

MIP Superfamily Diversity: Identification of Determinant Structural Motifs

Bussolini Lizundia R^{1,2}, Ramoa U^{1,2}, Canessa Fortuna A^{1,2}, Fox R³, Vitali V^{1,2}, Alleva K^{1,2}

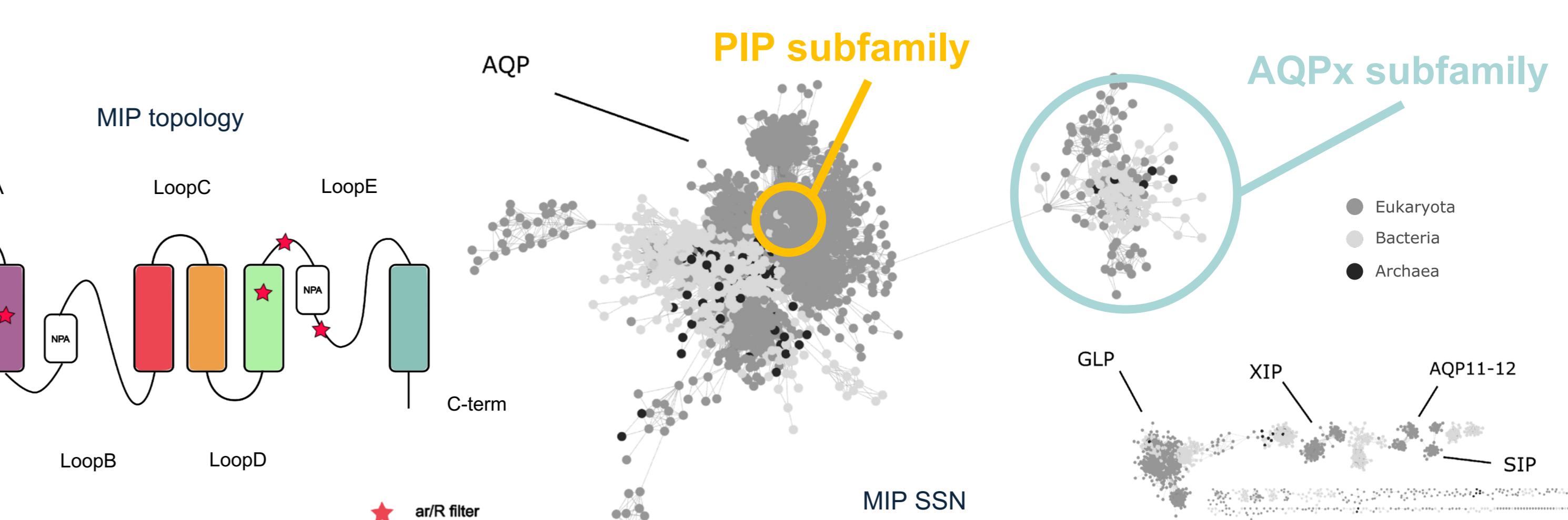
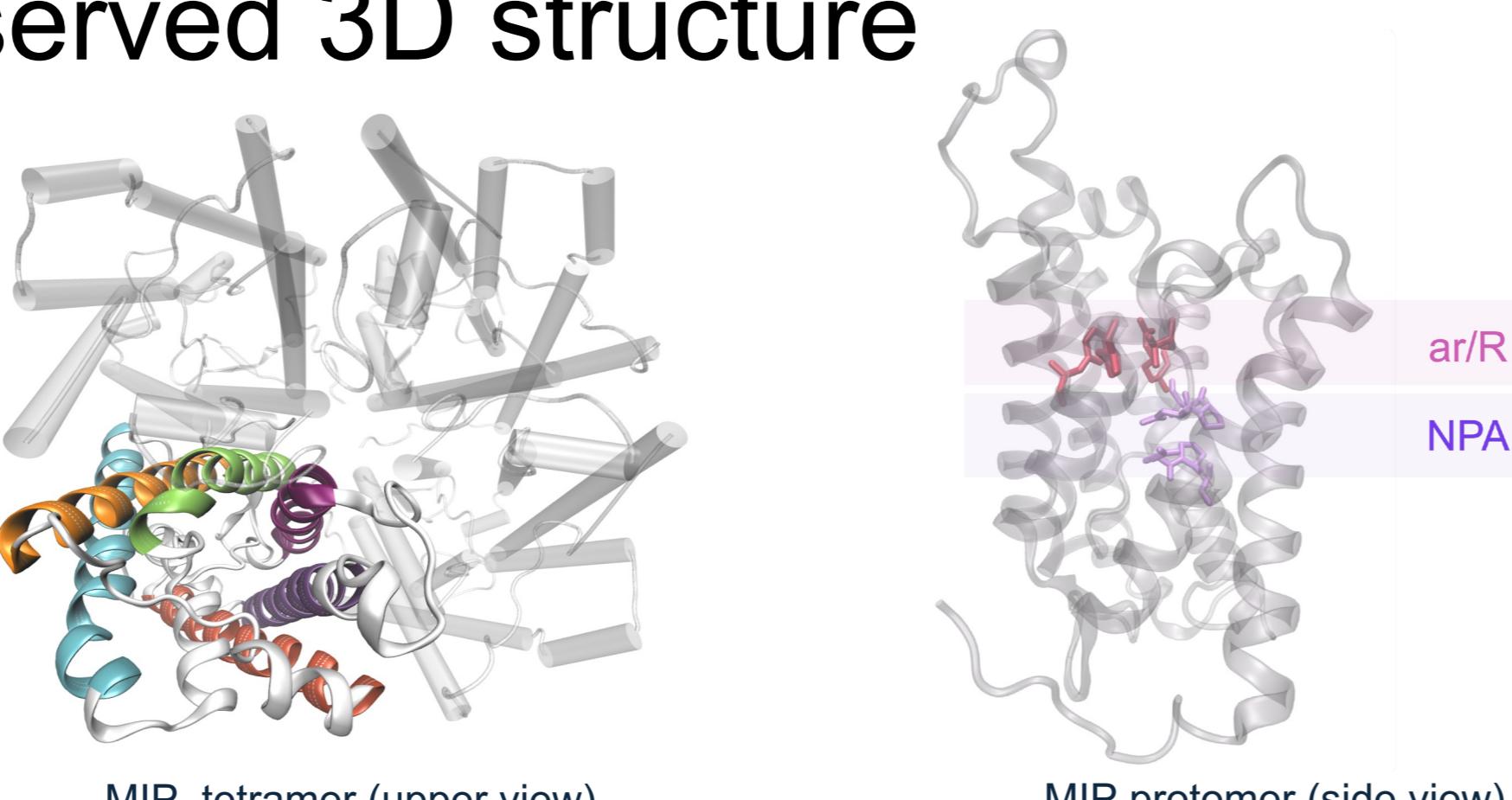
¹Universidad de Buenos Aires, Facultad de Farmacia y Bioquímica, Departamento de Fisicomatemática, Cátedra de Física, Buenos Aires, Argentina. ²Universidad de Buenos Aires-CONICET, Instituto de Química y Fisicoquímica Biológicas "Prof. Alejandro C. Paladini", Buenos Aires, Argentina. ³Louvain Institute of Biomolecular Science and Technology, UCLouvain, Louvain-la-Neuve, Belgium.

Funding: PICT 2021 (ANPCyT), UBACyT 2023 (UBA), PIP 2021 (CONICET)

rbusolini@docente.ffyb.uba.ar

MIP channels share a conserved 3D structure

Figure 1. Major Intrinsic Protein (MIP) channels assemble as tetramers within lipid membranes, forming a central pore. Each protomer features its own pore where there are two key selectivity filters: the Aromatic-Arginine (ar/R, pink stars) filter, serving as a steric and cations barrier, and two conserved Asn-Pro-Ala (NPA) motifs, which prevent the passage of protons. The Sequence Similarity Network (SSN) highlights the extensive distribution of these channels across domains of life. Dataset: 34,263 seq.; E-value: 10^{-5} ; Alignment Score threshold: 60.



PIP subfamily

Despite their high sequence identity and extensive study, the **distinct trafficking** between the PIP1 and PIP2 subfamily paralogues remain poorly understood.

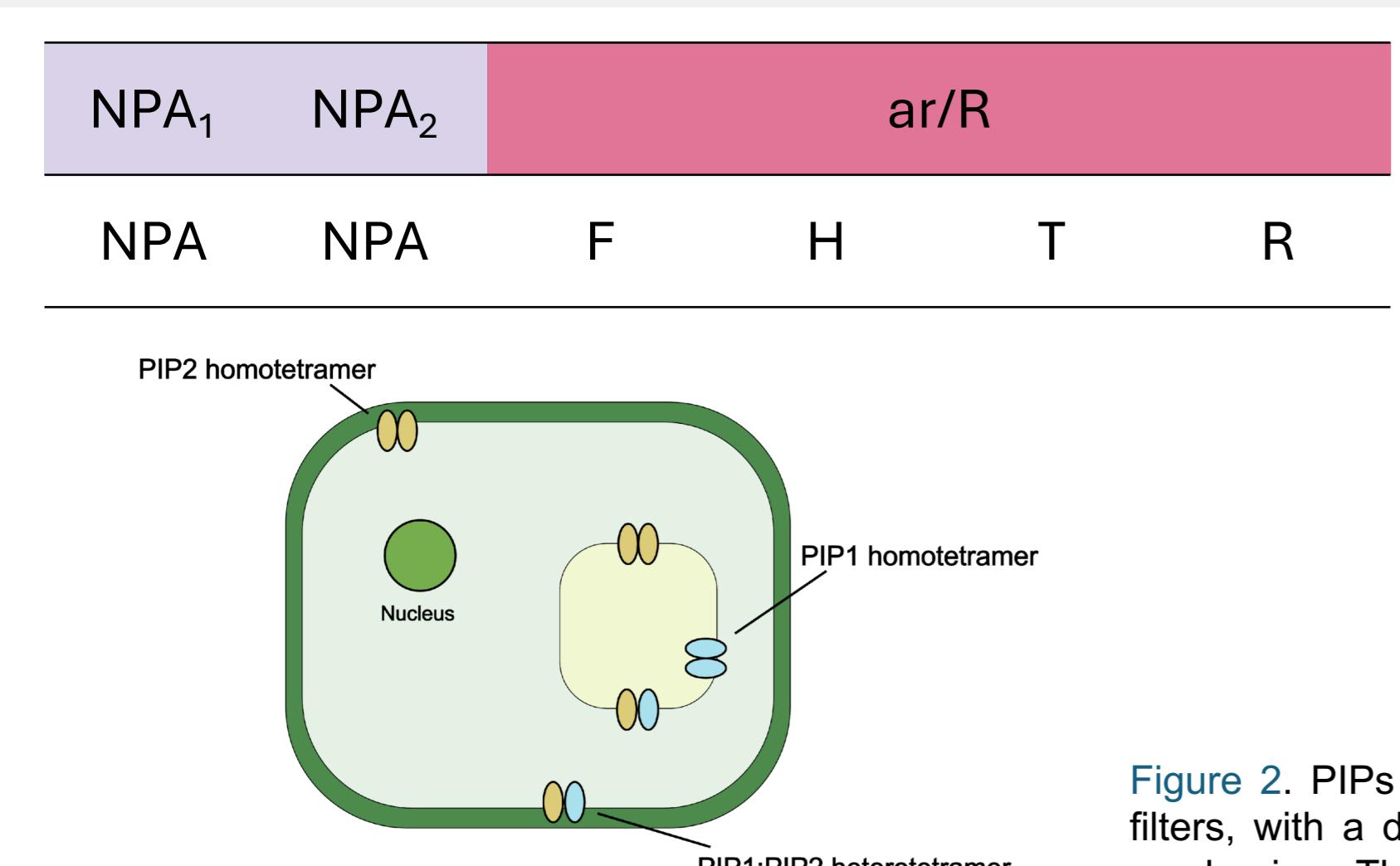
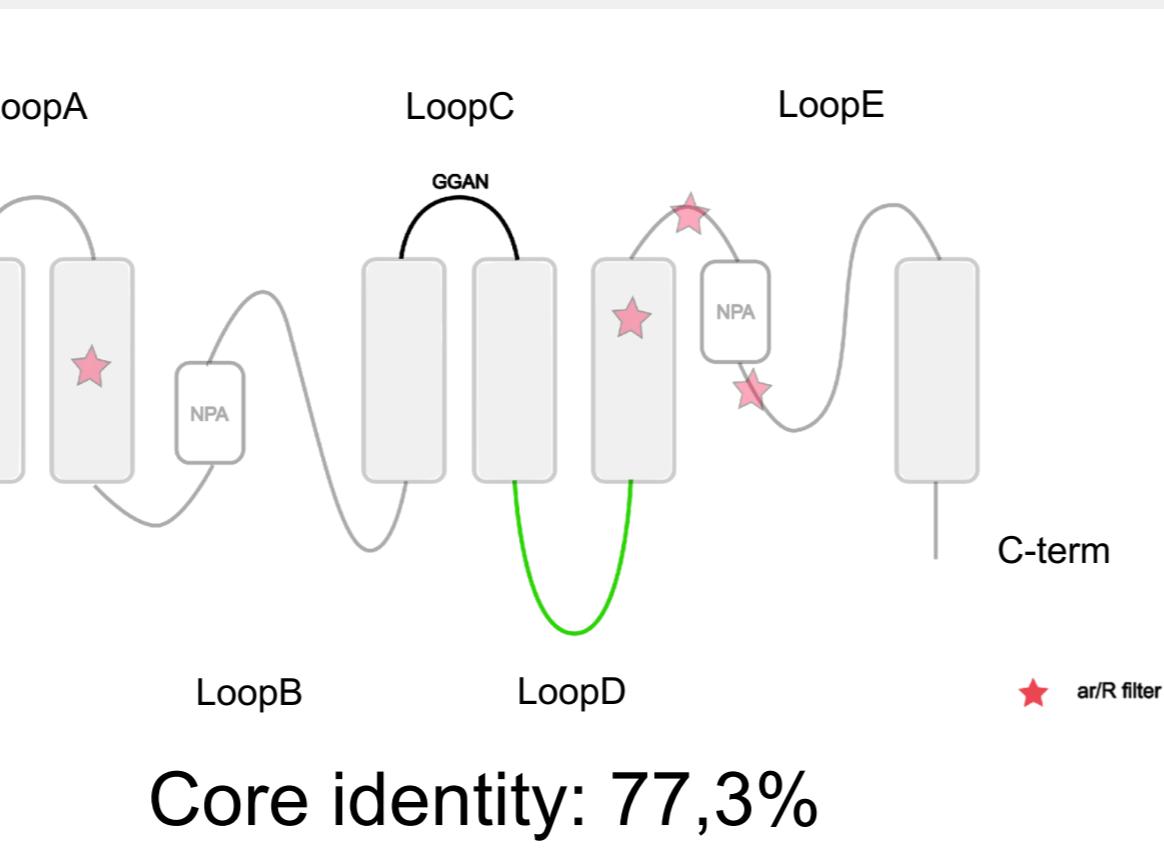


Figure 2. PIPs belong to the Plantae kingdom. Paralogues present conserved selectivity filters, with a distinctive loopC motif and a particularly long loopD implied in their gating mechanism. They exhibit different cellular localisation according to tetramer composition.

Jozefkowicz et al., 2017



Core identity: 77,3%

PIP2 channels exhibit greater **sequence heterogeneity**, in contrast to the more conserved domains of PIP1. Both paralogues **conserve the length of transmembrane domains**.

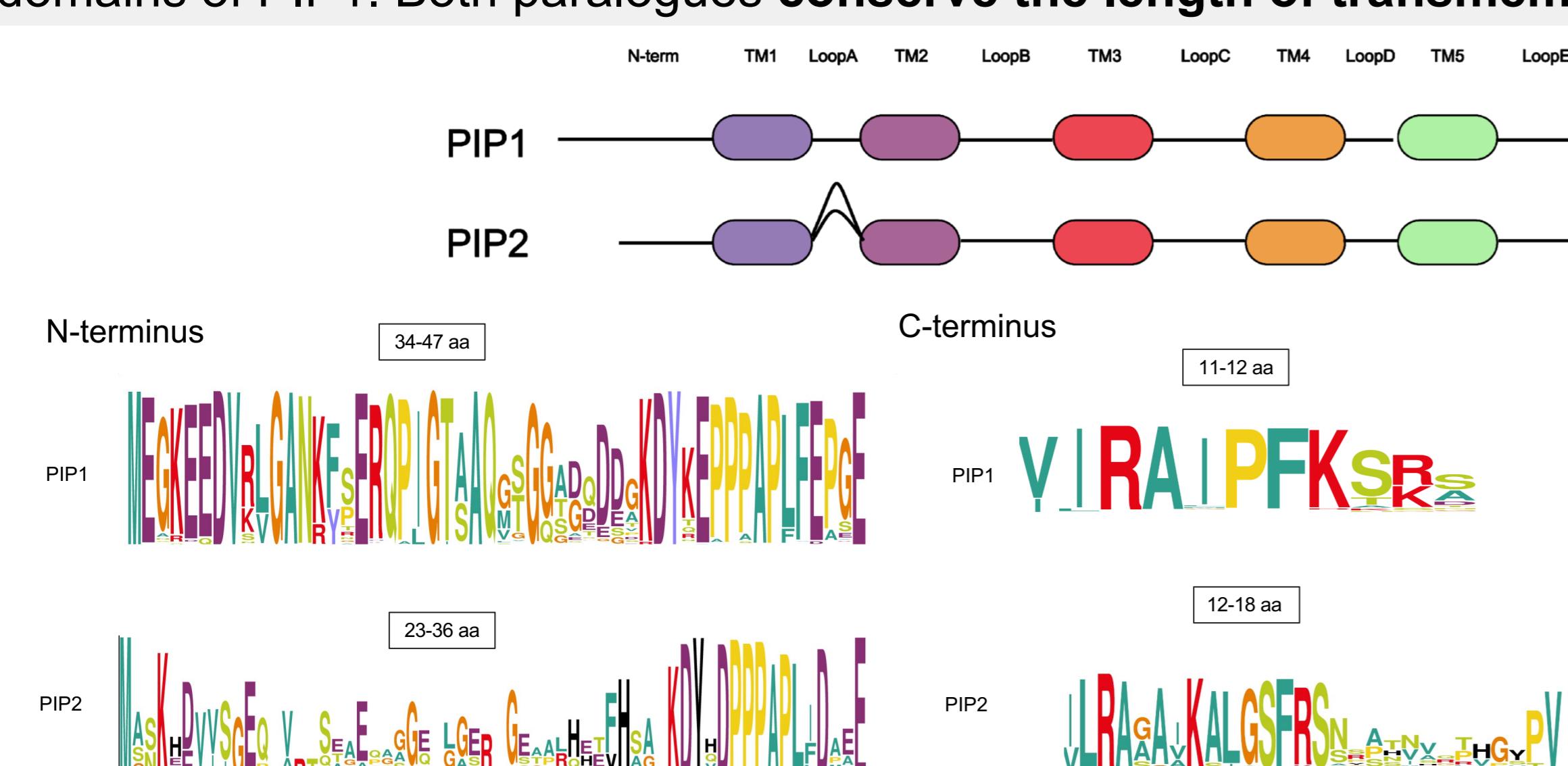


Figure 3. The length of PIP transmembrane domains is conserved, while the length of their terminal ends and loopA varies. PIP1 presents more conserved domains, with a longer N-terminus and shorter C-terminus, while PIP2 presents more heterogeneity in its loopA and N and C-termini.

The differential clustering of PIP1 and PIP2 is influenced by more than just variations in their N- and C-termini, reflecting additional structural distinctions between the paralogues.

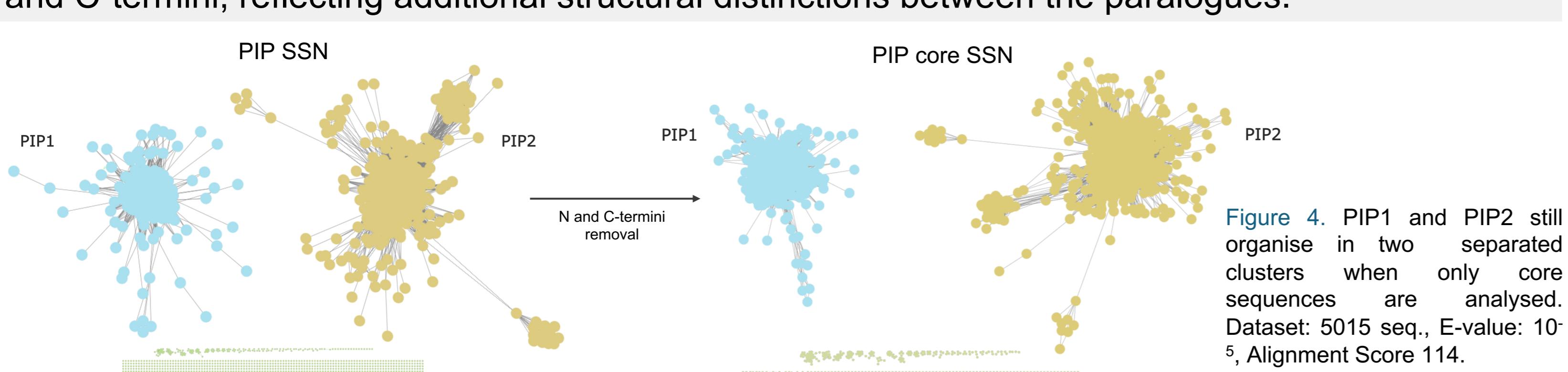


Figure 4. PIP1 and PIP2 still organise in two separated clusters when only core sequences are analysed. Dataset: 5015 seq., E-value: 10^{-5} , Alignment Score 114.

The **determinant structural motifs** of PIP1 and PIP2 are located in their **transmembrane domains**, while their protomeric pores remain identical.

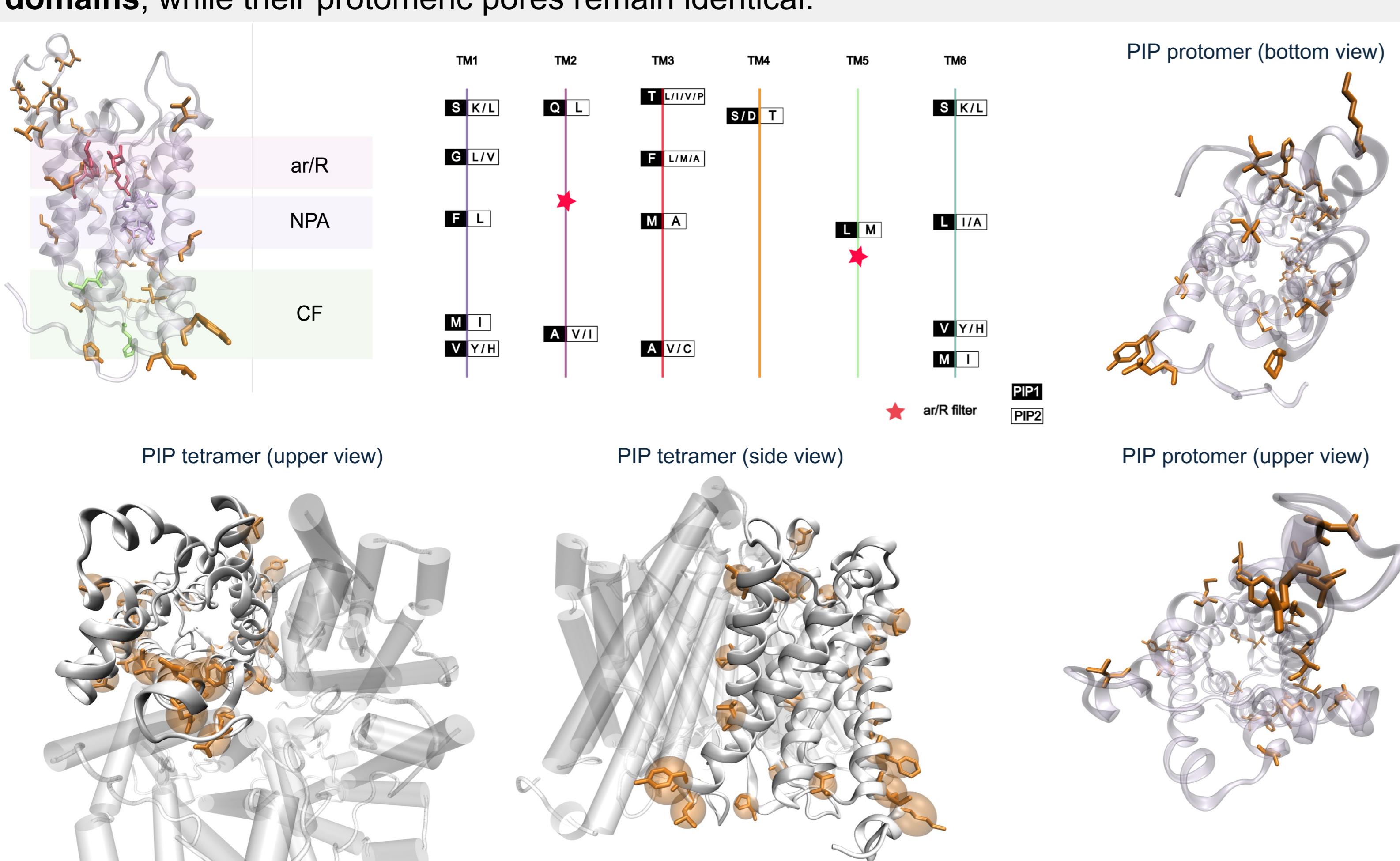


Figure 5. PIP paralogues' determinant structural motifs (orange) were selected based on Sequence Harmony and multiRelief parameters ($SH<0,2$; $mR>0,8$).

Materials and methods

Datasets Obtention: MIP dataset was built with PFAM00230 sequences in UniProt. PIP dataset was obtained from this dataset, adding NCBI and EukProt sequences for enrichment, filtered by length (200-400aa), PIP motifs and redundancy. AQPX dataset also derived from the MIP dataset and was complemented with NCBI and TriTryp sequences.

Sequence Similarity Network: SSNs were generated with EPI-EST using an E-value of 10^{-5} and an appropriate Alignment Score for each dataset (MIP SSN: 60 - UniRef90, PIP SSN: 122, AQPX SSN: 39). Cytoscape was used for SSN visualisation. Multiple Sequences Alignments were performed with MAFFT and treated in Jalview.

Identification of Motifs: Distinctive and conserved paralogues residues were identified with Multi-Harmony.

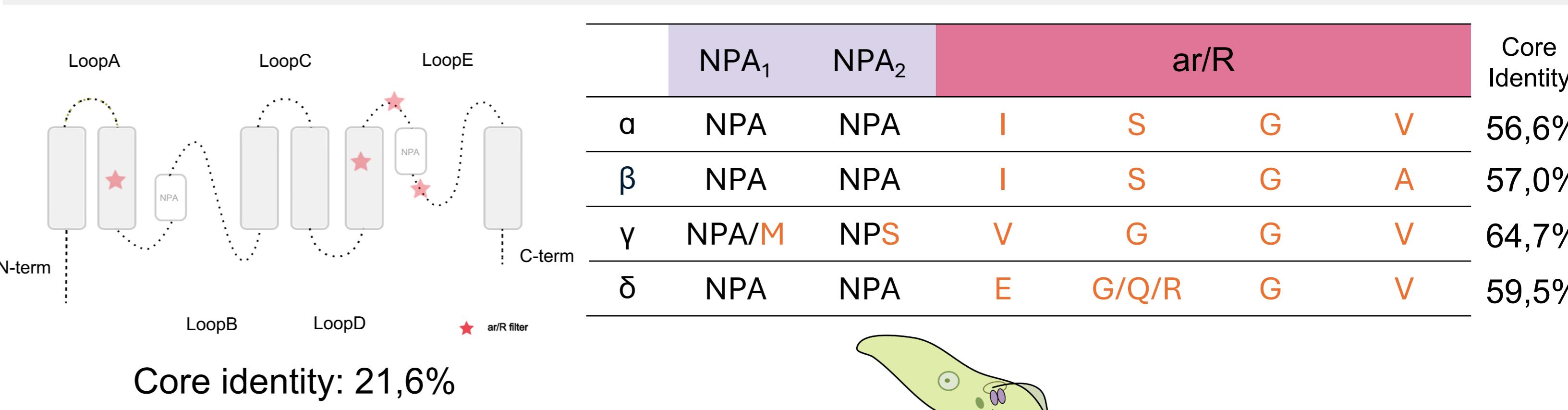
Bibliography

Jozefkowicz, C., Berry, M.C., Chaumont, F., Alleva, K. (2017). Heteromericity of Plant Aquaporins. In: Chaumont, F., Tyerman, S. (eds) Plant Aquaporins. Signaling and Communication in Plants. Springer, Cham. doi: 10.1007/978-3-319-49395-4_2

Montalvetti A, Rohloff P, Docampo R. (2004). A functional aquaporin co-localizes with the vacuolar proton pyrophosphatase to acidocalcisomes and the contractile vacuole complex of Trypanosoma cruzi. J Biol Chem. doi: 10.1074/jbc.M406304200.

AQPX subfamily

AQPX is a recently described and still functionally uncharacterised subfamily, with its four paralogues exhibiting **distinctive ar/R filters**.



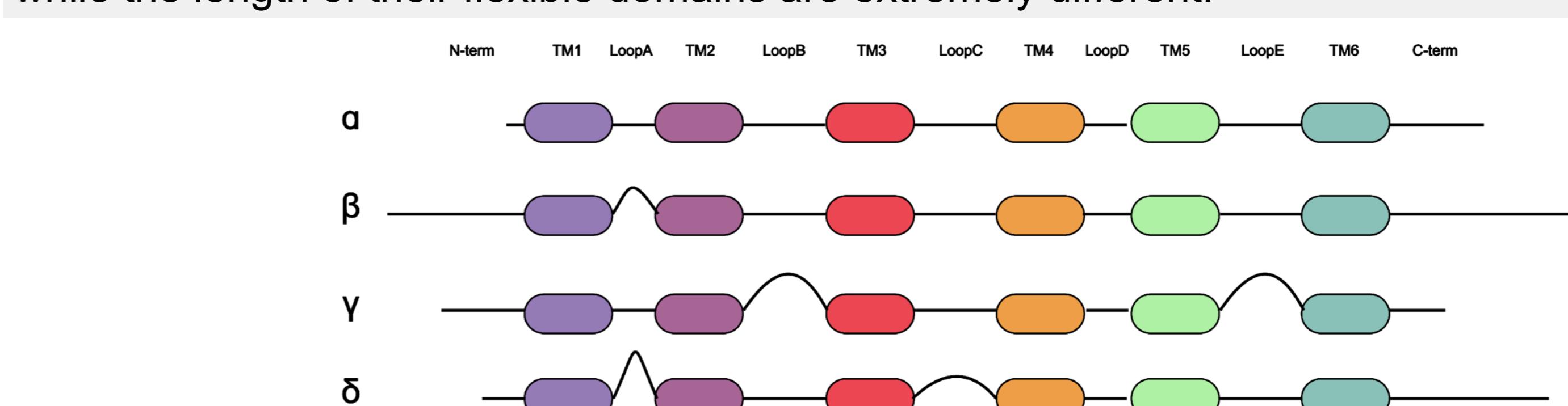
Core identity: 21,6%

Figure 6. AQPXs are present in unicellular organisms. No 3D structure has been obtained for this subfamily. There are few reports regarding localisation, but it has been suggested that AQPX representatives locate in intracellular membranes.

Ramoa 2024, unpublished results

Montalvetti et al., 2004

The four AQPX paralogues **conserve the length of their transmembrane domains and loopD**, while the length of their flexible domains are extremely different.



AQPX SSN includes unicellular organisms across all domains of life.

There's one determinant structural residue of AQPX paralogues, located next to an ar/R filter residue.

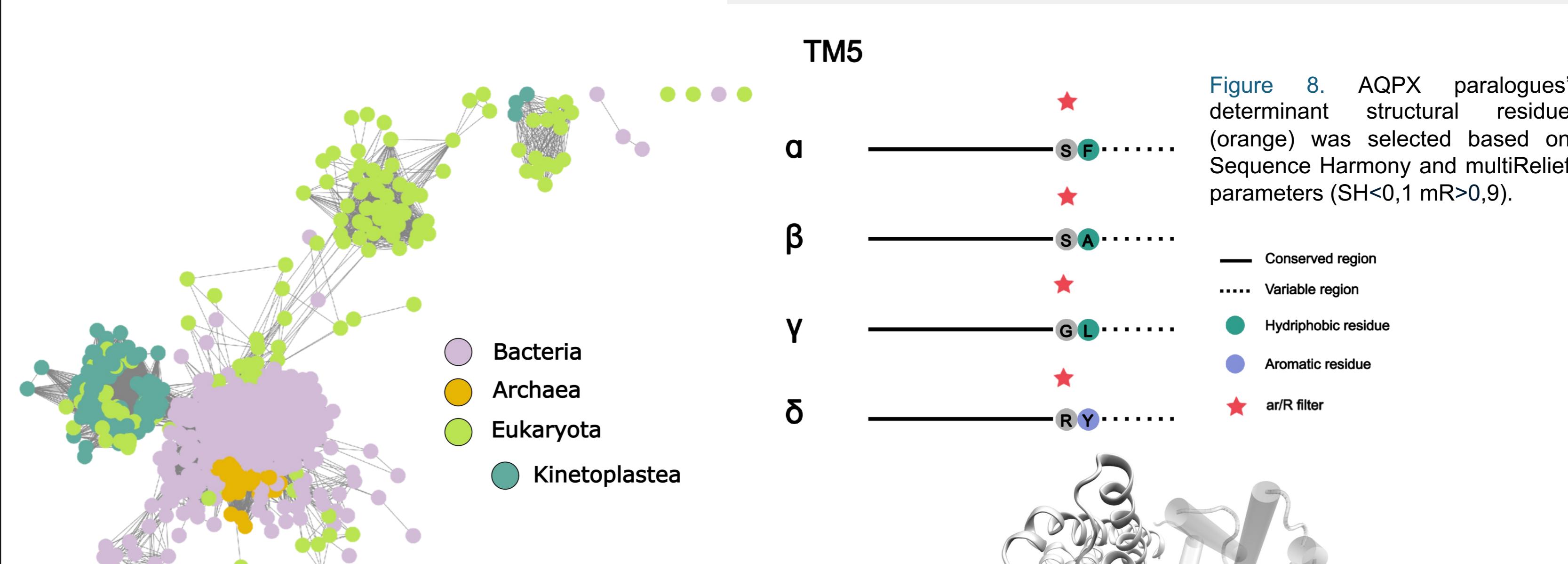


Figure 7. AQPX core SSN. The Kinetoplastea class, which includes multiple parasites responsible of human diseases, group in a subcluster of the AQPX SSN. Dataset: 1034 seq., E-value: 10^{-5} , Alignment Score 39.

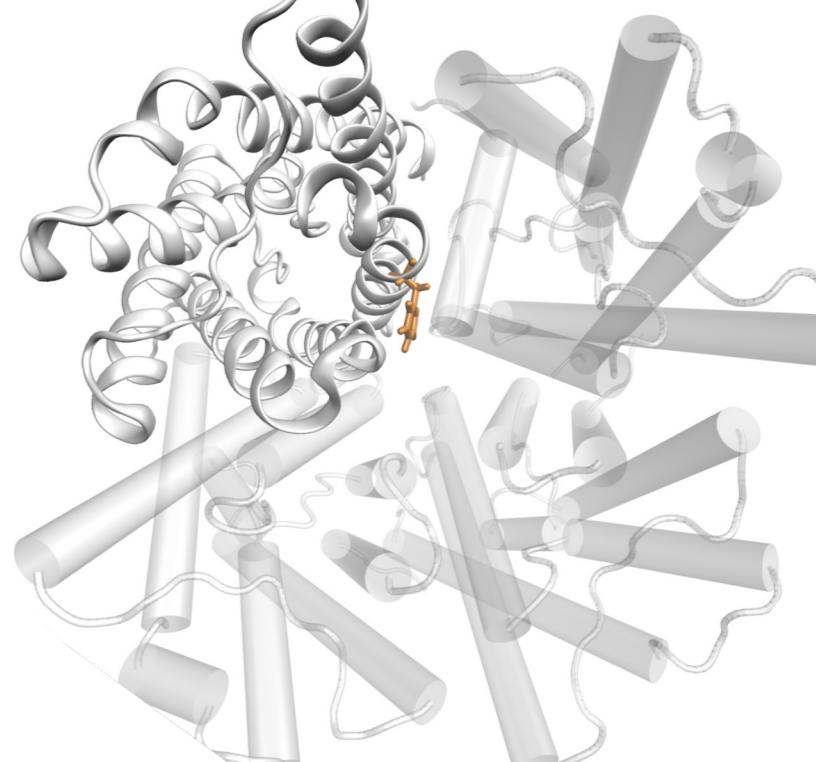
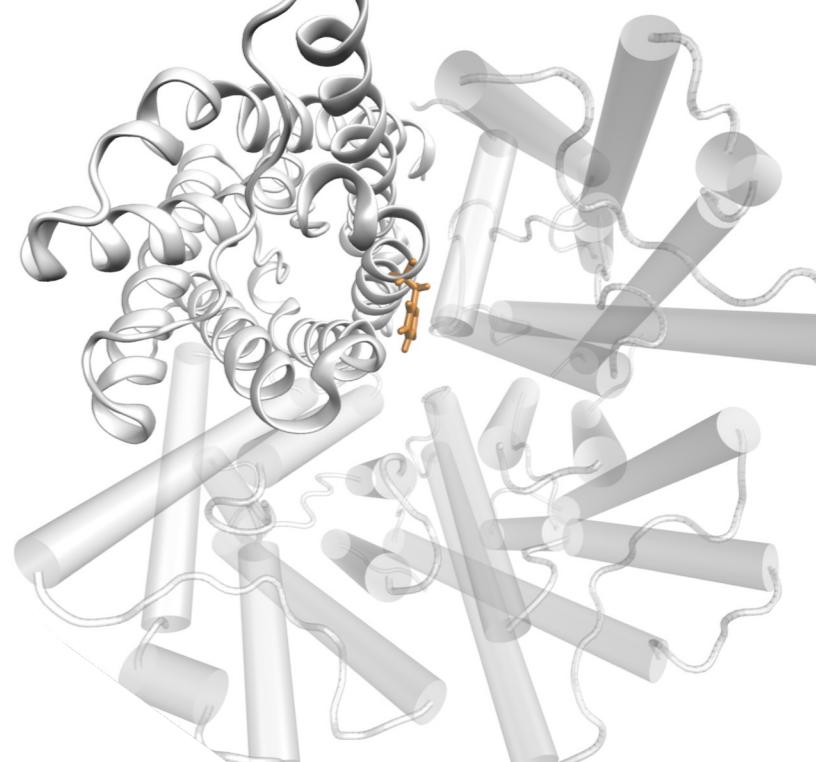


Figure 8. AQPX paralogues' determinant structural residue (orange) was selected based on Sequence Harmony and multiRelief parameters ($SH<0,1$, $mR>0,9$).

— Conserved region
.... Variable region
● Hydrophobic residue
● Aromatic residue
★ ar/R filter



There are **highly conserved distinctive motifs** within variable loops in the AQPX cluster.

Figure 9. Variable loopA does not present conserved motifs, while loopC presents an unusual proline motif and loopD presents conserved polar and charged residues. AQPX conserved motifs were selected based on Sequence Harmony and multiRelief parameters ($SH>0,8$, $mR<0,2$).

	LoopA	LoopB	LoopC	LoopD	LoopE
α		Y_SG_H_NPA	P(A/V)P	S(R/Q)Q_(R/S)_N	G_FF_NPA_AT
β		Y_SG_H_NPA	(ar)VP	S(R/Q)Q_(R/S)_N	G_FF_NPA_AT
γ		Y_SG_H_NP(A/M)	PAP	S(R/Q)Q_(R/S)_N	G_FF_NPM_AT
δ		Y_SG_H_NPA	PVP	S(R/Q)Q_(R/S)_N	G_FF_NPA_AT

Conclusions

PIP1 and PIP2 determinant structural motifs in their transmembrane domains could be responsible of their differential trafficking, central pore transport and/or tetramer assembly.

Conserved motifs in AQPX variable loops could be determinant motifs of this subfamily. The distinctive residue of AQPX paralogues' role is yet to be explored.

These different structural motifs within MIP subfamilies are our starting point to explore the biotechnological potential of these channels for targeted transport, directed localisation and altering their physiological function.