Relating amino acid patterns to successful high-level protein secretion in *Aspergillus niger*

B.A. van den Berg1,3,4, M. Hulsman1,4, M.J.T Reinders1,3,4, L. Wu1, H.J. Pe1, J.A. Roubos1, D. de Ridder1,3,4

1 The Delft Bioinformatics Lab, Faculty of Electrical Engineering, Mathematics & Computer Science, Delft University of Technology, Delft, The Netherlands; 2 DSM Biotechnology Center, Delft, The Netherlands; 3 Netherlands Bioinformatics Centre, Nijmegen, The Netherlands; 4 Kluyver Centre for Genomics of Industrial Fermentation, Delft, The Netherlands

* b.a.vandenberg@tudelft.nl

Introduction

*Aspergillus niger* is widely used for industrial enzyme production. Knowledge on high-level protein secretion could be useful to improve production rates. We used sequence-based classification methods to relate amino acid patterns to successful high-level secretion.

Methods & Results

Success of high-level secretion of 416 over-expressed homologous proteins was tested in the lab and used as data set. To obtain defining ‘generalized’ amino acid patterns, we aimed to improve similarity within a class (Fig.1).

We developed a hierarchical feature construction method that clusters amino acids (Fig.2). The resulting clusters seem to correspond to physicochemical clusters (Fig.3). Some interesting generalized amino acid patterns that are predictive for successful high-level secretion were found (Fig.4).

Conclusions

We showed that hierarchical feature construction can be used to obtain generalized amino acid patterns predictive for successful high-level secretion. The amino acid clusters found by the method seem to correspond to known physicochemical clusters, indicating biological relevance. Currently we are investigating if the occurrence of the patterns correlate to specific structural regions, e.g. if a pattern consistently occurs in a helix region. In a later stage, this knowledge could possibly be used to rationally redesign proteins for improved secretion.