

(a) Shisa-like protein family: FAM159

Table with columns for accession number, species, and protein sequence. Sequences are color-coded with green, red, and magenta highlights. Includes sub-section 'FAM159B'.

(b) Shisa-like protein family: KIAA1644

Table with columns for accession number, species, and protein sequence. Sequences are color-coded with green, red, and magenta highlights.

(c) Shisa and Shisa-like proteins in Branchiostoma floridae

Table with columns for accession number, species, and protein sequence. Sequences are color-coded with green, red, and magenta highlights.

(d) Shisa-like proteins in Saccoglossus kowalevskii

Table with columns for accession number, species, and protein sequence. Sequences are color-coded with green, red, and magenta highlights.

Figure S2. Shisa-like protein families FAM159 and KIAA1644 and lineage-specific expansions of Shisa-like proteins in Branchiostoma floridae and Saccoglossus kowalevskii. Sequences are denoted by NCBI accession numbers (sequences translated from EST sequences have italic and underlined accession numbers). An EST-derived protein with two cysteine-rich domains (labeled domain1 and domain2) has its names colored in green. Cysteine-rich domains are aligned in these proteins, while the segments after cysteine-rich domains (including predicted transmembrane segments and C-terminal low complexity regions showing great sequence divergence) are not aligned for Branchiostoma floridae and Saccoglossus kowalevskii proteins. Conserved cysteines in Shisa-like proteins are shaded in yellow. Branchiostoma-specific cysteines are shaded in magenta. Shisa-specific cysteines are shaded in red. Predicted transmembrane segments (by the program Phobius) are underlined. Cysteines and specifically charged residues (lysines and arginines) near the C-terminal ends of the predicted transmembrane segments are colored red and blue respectively. PY motifs ([LP]PxY) are shaded in cyan. C-terminal PDZ-binding motifs (ITSDVx[LVIL]) are shaded in red. Corrected sequence segments determined by the gene prediction program fgenesh are shaded in gray. Species abbreviations are as follows: Ail.meI, Ailuropoda melanoleuca; Bos.tau, Bos taurus; Cal.jac, Callitrix jacchus; Can.lup, Canis lupus; Mac.mul, Macaca mulatta; Mon.dom, Monodelphis domestica; Mus.mus, Mus musculus; Oik.dio, Okoloplea dioica; Ory.cun, Oryctolagus cuniculus; Pan.tro, Pan troglodytes; Pon.abe, Pongo abelii; Rat.nor, Rattus norvegicus; Sus.scr, Sus scrofa; Tae.gut, Taeniopygia guttata; Tet.nig, Tetraodon nigroviridis; and Xen.tro, Xenopus tropicalis.