

# The genome of fish tapeworm *Nippotaenia percotti* as a potential bookmark for gene loci that facilitate anthropogenic infection

Tyler K. Chafin, Bhavya Papudeshi, Sheri Sanders, Carrie Ganote, Andrey N. Reshetnikov, Sergey G. Sokolov, Thomas G. Doak, Jeff F. Pummill, Marlis R. Douglas, Michael E. Douglas

## Introduction

- Tapeworms are among the first known parasites of humans, but a safe and efficient cure for larval tapeworm infection in humans has yet to be found.
- Cestoda is a class of parasitic tapeworms that are endoparasites infecting all vertebrates.
- More than 50 % of the diversity described for Cestoda resides within the group Cyclophyllidea, a common, chronic source of anthropogenic infection.
- Phylogenetic relationships within the Cestoda remain unstable, with very little resolution in certain parts of the tree, due to lack of genomic information for many families.
- The tapeworm *Nippotaenia percotti* belongs to Nippotaeniidea (putative sister group to Cyclophyllidea), with only 6 identified species.
- **The *Nippotaenia percotti* genome was studied for comparative purposes, to explore evolutionary change in functional gene loci of host-parasite interactions.**

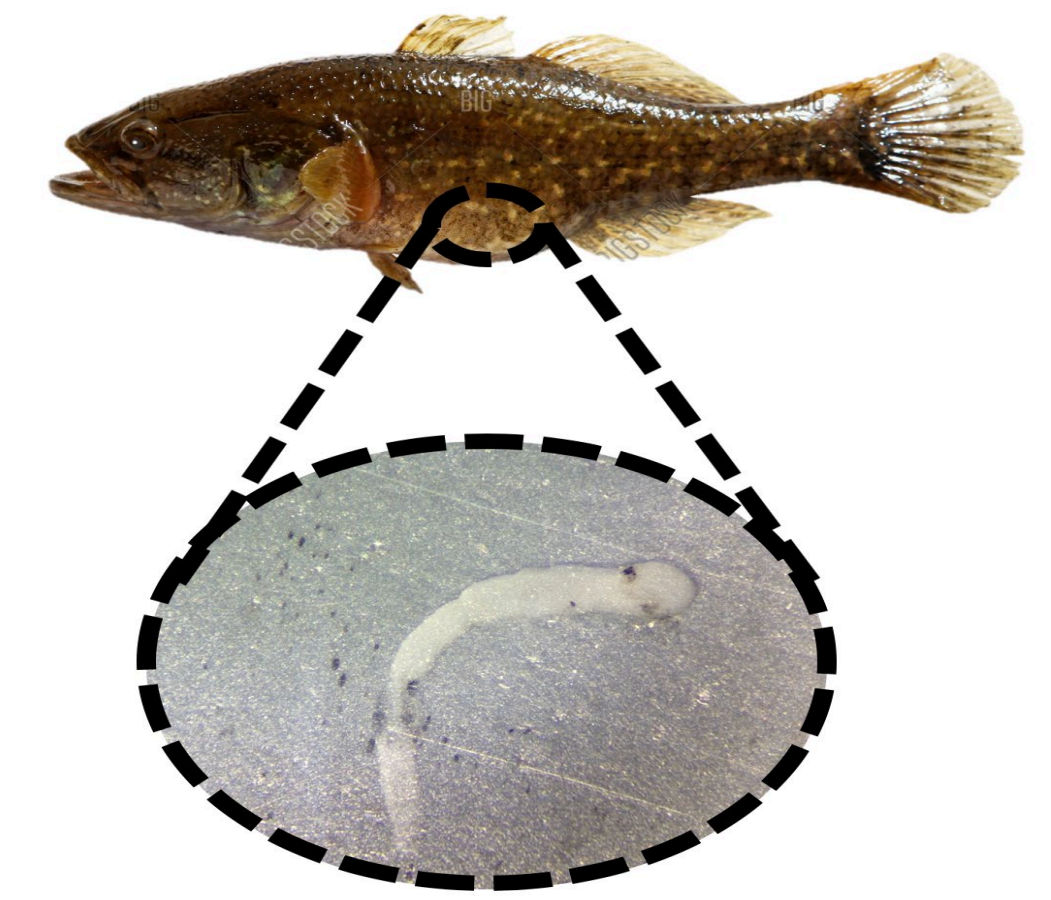


Figure 1- The tapeworm *Nippotaenia percotti* is host-specific to a fish in the Amur River

## Methods

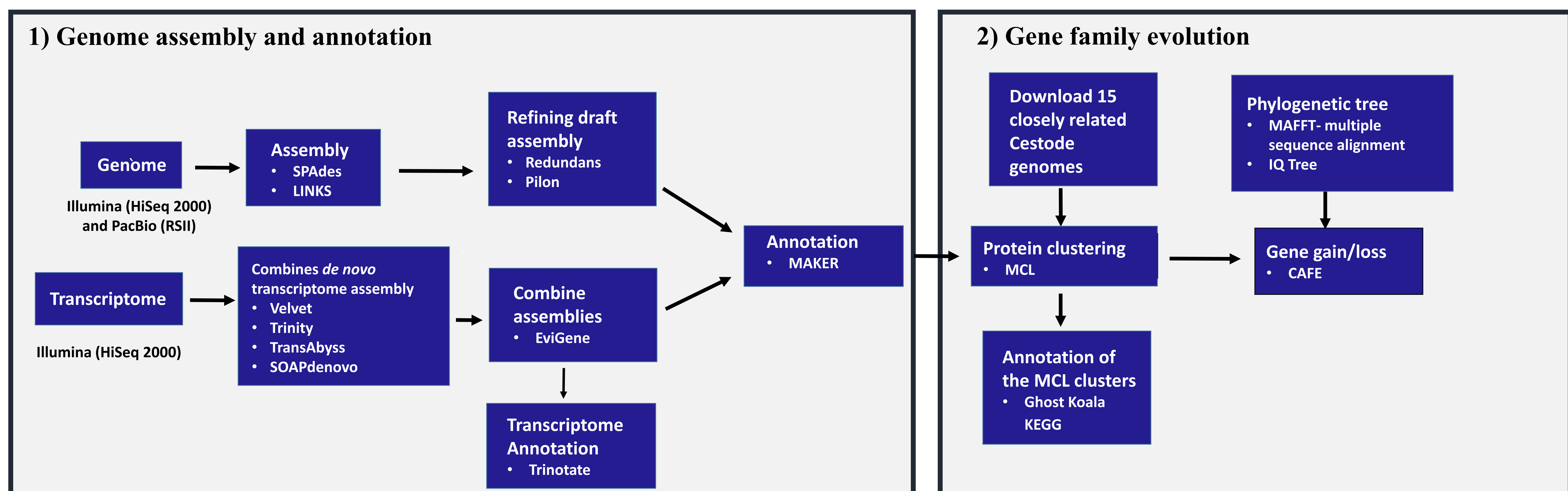


Figure 2 - Workflow applied: 1) assemble and annotate the *Nippotaenia percotti* reads; and 2) study the gene gain/loss among species within Cestoda.

## Results

### 1) *Nippotaenia percotti* genome

- The *Nippotaenia percotti* genome was assembled to generate 3,410 contigs at an  $N_{50}$  of 209,561bp, and 54.2 % genome completeness (Metazoa; *Schistosoma*).
- The transcriptome was assembled to generate 28,226 transcripts at an  $N_{50}$  of 2,290bp, and 82.7 % transcriptome completeness (Metazoa; *Schistosoma*).
- MAKER identified 30,671 genes with a mean length of 1,824 bp.

### 2) Cestoda gene family evolution

- 21,861 orthologous gene clusters were identified using MCL. The gene clusters were annotated against KEGG database using Ghost Koala.
- The rate of change as the probability that a gene is gained ( $\lambda$ ) or lost ( $\mu$ ) was estimated to be 0.0009, and 0.0024 respectively by CAFE.
- This suggests that Cestoda genomes are undergoing more gene loss ( $\mu$ ) than gene gain ( $\lambda$ ), which could be due to their parasitic nature.

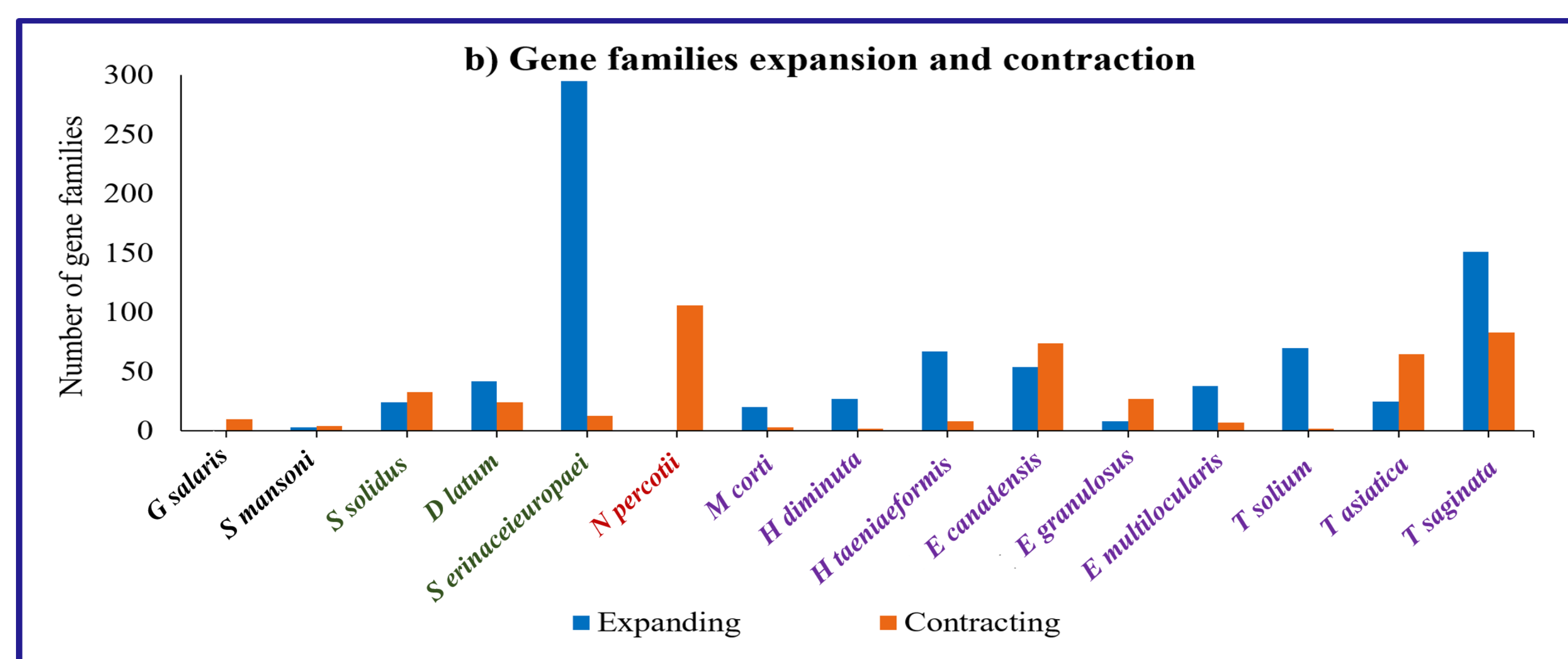
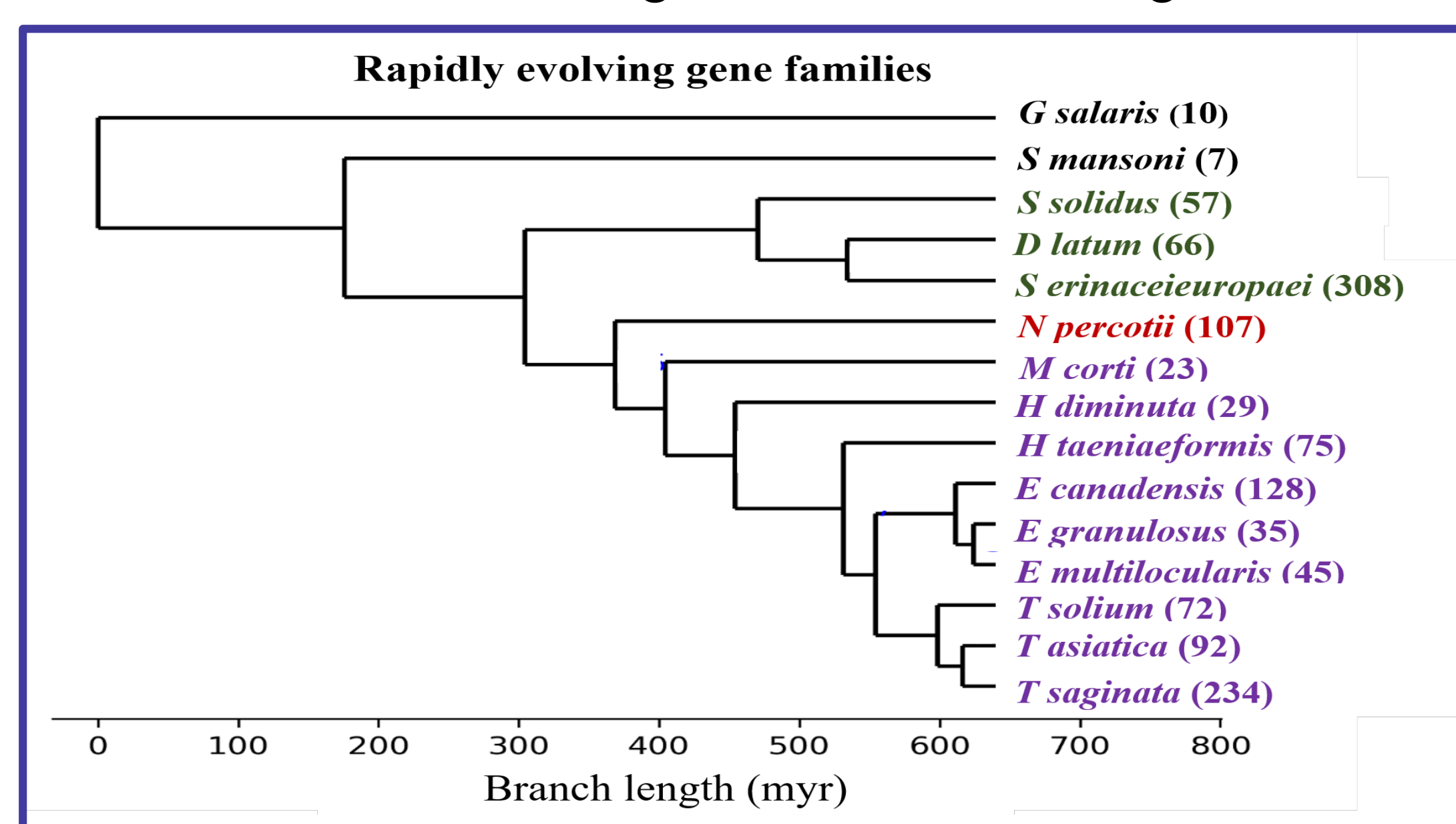


Figure 3 - a) Phylogenetic tree of the Cestoda constructed from representatives of Cyclophyllidea (purple), Nippotaeniidae (red), and Pseudophyllidae (green) and outgroups *G salaris*, and *S mansoni* (black), with the number of rapidly evolving gene families indicated for each species; b) The number of significantly evolving gene families from CAFE, expanding (blue), and contracting (orange) for each species included in this study.

## Conclusions and future directions

- We report the first sequenced genome for the tapeworm *Nippotaenia percotti*, belonging to the group Nippotaeniidea. Further work is currently being done to improve the genome assembly.
- A phylogenetic tree constructed with core genes identifies Nippotaeniidea as a sister group to the clade Cyclophyllidea, which can now be studied as a potential bookmark to identify gene loci that facilitate anthropogenic infection.
- Cestode genomes are undergoing more gene loss ( $\mu$ ) than gene gain ( $\lambda$ ), which is expected, due to their host-parasite interactions.
- Our next step will examine gene family annotations to study the evolution of parasitism and identify potential gene loci that could help with tapeworm treatment/control.