

# signalHsmm: prediction of malarial signal peptides

---

Michał Burdukiewicz<sup>1\*</sup>, Piotr Sobczyk<sup>2</sup>, Paweł Błazej<sup>1</sup>, Paweł Mackiewicz<sup>1</sup>

<sup>1</sup>University of Wrocław, Department of Genomics,

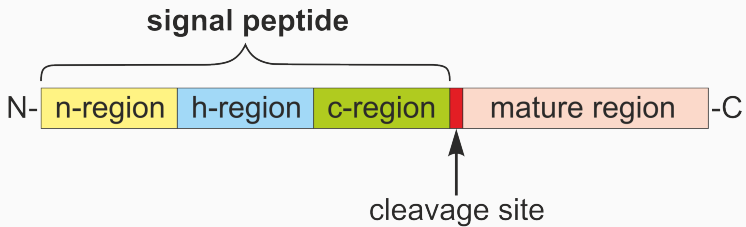
<sup>2</sup>Wrocław University of Science and Technology, Faculty of Pure and Applied Mathematics,

## Signal peptides

---

## Secretory signal peptides:

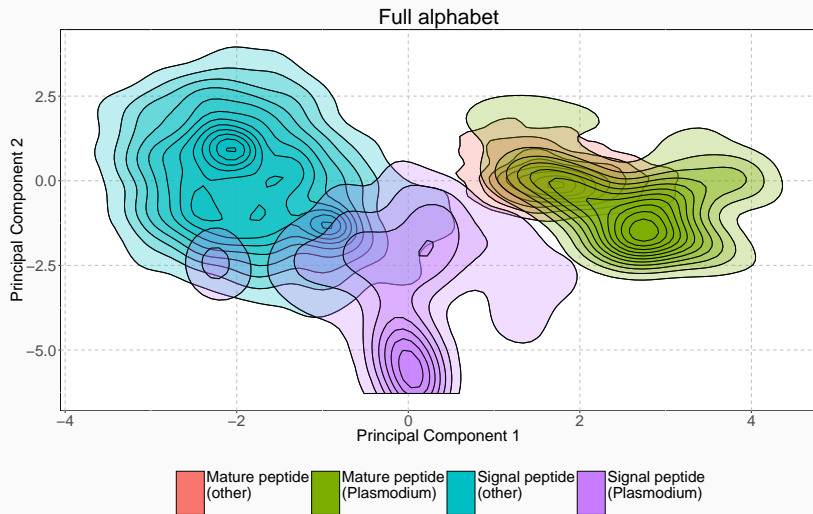
- are short (20-30 residues) N-terminal amino acid sequences,
- direct a protein to the endomembrane system and next to the extracellular localization,
- are universal enough to direct properly proteins in different secretory systems; artificially introduced bacterial signal peptides can guide proteins in mammals (Nagano and Masuda, 2014) and plants (Moeller et al., 2009),
- tag among others hormones, immune system proteins, structural proteins, and metabolic enzymes.

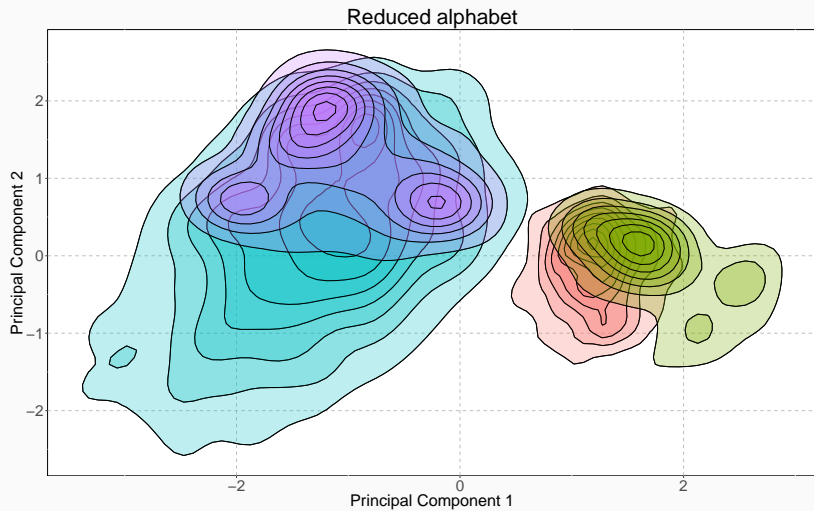


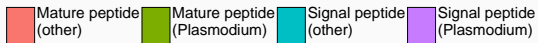
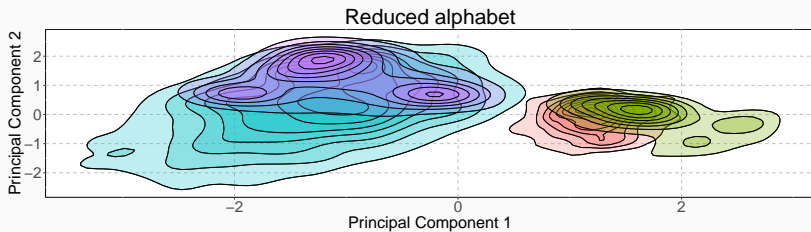
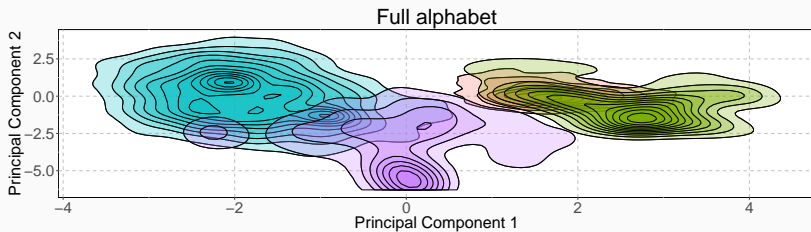
Signal peptides possess three distinct domains with variable length and characteristic amino acid composition (Hegde and Bernstein, 2006):

- 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

---

- Hegde, R. S. and Bernstein, H. D. (2006). The surprising complexity of signal sequences. *Trends in Biochemical Sciences*, 31(10):563–571.
- 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.

## References II

- Moeller, L., Gan, Q., and Wang, K. (2009). A bacterial signal peptide is functional in plants and directs proteins to the secretory pathway. *Journal of Experimental Botany*, 60(12):3337–3352.
- Nagano, R. and Masuda, K. (2014). Establishment of a signal peptide with cross-species compatibility for functional antibody expression in both escherichia coli and chinese hamster ovary cells. *Biochemical and Biophysical Research Communications*, 447(4):655 – 659.

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.