

Bioinformatic pipeline for the extraction of molecular markers from Nanopore transcriptomic data in *Pomacea canaliculata*

Filippo Bertolasi¹, Davide Malagoli¹, Nicola Franchi¹

¹Department of Life sciences, University of Modena and Reggio Emilia

Introduction

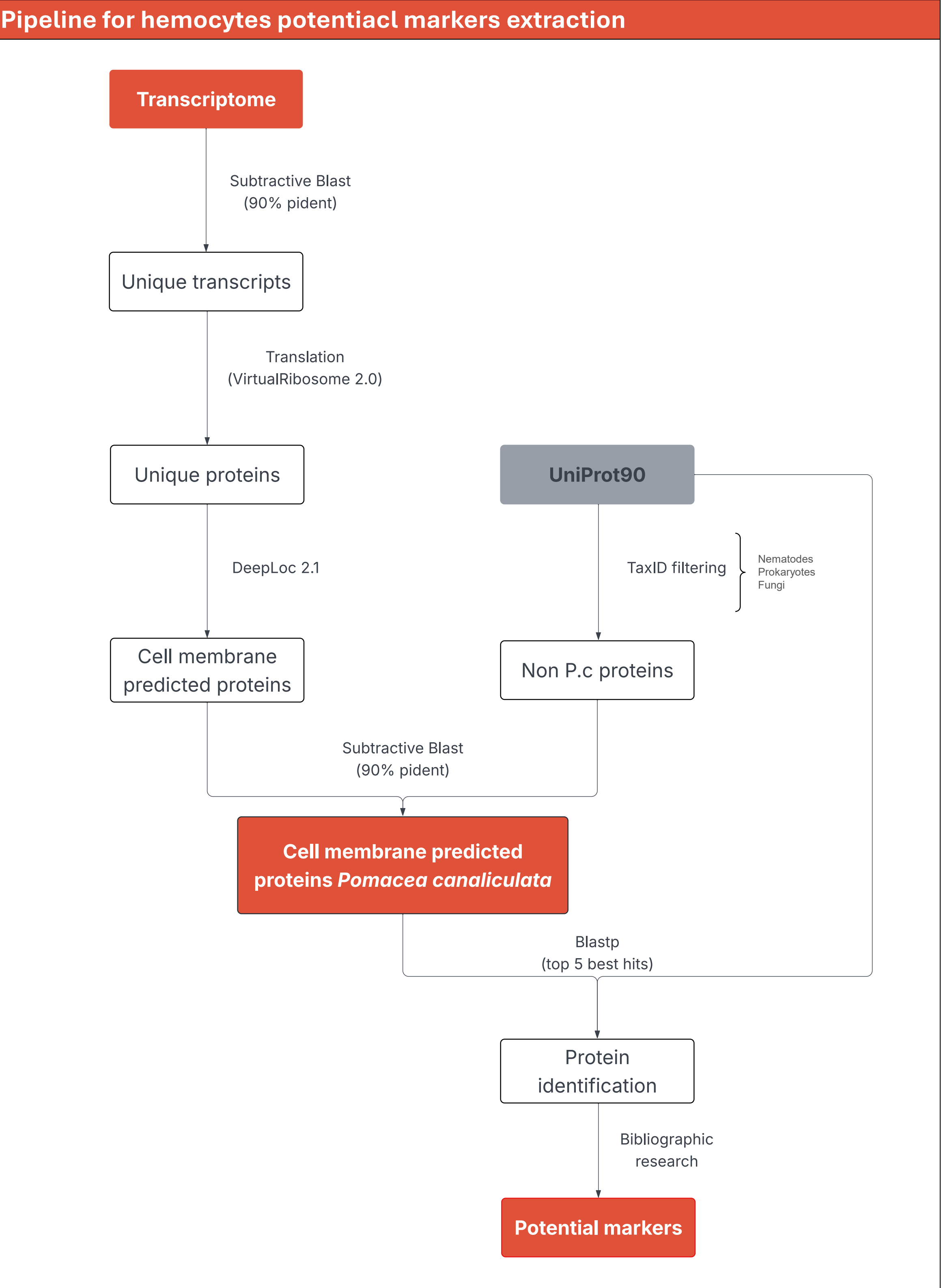
Pomacea canaliculata is a freshwater gastropod mollusk of interest from multiple perspectives, especially due to its high resistance to stress conditions and remarkable regenerative abilities in the adult organism (Bever et al., 1988. Accorsi et al., 2017). Its high fecundity and adaptability to a wide variety of environments make *Pc* one of the 100 most invasive species according to the ISSG (Invasive Species Specialist Group) (View 100 of the world's worst invasive alien species, 2000). Molecular data on these organisms is very limited. In this study, we used 4th generation Oxford Nanopore sequencing technology to generate tissue-specific transcriptomes and perform an initial analysis to extract potential markers useful for in vitro analyses.

Objective

The aim of this study is to leverage bioinformatic resources available in the literature to develop an analysis pipeline for the identification of molecular markers in transcriptomes generated using Nanopore sequencing technology. Samples were selected based on their functional relevance: ganglion (central nervous system markers), hemocytes (immune-related markers), and posterior kidney (cell proliferation markers). Additionally, an *in silico* analysis of transcript expression was conducted to obtain preliminary insights into expression patterns, which will serve as a foundation for subsequent in vitro validation.

Method

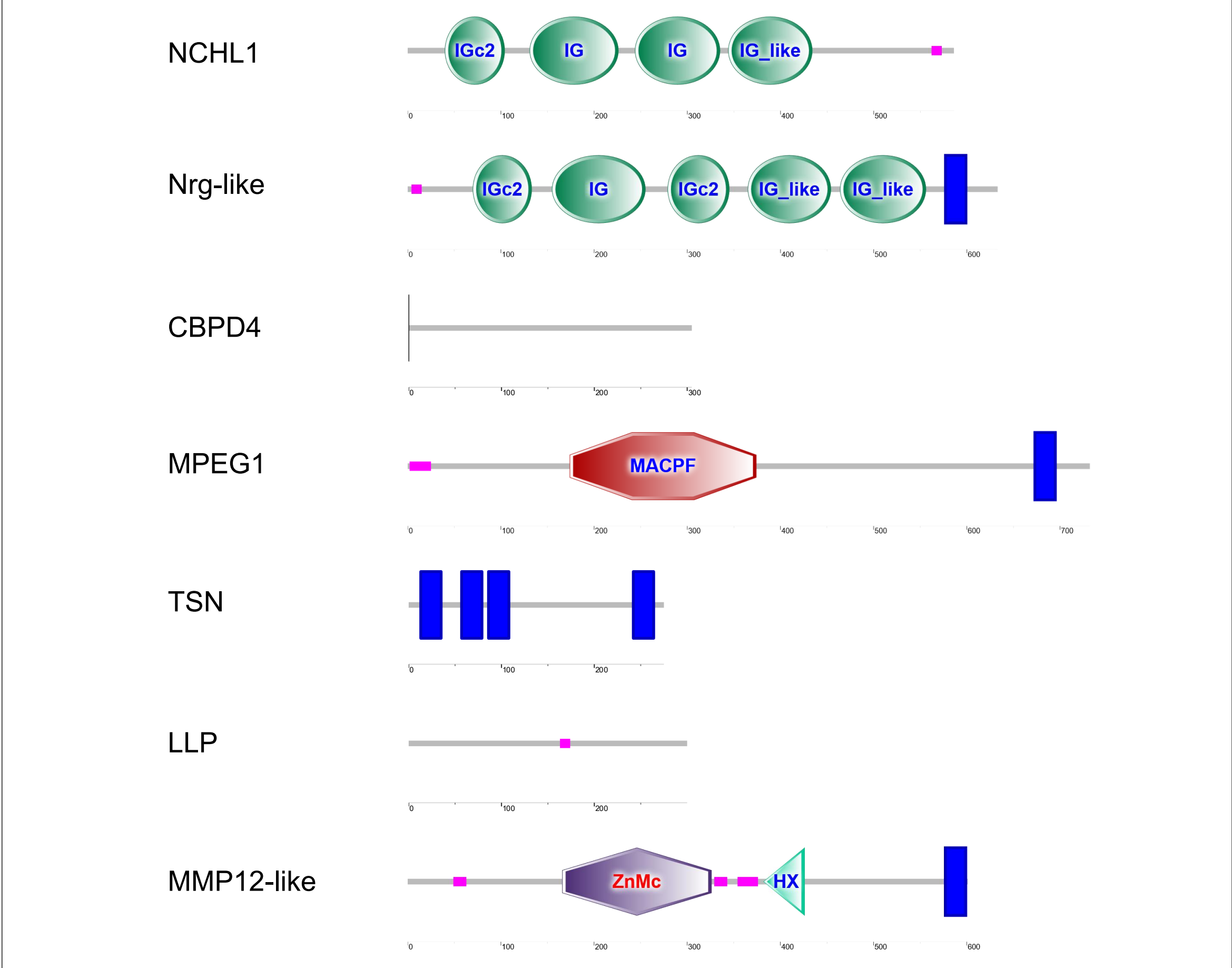
Transcripts were generated using the Nanopore MinION Mk1C platform, leveraging its capacity to produce long reads relative to short-read sequencing technologies. To compensate for the higher intrinsic error rate associated with Nanopore reads, we performed a polishing step using raw Illumina data available from the public SRA database. The processed data was then analyzed using a custom pipeline I developed, consisting of multiple filtering stages to progressively refine the dataset. The final output was a pool of sequences fulfilling specific criteria: uniqueness to the hemocyte transcriptome, predicted localization at the cell membrane, and confirmed as belonging to *Pomacea canaliculata* by subtracting potential contaminants.



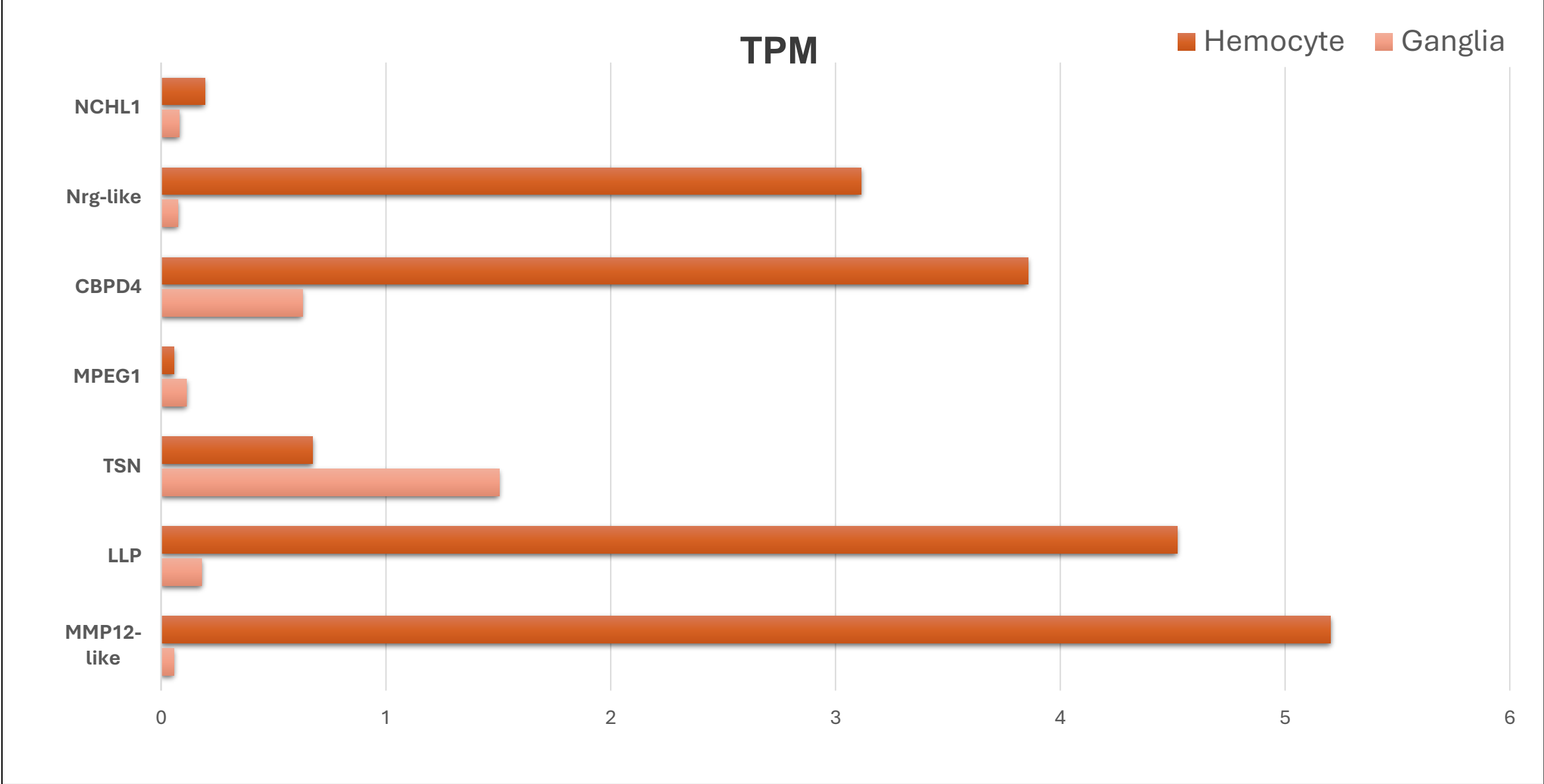
Results

UniProt90 best hit	Brief description
Neural cell adhesion molecule L1-like protein (NCHL1)	Expressed on NK cells
Neuroglian-like (Nrg-like)	Promotes interaction neuron-neuron and neuron-glia cells. Essential for the formation of axon bundles
Chitin-binding type-4 domain-containing protein (CBPD4)	Related to pathogen recognition. Binds fungi and bacteria.
Macrophage-expressed gene 1 protein-like (MPEG1)	Macrophage expressed antimicrobial pore-forming protein.
Tetraspanin (TSN)	Ancient conserved protein family, some of them are associated with immune response. They are sometimes upregulated after PAMPs stimulus.
Leishmanolysin-like peptidase isoform X1 (LLP)	Complement inhibition activity without presenting CCP domains. Mediate parasite invasion of macrophages.
Macrophage metalloelastase-like (MMP12-like)	Key enzyme for macrophage-mediated tissue remodeling and migration.

SMART domain architecture analysis of the potential markers



Salmon expression analysis of Hemocytes and Ganglia



Conclusion

The study revealed seven potential immune markers specific to hemocytes, which will be further investigated in vitro. Additionally, a comprehensive expression profile of the full transcriptome was generated, now available for quantitative consultation of any transcript. The research produced the most advanced, tissue-specific transcriptome to date for this organism. The presence of all seven sequences in both hemocyte and ganglion transcriptomes suggests immune cell infiltration into the central nervous system. These findings contribute new insights into neuroimmune interactions in *Pomacea canaliculata*.