

Daniela Albrecht

Integration of transcriptome and proteome data from human-pathogenic fungi

Many research groups use high-throughput methods for studies at the transcriptome and proteome level, such as microarrays and two-dimensional gel electrophoresis (2D-GE). Data created by these techniques are very large and hardly inspected by eye or analysed by hand. Bioinformatic methods are required to filter out the biological meaning from the wealth of information. The final aim of my dissertation is to combine transcriptomic and proteomic data in an attempt to get a more holistic view on the fungal infection process. However, the proteomic data often do not have a sufficient quality. In contrast to microarray data, preprocessing of 2D-GE data has only rarely been a research subject.

That is why the first part of my work focuses on this topic and aims to create a standardised workflow for the analysis of such data. A number of different factors must be taken into consideration to gain proteomic data of high quality. This includes the experimental design, handling of missing values, normalization and filtering of the raw data. Each single step of the proposed workflow greatly influences the number of differentially regulated proteins. Therefore, it should use dataset specific parameters. The result of the whole pre-processing procedure is a list of potentially interesting proteins that has to be interpreted. Assigning functional annotations to proteins and categorising them into broader categories is a promising approach.

In the second part of my work, a transcriptomic and a proteomic time series dataset is analysed. The research topic of both datasets is the response of *Aspergillus fumigatus* to a temperature shift from 30°C to 48°C. The proteomic dataset is pre-processed using the new standardised workflow to get a list of differentially regulated proteins. Transcriptome and proteome data are compared using two correlations and one information theoretical measure. Additionally, Coinertia analysis is used for visualisation of both datasets. Results are augmented by bioinformatical search for transcription factor binding sites (TFBSs) of heat shock regulators and comparison to *Saccharomyces cerevisiae* and other *Aspergillus* species. As third part of my dissertation a data warehouse as central store for transcriptomic and proteomic data from different working groups is established and maintained. I implement routines for importing and exporting of certain data formats and collect datasets. International standards are considered for data annotation. Several analysis tools complete the database. The proteome workflow including functional analysis is implemented. Additionally a tool for promoter analysis and one for analysis of infection models are created and further tools are in progress.

Publications

Priebe S, Linde J, Albrecht D, Guthke R, Brakhage AA (2011) FungiFun: a web-based application for functional categorization of fungal genes and proteins. *Fungal Genet Biol* 48(4), 353-358. [Details](#) [PubMed](#)

Albrecht D, Guthke R, Brakhage AA, Kniemeyer O (2010) Integrative analysis of the heat shock response in *Aspergillus fumigatus*. *BMC Genomics* 11, 32. [Details](#) [PubMed](#)

Albrecht D, Kniemeyer O, Brakhage AA, Guthke R (2010) Missing values in gel-based proteomics. *Proteomics* 10(6), 1202-1211. [Details](#) [PubMed](#)

Bruns S, Seidler M, Albrecht D, Salvenmoser S, Remme N, Hertweck C, Brakhage AA, Kniemeyer O, Müller FM (2010) Functional genomic profiling of *Aspergillus fumigatus* biofilm reveals enhanced production of the mycotoxin gliotoxin. *Proteomics* 10(17), 3097-3107. [Details](#) [PubMed](#)

Shelest V, Albrecht D, Shelest E (2010) DistanceScan: a tool for promoter modeling. *Bioinformatics* 26(11), 1460-1462. [Details](#) [PubMed](#)

Teutschbein J, Albrecht D, Pötsch M, Guthke R, Aimaganianda V, Clavaud C, Latgé JP, Brakhage AA, Kniemeyer O (2010) Proteome profiling and functional classification of intracellular proteins from conidia of the human-pathogenic mold *Aspergillus fumigatus*. *J Proteome Res* 9(7), 3427-3442. [Details](#) [PubMed](#)

Vödisch M, Albrecht D, Lessing F, Schmidt AD, Winkler R, Guthke R, Brakhage AA, Kniemeyer O (2009) Two-dimensional proteome reference maps for the human pathogenic filamentous fungus *Aspergillus fumigatus*. *Proteomics* 9(5), 1407-1415. [Details](#) [PubMed](#)

Albrecht D, Guthke R, Kniemeyer O, Brakhage AA (2008) Systems biology of human-pathogenic fungi. In: Daskalaki A (ed.) Handbook of research on systems biology applications in medicine. 1, pp. 400-418. IGI Global. [Details](#)

Albrecht D, Kniemeyer O, Brakhage AA, Guthke R (2007) Integration of transcriptome and proteome data from human-pathogenic fungi by using a data warehouse. *J Integrative Bioinf* 4, 52. [Details](#)

Guthke R, Kniemeyer O, Albrecht D, Brakhage AA, Moeller U (2007) Discovery of gene regulatory networks in *Aspergillus fumigatus*. *Lect Notes Bioinf* 4366, 22-41. [Details](#)

Supervisor

[Reinhard Guthke](#)

Start of PhD

May 1, 2006

Doctoral Disputation

June 1, 2010