

***In silico* identification of tetraspanins in monopisthocotylean (Platyhelminthes: Monogenea) parasites of fish**

Víctor Caña-Bozada^{1*}, F. Neptalí Morales-Serna², José Ángel Huerta-Ocampo^{3,4}, Anaguiven Avalos-Soriano^{1,4}

¹Centro de Investigación en Alimentación y Desarrollo, A.C., Mazatlán 82112, Sinaloa, Mexico

²Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de México, Mazatlán 82040, Sinaloa, Mexico

³Centro de Investigación en Alimentación y Desarrollo, A.C., Hermosillo 83304, Sonora, Mexico

⁴Consejo Nacional de Ciencia y Tecnología (CONACyT), Ciudad de México, Mexico

*** Correspondence:**

V. Caña-Bozada

victorcana1991@hotmail.com

Supplementary fig. S1. Phylogenetic analysis performed using the protein domains of 64 tetraspanins belonging to *Homo sapiens* and *Mus musculus*. The midpoint-rooted phylogenetic tree was constructed using 1000 replicates of the approximate likelihood ratio test (similar to the Shimodaira–Hasegawa test). The JTTDCMut+F+I+G4 model was implemented. According to our analysis, the metazoan tetraspanins can be divided into 17 monophyletic groups. Each group is represented in a different coloured text.

Supplementary fig. S2. Phylogenetic analysis performed using the protein domains of 159 tetraspanins belonging to 24 metazoan species. The midpoint-rooted phylogenetic tree was constructed using 1000 replicates of the approximate likelihood ratio test (similar to the Shimodaira–Hasegawa test). The LG+R5 model was implemented. According to our analysis, the metazoan tetraspanins can be divided into 17 monophyletic groups. Each group is represented in a different coloured text. Proteins that did not fit into one of the 17 monophyletic groups of metazoans are shown in grey text.

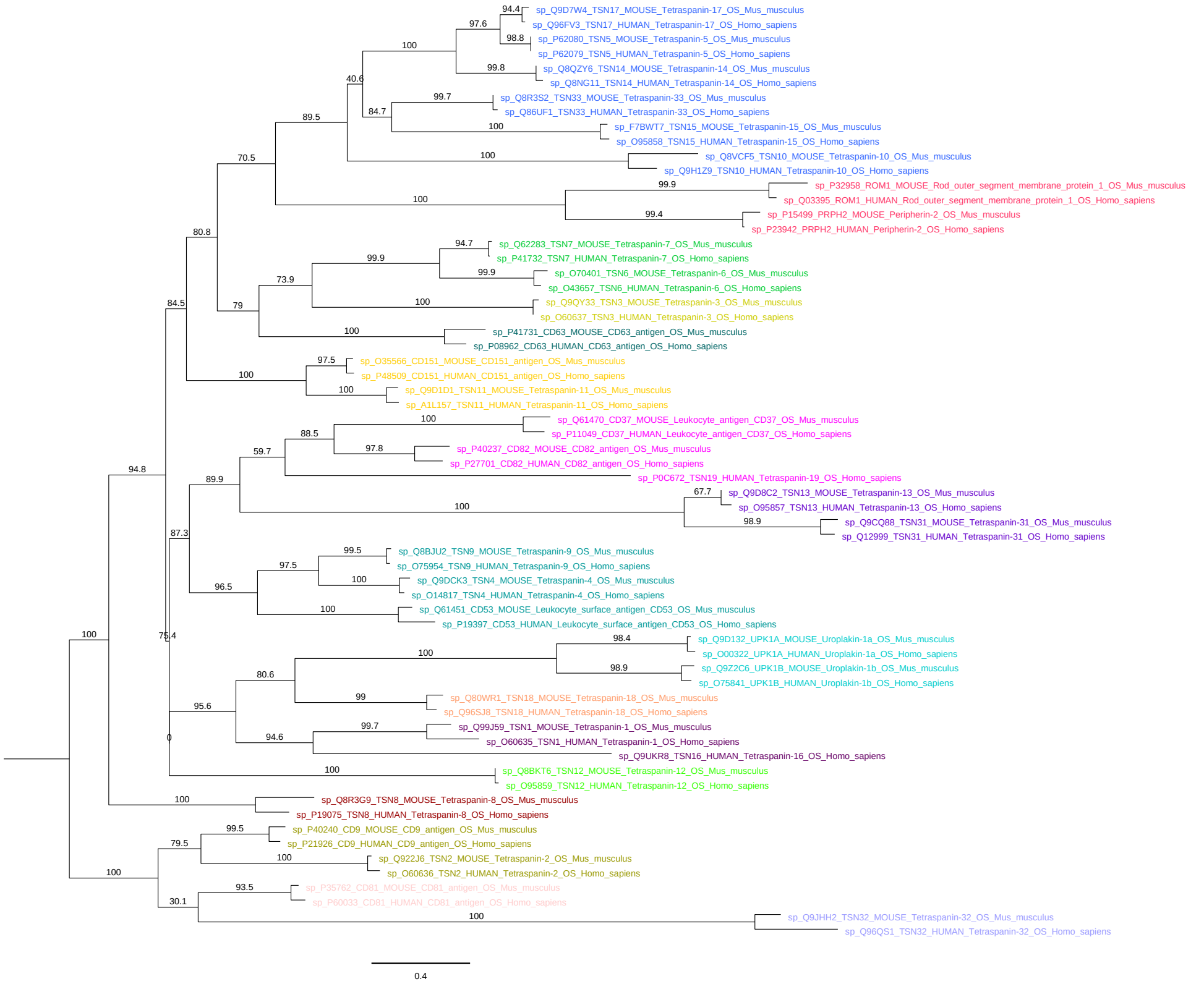
Supplementary fig. S3. Phylogenetic analysis performed using the protein domains of 184 tetraspanins belonging to 26 eukaryota. The midpoint-rooted phylogenetic tree was constructed using 1000 replicates of the approximate likelihood ratio test (similar to the Shimodaira–Hasegawa test). The JTT+F+I+G4 model was implemented. Proteins that did not group into one of the 17 monophyletic groups of metazoans are shown in black (plant and amoeba proteins) and grey text.

Supplementary fig. S4. Phylogenetic analysis performed using the complete sequence of 159 tetraspanins belonging to 24 metazoan species. The midpoint-rooted phylogenetic tree was constructed using 1000 replicates of the approximate likelihood ratio test (similar to the Shimodaira–Hasegawa test). The JTTDCMut+F+R4 model was implemented. Proteins that did not fit into one of the 17 monophyletic groups of metazoans are shown in black and grey text.

Supplementary fig. S5. Numbers of transmembrane domains present in the platyhelminth tetraspanins.

Supplementary fig. S6. Phylogenetic analysis performed using the protein domains of 114 monopisthocotylean tetraspanins and 64 tetraspanins belonging to *Homo sapiens* and *Mus musculus*. The midpoint-rooted phylogenetic tree was constructed using 1000 replicates of the approximate likelihood ratio test (similar to the Shimodaira–Hasegawa test). The WAG+F+R4 model was implemented. The monopisthocotylean proteins are shown in black text, and the metazoan proteins are shown in different colours according to each monophyletic group. Blue-coloured branches indicate phylogenetic closeness of monogenean tetraspanins to metazoan tetraspanins. Because the TRINITY_DN12906_c0_g2_i1.p4 protein did not show agreement in further phylogenetic analyses, this protein was not considered close to metazoan tetraspanins.

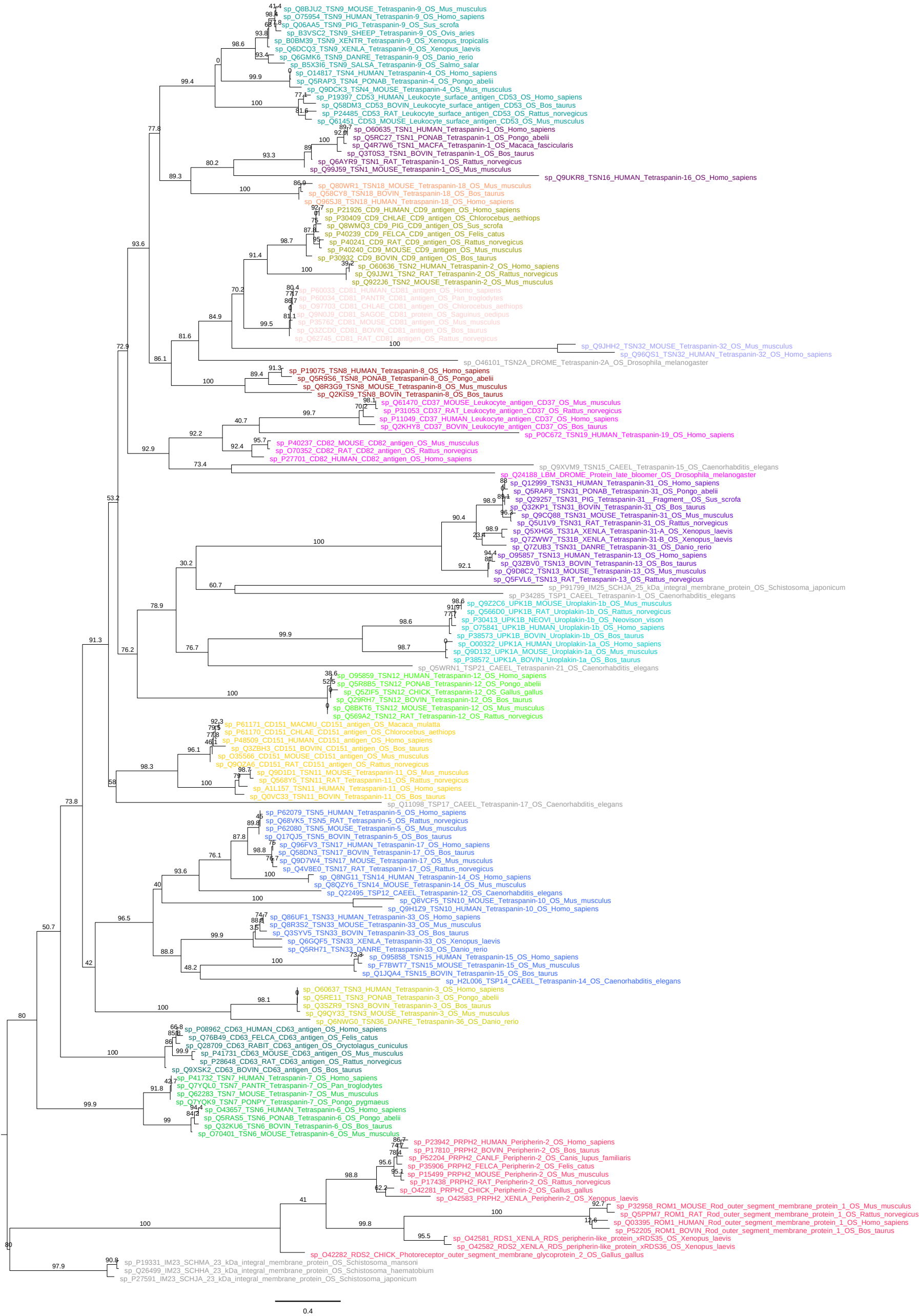
Supplementary fig. S7. Phylogenetic analysis performed using the protein domains of 114 monopisthocotylean tetraspanins and 354 belonging to other platyhelminths. The midpoint-rooted phylogenetic tree was constructed using 1000 replicates of the approximate likelihood ratio test (similar to the Shimodaira–Hasegawa test). The WAG+F+R7 model was implemented. The coloured boxes indicate the \log_{10} -transformed e-values obtained from the alignment of the monogenean sequences against sequences of different taxa using the NCBI database (Tre, Trematoda; Ces, Cestoda; Rha, Rhabditophora; Lop, Lophotrochozoa; Spi, Spiralia; Pro, Protostomia; Bil, Bilateria). The sequences of the proteins Sm23, Sm-TSP-1 and Sm-TSP-1 from *S. mansoni* were used as references.



Supplementary fig. S1

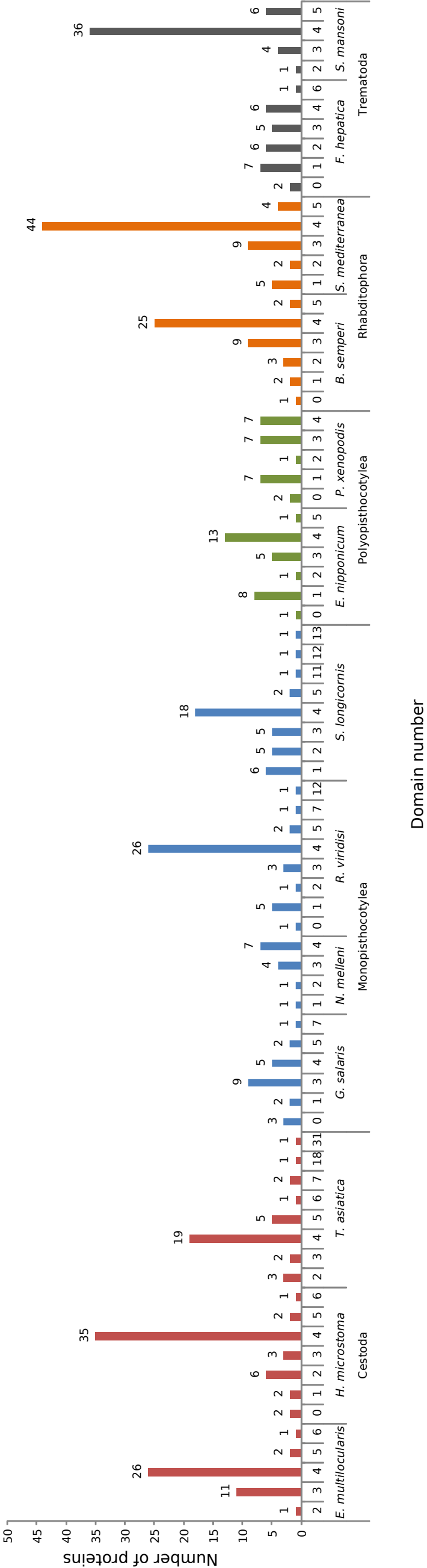






Supplementary fig. S4

Supplementary fig. S5





Supplementary fig. S7

