Using an adaptive lifting scheme in predicting transmembrane helix locations along a protein sequence

Marina Popa* & Guy P. Nason

University of Bristol

1 Introduction

Classical wavelet techniques and various thresholding methods have been used so far in smoothing protein hydropathy profiles. The current literature mentions possible improvements through the use of sequentially aligned proteins, but classical wavelets do not allow for such extensions.

2 Outline

Starting from the lifting scheme, capable to work on irregularly spaced grids (introduced by Sweldens in 1994), we have developed an adaptive version of it. Wavelet functions will be adaptively constructed at every step, by taking into account the signal particularities within a selected window.

Next, a hydropathy profile will be constructed for each protein of interest through incorporating the information conveyed by sequentially aligned proteins with resolved three-dimensional structure. This will result in having to smooth an unequally spaced signal on the horizontal axis. Our adaptive lifting transform will be used to decompose the signal into wavelet and scaling coefficients. The wavelet coefficients will then be thresholded by means of an empirical Bayes procedure (Johnstone and Silverman, 2002) and the algorithm inverted, generating a smoothed profile.

Based on the centred smoothed profile, segments incorporating amino acids with hydropathy values above zero will be classified as transmembranar segments, without using a further biochemical filtering step.

3 Conclusions

Being able to use further information, provided by the usage of aligned sequences, improves the accuracy of prediction of the transmembrane segments.

References

Fisher, P., Baudoux, G. and Wouters, J. (2003). WAVPRED: a Wavelet-Based Algorithm for the Prediction of Transmembrane Proteins. *Comm. Math. Sci.* **Vol. 1**, 44-56.

Jansen, M., Nason, G.P. and Silverman, B.W. (2003). To come.

Johnstone, I.M. and Silverman, B.W. (2002). EbayesThresh: R and S-PLUS software for Empirical Bayes Thresholding. As yet unpublished.

Lio, P. and Vannucci, M. (2000). Wavelet change-point prediction of transmembrane proteins. *Bioinformatics*, **Vol. 16**, 376-382.

Nason, G.P., Nunes, M.A. and Popa, M.I. (2004). To come.

Sweldens, W. (1997). The lifting scheme: a construction of second generation wavelets. *SIAM J. Math. Anal.*, **Vol. 29**, 511-546.