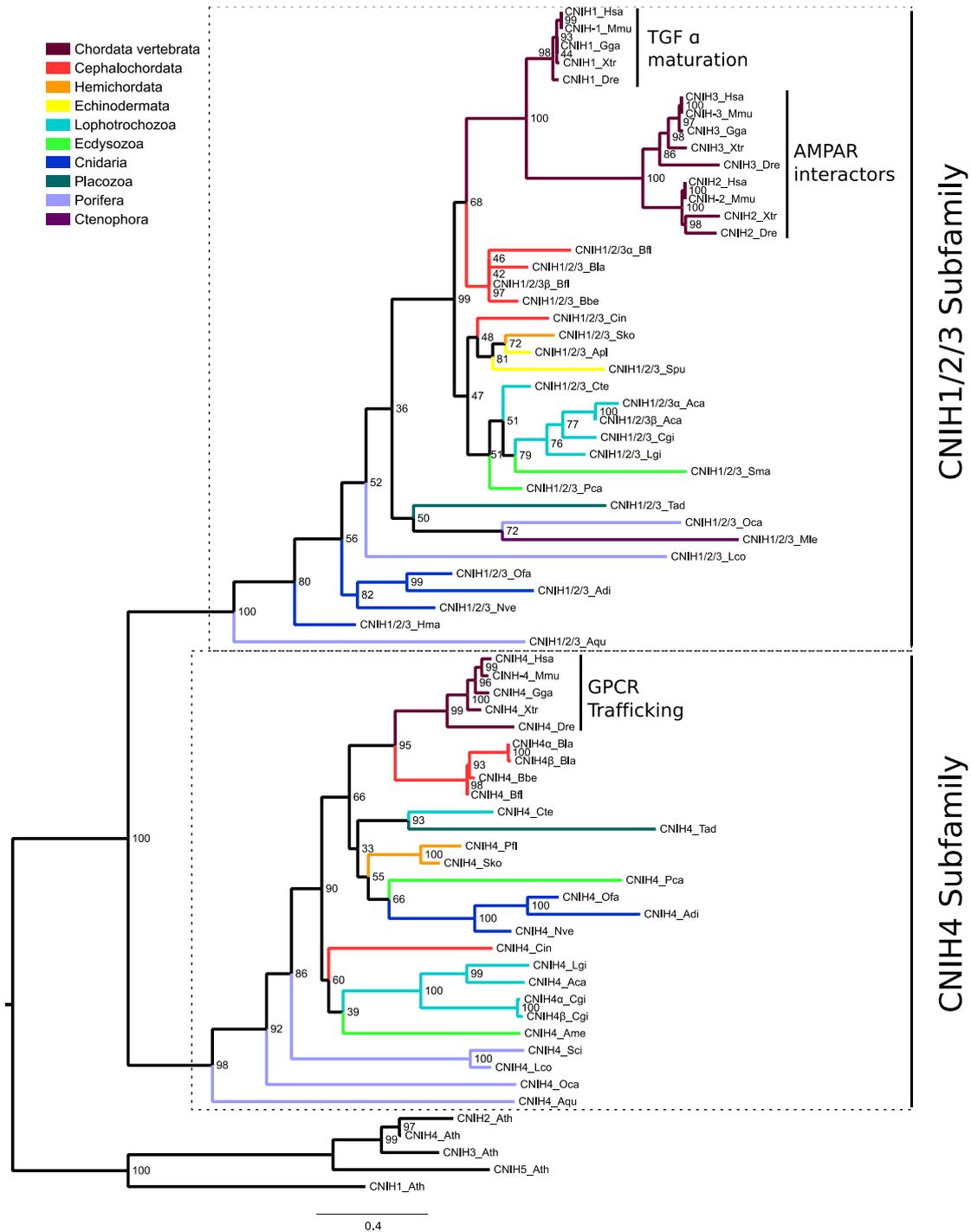


**Supplementary Figure 2. Schemes of the evolution of the three protein families and the Dispanin C subfamily.**

Schematic trees showing the evolution of (A) CACNG-GSG1, (B) Cornichon, (C) Sisha and (D) Dispanin C families. Each branch corresponds with one lineage. Phylogenetic subfamilies, and classes in the case of Dispanin C tree, are represented by blue boxes. The ancestral SynDIG1 gene is represented by a black box. When a subfamily or class is lost in a lineage or in an ancestor, the corresponding box is crossed out by a red cross. Mollusca and annelida are lophotrochozoans and priapulida and arthropoda ecdysozoans.



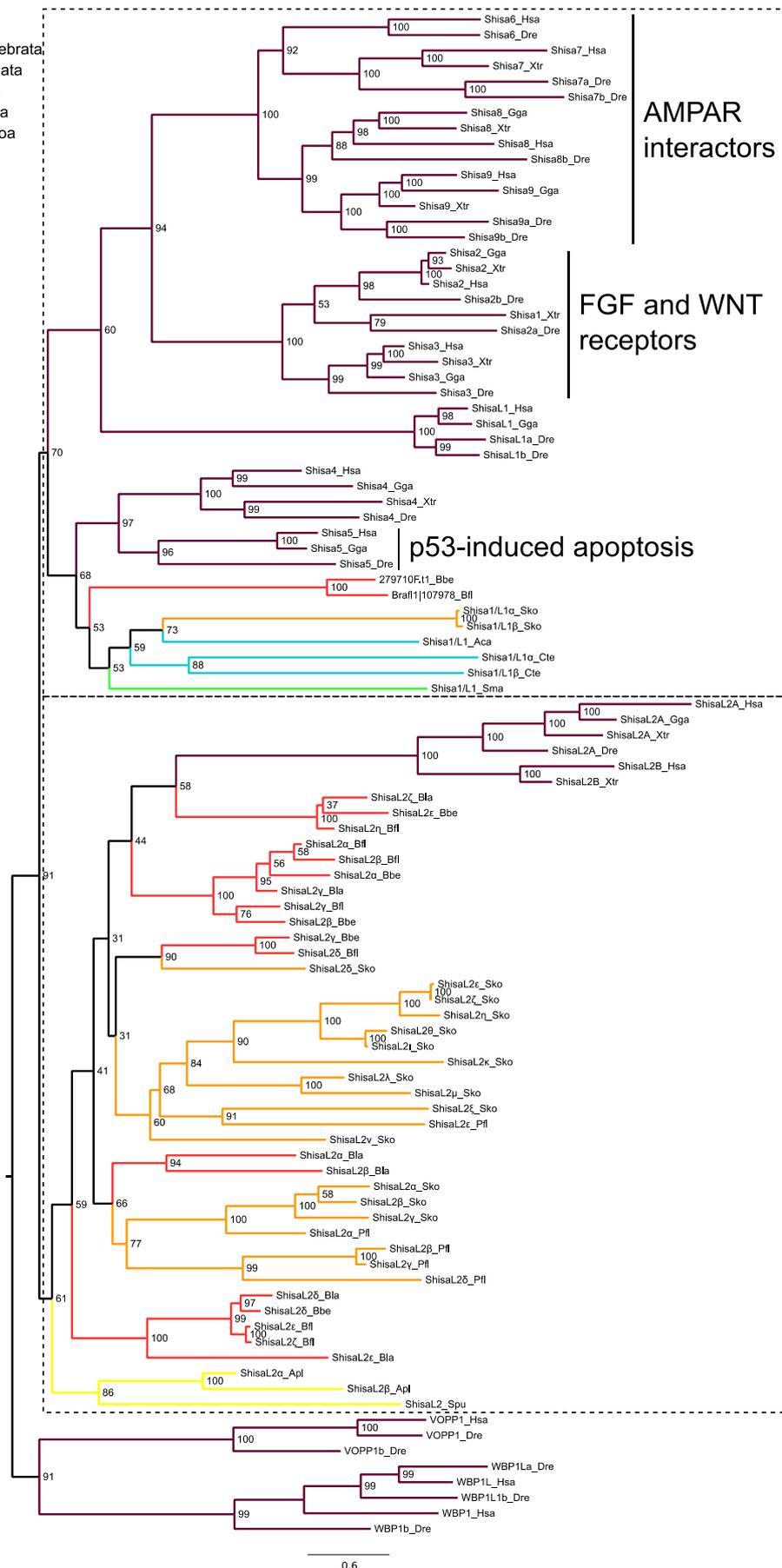




**Supplementary Figure 5. Maximum likelihood phylogenetic tree of Cornichon protein family.**

The two subfamilies in which the Cornichon family is divided are highlighted by dashed line boxes. The name of each subfamily is presented at the right of the box. Bootstrap values are shown at tree nodes and protein names at the end of each branch. Tree branches are coloured based on phylum, as indicated in the legend. Cornichon proteins from *Arabidopsis thaliana* were used as outgroup. The function of vertebrate sequences is indicated. Scale bar denotes number of amino acid substitutions per site. The amino-acid substitution model used was Lg+G, branch support was obtained after 1000 iterations of ultrafast bootstrapping.

- Chordata vertebrata
- Cephalochordata
- Hemichordata
- Echinodermata
- Lophotrochozoa
- Ecdysozoa

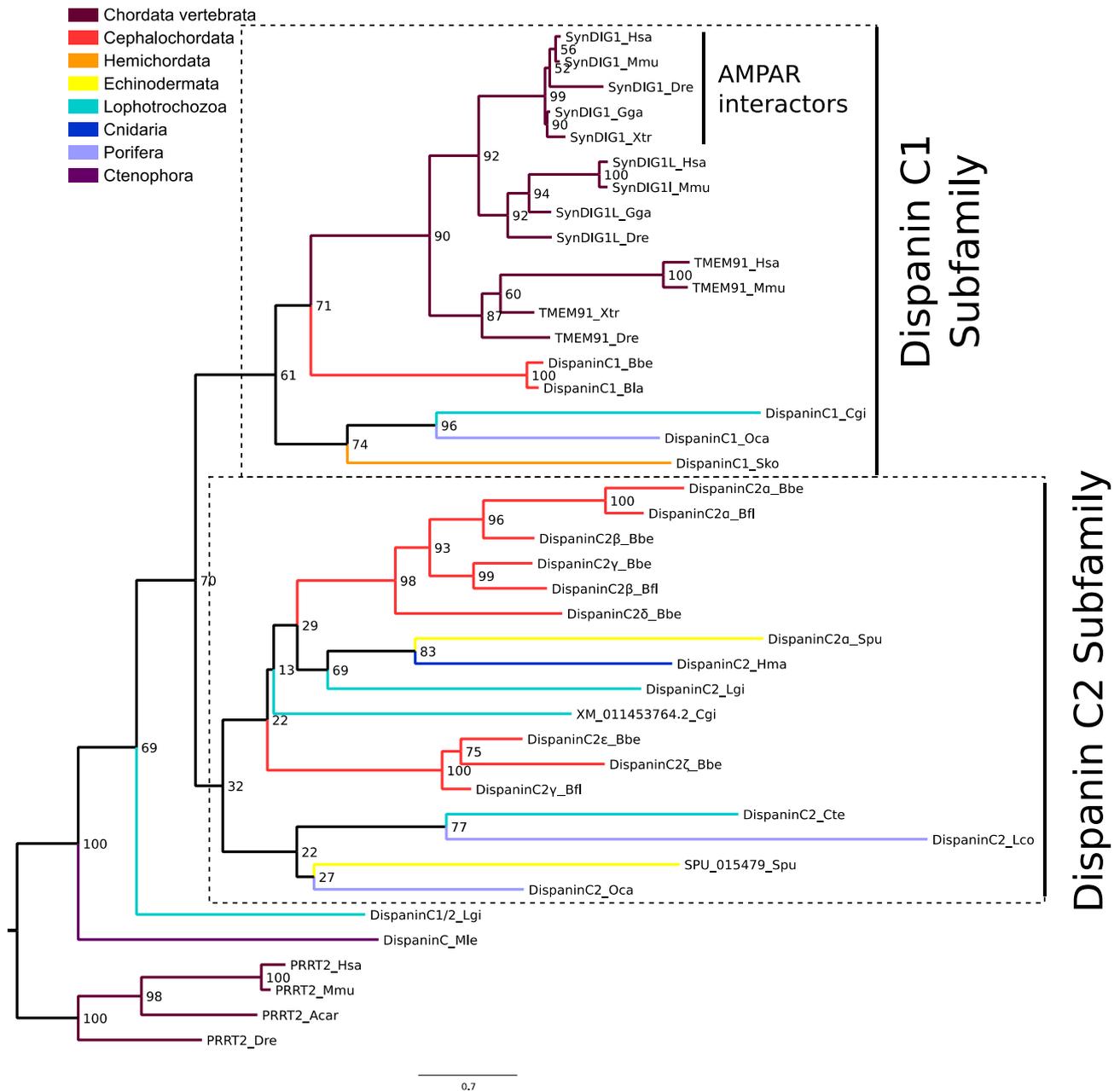


Shisa1/L1 Subfamily

ShisaL2 Subfamily

**Supplementary Figure 6. Maximum likelihood phylogenetic tree of Shisa protein family.**

The two subfamilies in which the Shisa family is divided are highlighted by dashed line boxes. The name of each subfamily is presented at the right of the box. Posterior probabilities are shown at tree nodes and protein names at the end of each branch. Tree branches are coloured based on phylum, as indicated in the legend. The closely related vertebrate proteins VOPP and WBP1 were used as the outgroup. The function of vertebrate sequences, if known, is indicated. Scale bar denotes number of amino acid substitutions per site. The amino-acid substitution model used was Vt+I+G+F, branch support was obtained after 1000 iterations of ultrafast bootstrapping.



**Supplementary Figure 7. Maximum likelihood phylogenetic tree of Dispanin C subfamily.**

The two subfamilies in which the Dispanin C family is divided are highlighted by dashed line boxes. The name of each subfamily is presented at the right of the corresponding box. Posterior probabilities are shown at tree nodes and protein names at the end of each branch. Tree branches are coloured based on phylum, as indicated in the legend. The vertebrate proteins PRRT2 from the Dispanin B family were used as outgroup. The function of vertebrate sequences, if known, is indicated. Scale bar denotes number of amino acid substitutions per site. The amino-acid substitution model used was Vt+I+G+F, branch support was obtained after 1000 iterations of ultrafast bootstrapping.

## **OTHER SUPPLEMENATRY FILES**

**Supplementary Table 1. Reference Table for Protein names and Gene reference codes.**

**Supplementary File 1. Multiple sequence alignment of CACNG-GSG1 protein family.**

**Supplementary File 2. Multiple sequence alignment of Cornichon protein family.**

**Supplementary File 3. Multiple sequence alignment of Shisa protein family.**

**Supplementary File 4. Multiple sequence alignment of Dispanin C protein family.**