

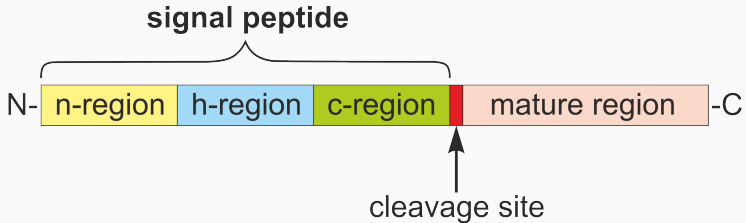
signalHsmm: prediction of malarial signal peptides

Michał Burdukiewicz^{1*}, Piotr Sobczyk², Paweł Błazej¹, Paweł Mackiewicz¹

¹University of Wrocław, Department of Genomics,

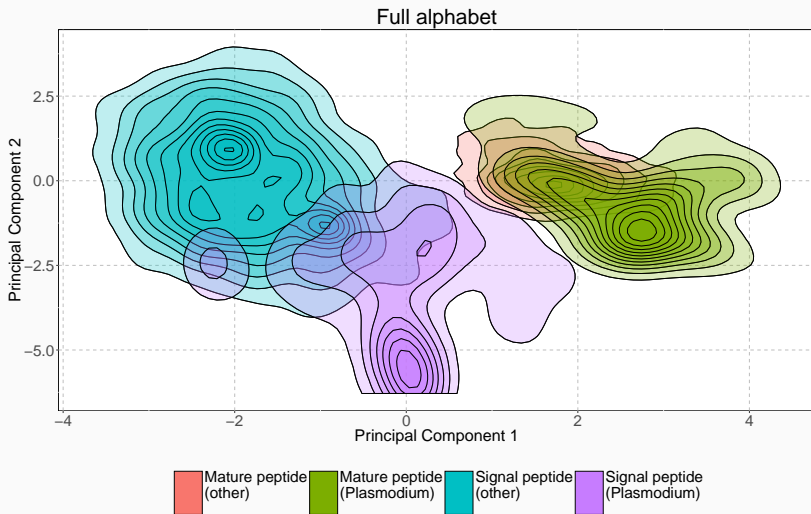
²Wrocław University of Science and Technology, Faculty of Pure and Applied Mathematics,

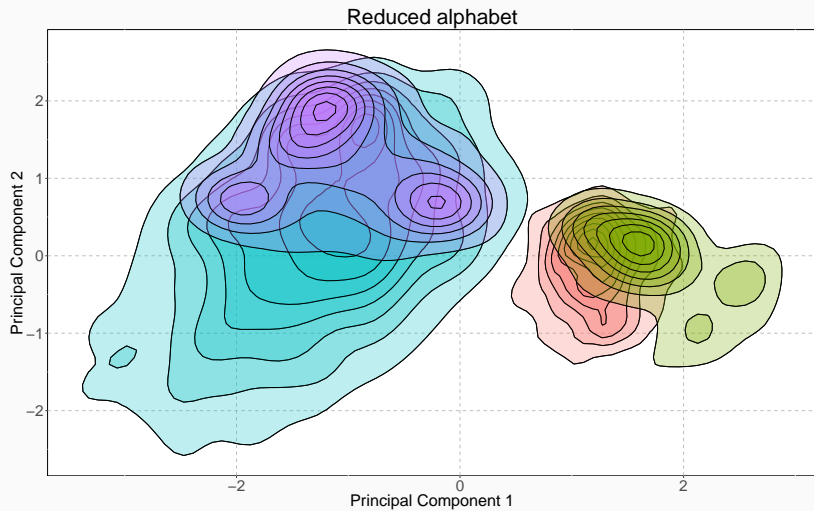
Signal peptides

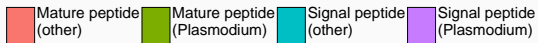
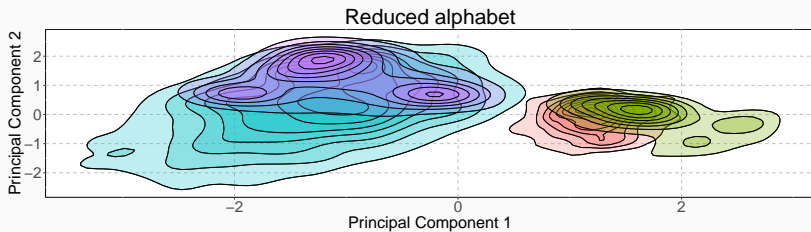
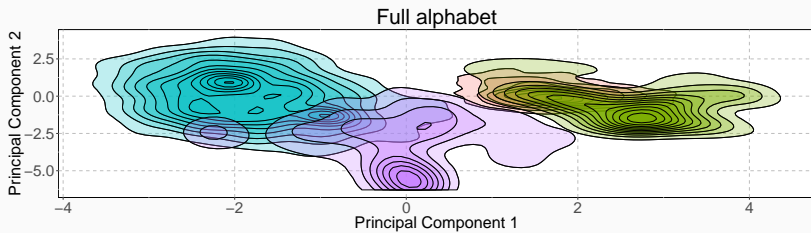


- n-region: mostly basic residues (Nielsen and Krogh, 1998),
- h-region: strongly hydrophobic residues (Nielsen and Krogh, 1998),
- c-region: a few polar, uncharged residues (Jain et al., 1994).

Group	Amino acids
1	D, E, H, K, N, Q, R
2	G, P, S, T, Y
3	F, I, L, M, V, W
4	A, C







Countour plot of first two components in Pricincipal Compoment Analysis of amino acid frequency. The signal peptides from malaria and other taxons differ significantly when the full amino acid alphabet is employed. After the reduction of the alphabet, the signal peptides group together despite their origin.

Acknowledgements and funding

This research was partially funded by the KNOW Consortium and National Science Center (2015/17/N/NZ2/01845).

- Paweł Mackiewicz.
- **biogram** package
(<https://cran.r-project.org/package=biogram>):
 - Piotr Sobczyk,
 - Chris Lauber.

References

- Jain, R. G., Rusch, S. L., and Kendall, D. A. (1994). Signal peptide cleavage regions. functional limits on length and topological implications. *The Journal of Biological Chemistry*, 269(23):16305–16310.
- Nielsen, H. and Krogh, A. (1998). Prediction of signal peptides and signal anchors by a hidden markov model. *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology*, 6:122–130.