

Exploring Evolutionary Relationship within Neodermata using Putative Orthologous Groups of Proteins, with Emphasis on Peptidases

Víctor Caña-Bozada¹, Mark W. Robinson², David I. Hernández-Mena³, Francisco N. Morales-Serna⁴

¹Centro de Investigación en Alimentación y Desarrollo, A.C., Mazatlán 82112, Sinaloa, Mexico; ²School of Biological Sciences, Queen's University Belfast, 19 Chlorine Gardens, Belfast, Northern Ireland, UK; ³Centro de Investigación y de Estudios Avanzados, Instituto Politécnico Nacional, Unidad Mérida, Mérida 97310, Yucatán, Mexico; and ⁴Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de México, Mazatlán 82040, Sinaloa, Mexico

Corresponding author: Francisco N. Morales-Serna,
neptali@ola.icmyl.unam.mx

Supplementary Figures S1-S11

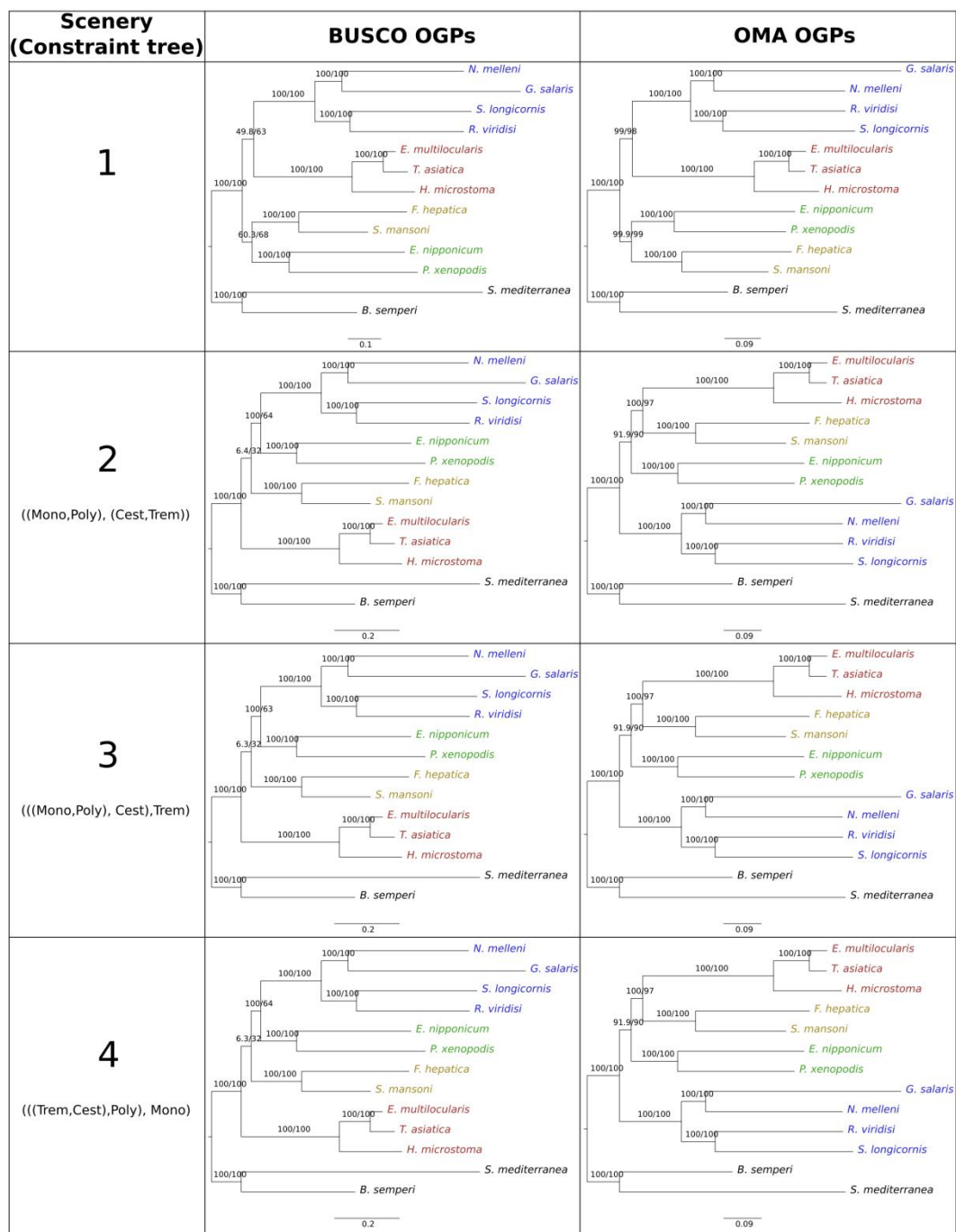
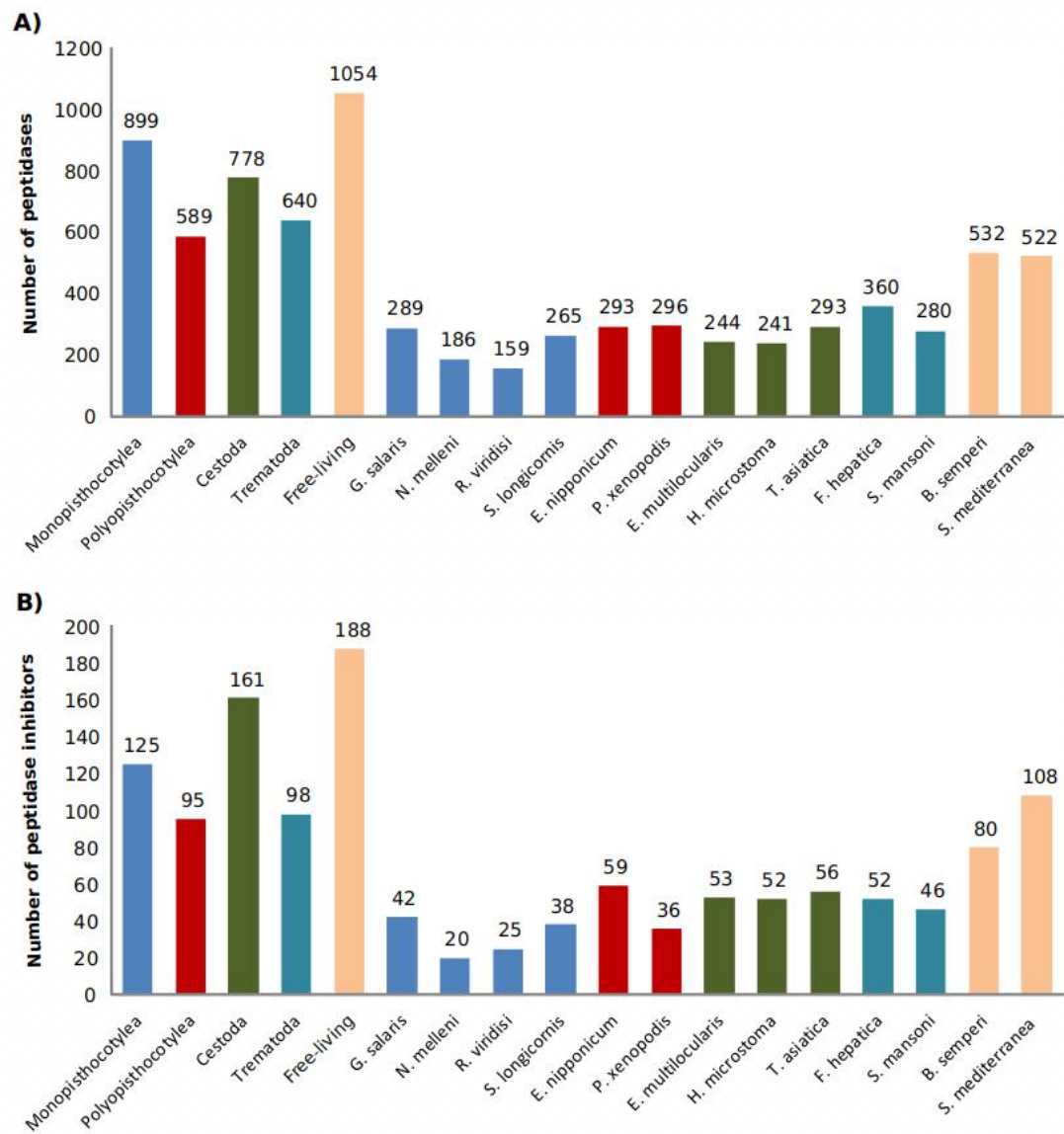
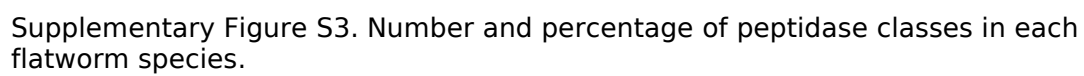
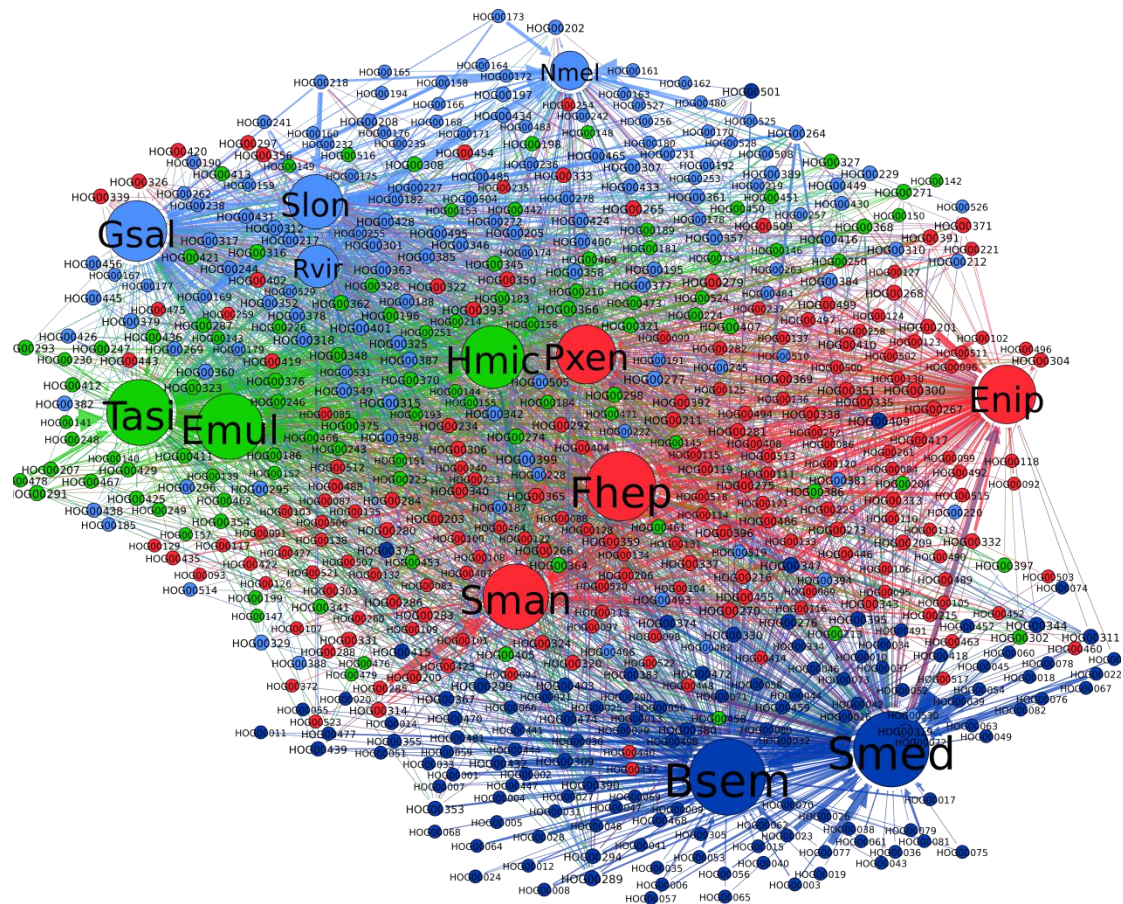


Figure S1: Constrained trees constructed of 13 platyhelminth species, based on single-copy BUSCO OGPs and simply OMA OGPs. Alternative scenarios obtained from the literature were used as guide (2, 3, 4) and the LG+F+R4 model. Values above represent SH-aLRT support (%). The free-living platyhelminths *Bothrioplana semperi*, belonging to the order Bothrioplanida, and *Schmidtea mediterranea*, belonging to the order Tricladida, were used as outgroups.

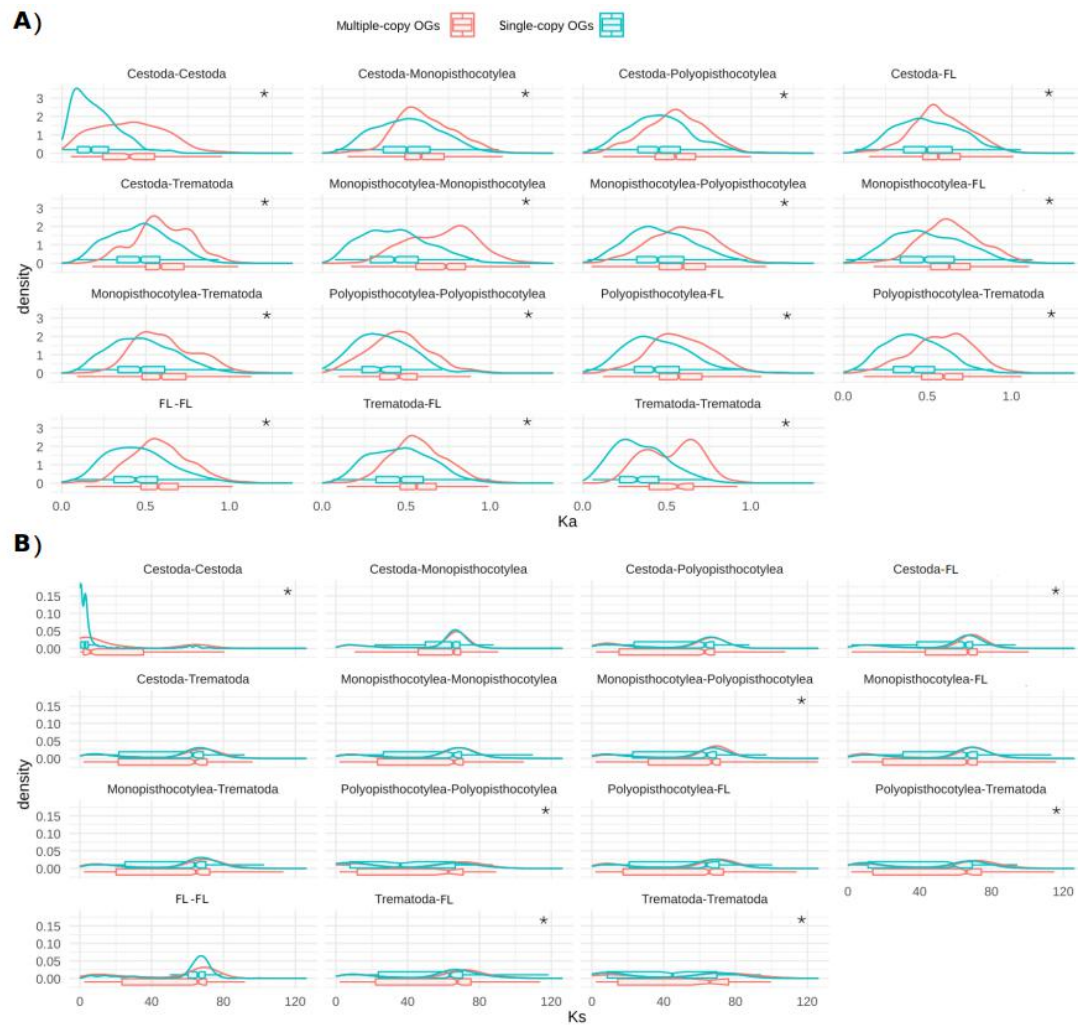


Supplementary Figure S2. Number of peptidases (A) and peptidase inhibitors (B) identified in the selected platyhelminth species.





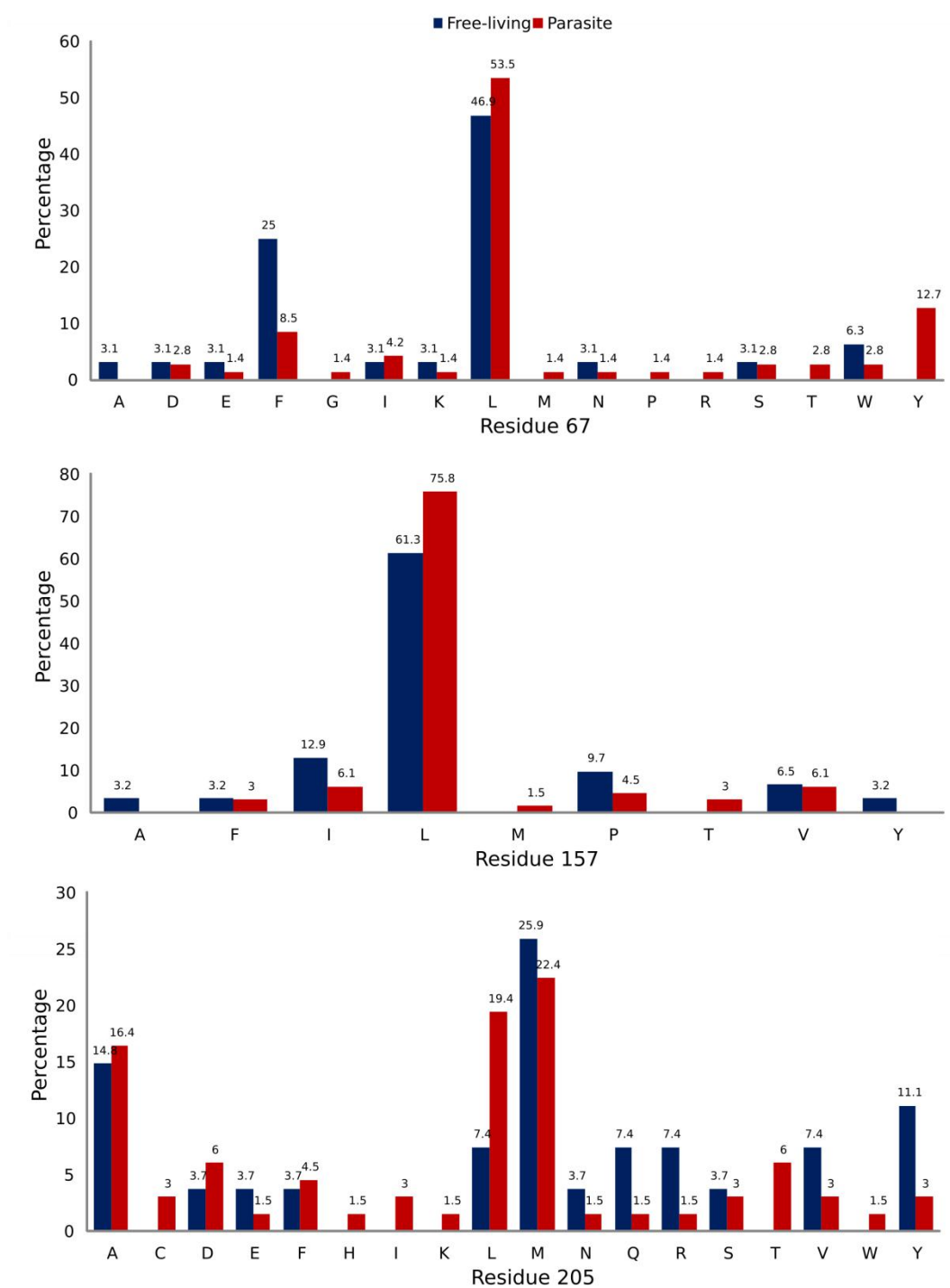
Supplementary Figure S4. Formation of clusters of peptidases/inhibitors OGs and platyhelminth species in network analysis. Monogenea Monopisthocotylea: NmeI, *Neobenedenia melleni*; Gsal, *Gyrodactylus salaris*; Slon, *Scutogyrus longicornis*; Rvir, *Rhabdosynochus viridisi*. Cestoda: Tasi, *Taenia asiatica*; Emul, *Echinococcus multilocularis*; Hmic, *Hymenolepis microstoma*. Monogenea Polyopisthocotylea: Enip, *Eudiplozoon nipponicum*; Pxen, *Protopolystoma xenopodis*. Trematoda: Fhe, *Fasciola hepatica*; Sman, *Schistosoma mansoni*. Free-living platyhelminths: Bsem, *Bothrioplana semperi*; Smed, *Schmidtea mediterranea*.



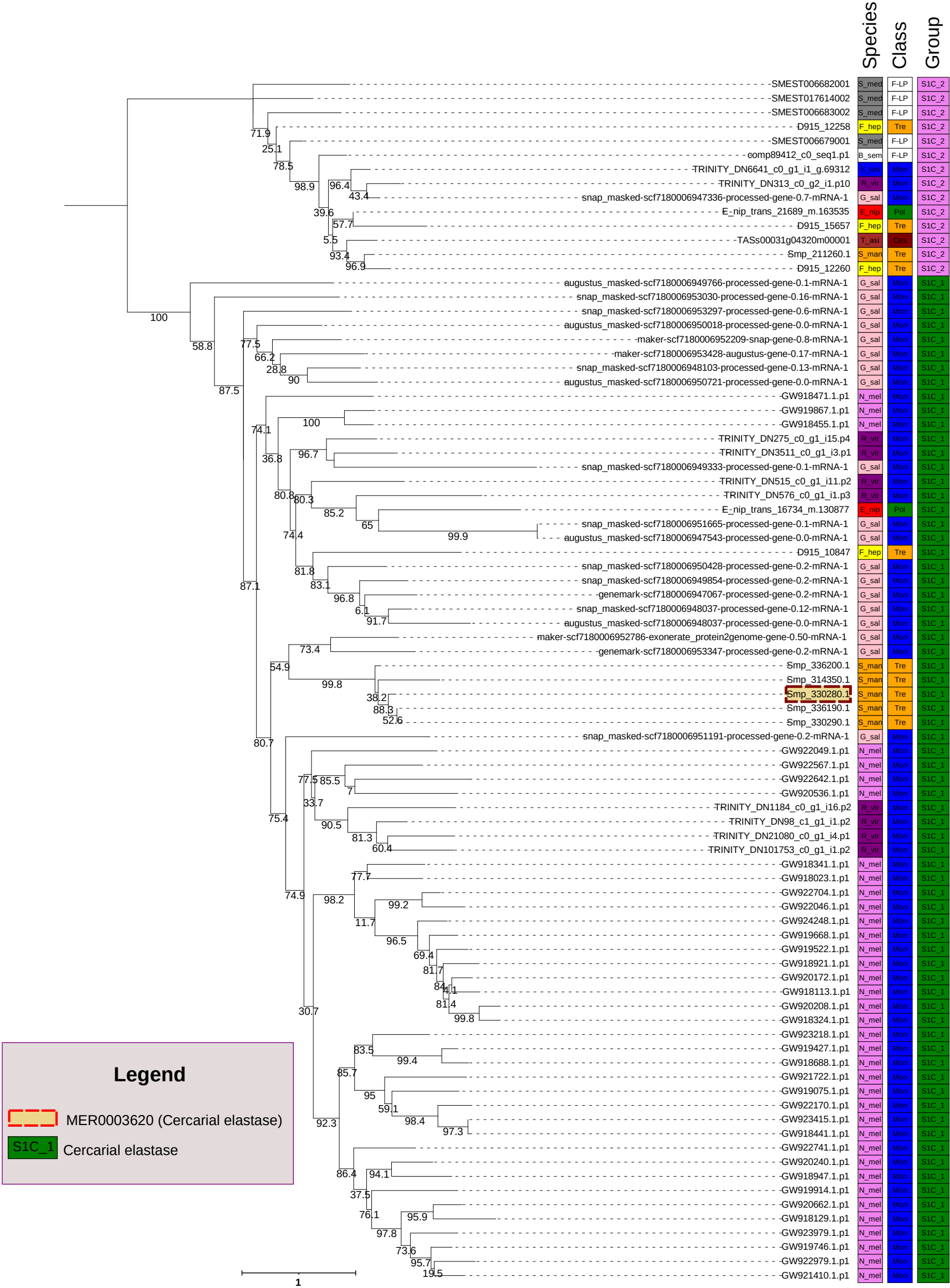
Supplementary Figure S5. Distribution of Ka and Ks in peptidases and inhibitors of platyhelminths. A) Ka was higher in multiple-copy OGs than in single-copy OGs. B) Ks was similar between single-copy OGs and multiple-copy OGs. FL, free-living platyhelminths. * indicate $p < 0.05$.



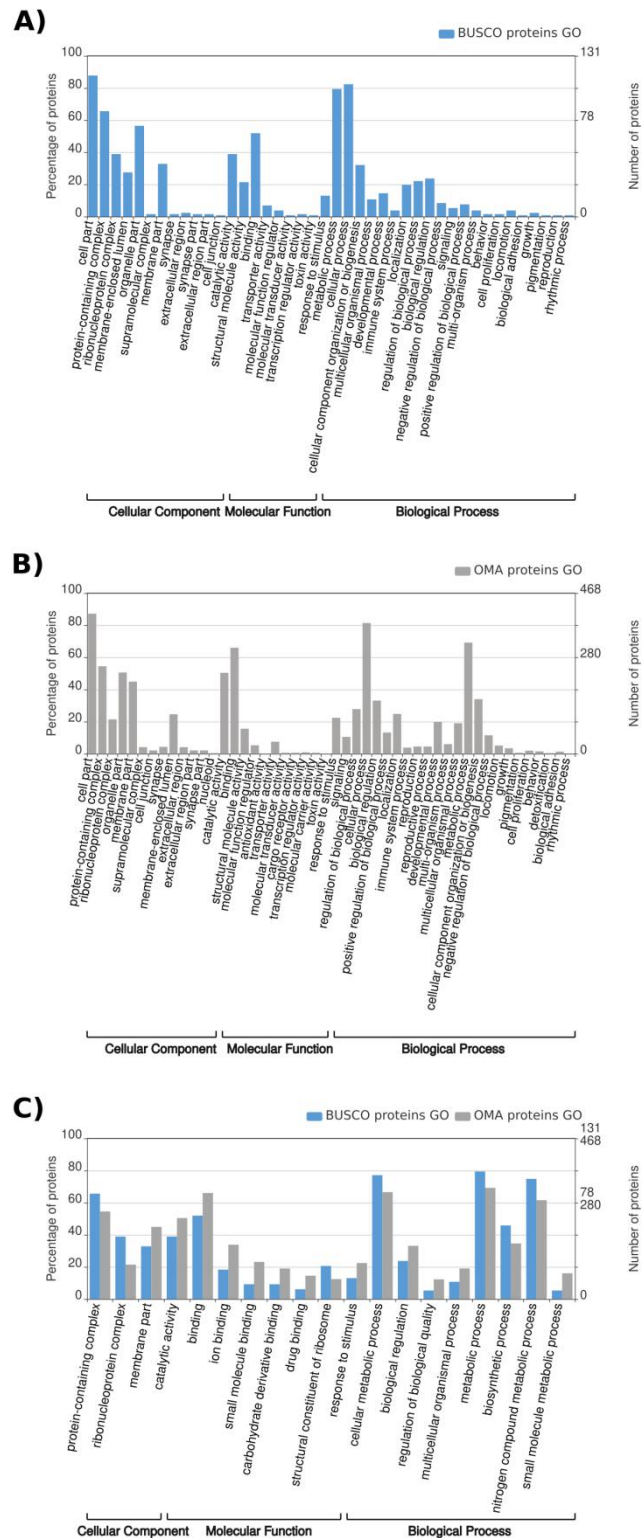
Age Group	Number of People
10-19	20
20-29	10
30-39	40
40-49	60
50-59	20
60-69	20
70-79	30
80-89	20
90-99	40
100-109	10



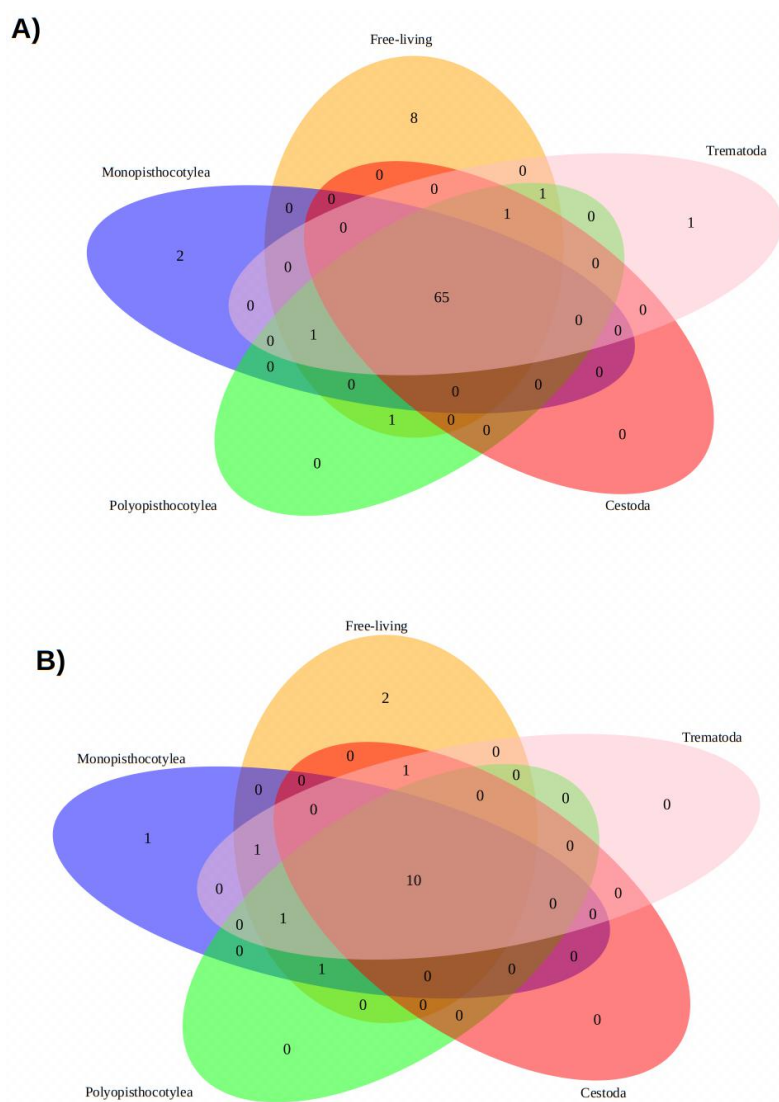
Supplementary Figure S8. Frequency of the amino acids comprising the S2 subsite (positions 67, 167 and 205) of cathepsins L belonging to the C01A subfamily in platyhelminths.



Supplementary Figure S9. Phylogenetic analysis of the S01C peptidase subfamily. The midpoint-rooted phylogenetic tree was constructed using 1000 replicates of the approximate likelihood ratio test (similar to the Shimodaira–Hasegawa test). The WAG+R5 model was implemented. The protein MER0003620 (Cercarial elastase) of the MEROPS database was used as reference sequences. Abbreviations: Mon, Monopisthocotylean; Pol, Polyopisthocotylea; Tre, Trematoda; Ces, Cestoda; CF-LP, Free-living platyhelminths. N_mel, *Neobenedenia melleni*; G_sal, *Gyrodactylus salaris*; S_lon, *Scutogyrus longicornis*; R_vir, *Rhabdosynochus viridis*; t_asi, *taenia asiatica*; E_mul, *Echinococcus multilocularis*; H_mic, *Hymenolepis microstoma*; E_nip, *Eudiplozoon nipponicum*; P_xen, *Protopolystoma xenopodis*; F_hep, *Fasciola hepatica*; S_man, *Schistosoma mansoni*; B_sem, *Bothrioplana semperi*; S_med, *Schmidtea mediterranea*. CTSL, cathepsins L; CTSC, cathepsins C.



Supplementary Figure S10. Gene Ontology annotation of proteins used in the phylogenomic analysis. A) GO of proteins obtained from BUSCO. B) GO of proteins obtained from OMA. C) GO term enrichment between BUSCO and OMA proteins.



Supplementary Figure S11. Number of shared peptidase families. A) Peptidase families. B) Peptidase inhibitor families in taxonomic groups within the Neodermata.