International Journal of Medical Microbiology 301 (2011) 453–459

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.de/ijmm

Mini review

On the way toward systems biology of *Aspergillus fumigatus* infection

Daniela Albrecht[a], Olaf Kniemeyer[b,c], Franziska Mech[a], Matthias Gunzer[d], Axel Brakhage[b,c], Reinhard Guthke[a,*]

[a] *Research Group Systems Biology/Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute, Beutenbergstr. 11a, 07745 Jena, Germany*
[b] *Dept. Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute, Jena, Germany*
[c] *Dept. Microbiology and Molecular Biology, Friedrich Schiller-University, Jena, Germany*
[d] *Institute for Molecular and Clinical Immunology, Otto-von-Guericke-University, Magdeburg, Germany*

ARTICLE INFO

Keywords:
Systems biology
Fungal infection
Data warehouse
Transcriptome
Proteome
Network inference
Movies

ABSTRACT

Pathogenicity of *Aspergillus fumigatus* is multifactorial. Thus, global studies are essential for the understanding of the infection process. Therefore, a data warehouse was established where genome sequence, transcriptome and proteome data are stored. These data are analyzed for the elucidation of virulence determinants. The data analysis workflow starts with pre-processing including imputing of missing values and normalization. Last step is the identification of differentially expressed genes/proteins as interesting candidates for further analysis, in particular for functional categorization and correlation studies. Sequence data and other prior knowledge extracted from databases are integrated to support the inference of gene regulatory networks associated with pathogenicity. This knowledge-assisted data analysis aims at establishing mathematical models with predictive strength to assist further experimental work. Recently, first steps were done to extend the integrative data analysis and computational modeling by evaluating spatio-temporal data (movies) that monitor interactions of *A. fumigatus* morphotypes (e.g. conidia) with host immune cells.

Introduction

'Systems biology' is a multidisciplinary research area that combines experimental discovery with mathematical modeling to assist the understanding of the dynamic global organization and function of a biological system. This emerging research field has demonstrated the most progress for unicellular microorganisms such as *Escherichia coli* and *Saccharomyces cerevisiae*. Recently, improvement has also been achieved in systems biology of filamentous fungi of molecular biological and industrial interest, in particular of *Aspergillus nidulans* (David et al., 2006), *A. niger* (Andersen et al., 2008), and others (Harris, 2009; Andersen and Nielsen, 2009). The largest reconstruction of a metabolic network reported for a filamentous fungus was presented for the industrial workhorse *A. niger* comprising 2240 enzymatic reactions (Andersen et al., 2008). Despite the assumption that there is an overlap in genes and metabolic reactions between different *Aspergillus* species of about 70% (Galagan et al., 2005; Andersen et al., 2008), it can be expected that genes and reactions responsible for the pathogenicity differ. Thus, analytical methods, but not the knowledge about molecular details gained from other filamentous fungi can be transferred to the human-pathogenic fungus *A. fumigatus*. For instance, the molecular approach to reconstruct the signaling pathways that control hyphal elongation and branching applied by Meyer et al. (2009) to *A. niger* can be used to study the mycelial growth of pathogenic filamentous fungi such as *A. fumigatus*.

Currently, systems biological research is more advanced for bacterial than for fungal infections. In particular, there are very promising results available from systems biology studies of bacterial pathogens such as *Escherichia coli* (Guthke et al., 2005), *Salmonella typhimurium* (Raghunathan et al., 2009), *Pseudomonas aeruginosa* (Oberhardt et al., 2008), *Mycobacterium tuberculosis* (Raman et al., 2008), *Helicobacter pylori* (Franke et al., 2008), and *Streptococcus pyogenes* (Klenk et al., 2005). Vodovotz et al. (2008) proposed the term 'translational systems biology' for the application of systems biological and engineering principles to biological systems with the primary goal of optimizing clinical practice in sepsis treatment. Day et al. (2010) applied this to tuberculosis research.

Host–pathogen systems biology has to model not only the molecular network in one organism or population, but also the interaction of the networks of pathogen and host (Franke et al., 2008). For example, the merging of a viral interactome and the human interactome has been used to simulate viral infection (Pelkmans, 2009).

Systems biology of fungal pathogenicity is still in its infancy (Albrecht et al., 2008a). First steps were taken in systems biological

* Corresponding author. Tel.: +49 3641 1083; fax: +49 3641 2083.
E-mail address: reinhard.guthke@hki-jena.de (R. Guthke).

1438-4221/$ – see front matter © 2011 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ijmm.2011.04.014

research of infections caused by *Candida albicans*, in particular by game theoretic modeling of the host–pathogen interplay (Hummert et al., 2010) and the prediction of multi-agent antifungal therapies (Jansen et al., 2009).

Researchers are only just starting to understand the complex interplay between the host and *A. fumigatus* at a genome-wide scale. Since fungal virulence is a multifactorial and complex process, holistic approaches such as transcriptomic and proteomic technologies seem to be most promising for its elucidation. High-throughput analyses allow the discovery-driven identification of candidate genes or proteins, while also providing the necessary context in which hypothesis-driven findings can be interpreted (Chan, 2006). Thus, systems biology leads to new models, new predictions, and new experiments to test those (Ideker et al., 2001).

When comparing transcriptomic and proteomic data, often poor correlation of expression is found (e.g., Jones et al., 2004). Thus, there is apparently no strict linear relationship between the transcripts and the proteins of a cell. In addition, observed dissimilarities between monitored behavior on transcriptomic and proteomic level can show important post-transcriptional or post-translational regulatory junctures. Hence, it is clear that transcriptomic and proteomic approaches complement each other in terms of the information produced and their relative advantages and disadvantages. So at best, both are used, and data are integrated.

Another approach to observe and document host–pathogen interaction during fungal infections is provided by techniques like confocal laser scanning microscopy or two-photon microscopy. Thus, images and movies of host immune cells and opposing pathogens are received. These enable researchers to investigate cellular dynamics and the links between cellular functions and spatio-temporal localization in the living system. The analysis and feature extraction is complicated by the fact that the data are typically very noisy (Sage et al., 2005). Many current surveys are based on the study of a few manually selected particles and thus, are time consuming and can produce a strong bias in the analysis (Genevesio et al., 2006). It is therefore important to develop automatic and reproducible methods to perform feature extraction and cell tracking.

Aspergillus fumigatus

Around 150–200 fungal species are known to be associated with human infections. During the past decades, the incidence of opportunistic invasive mycoses has risen to a considerable amount. For example, from 1980 to 1997 there was a dramatic increase in mycoses-associated mortality in the USA (McNeil et al., 2001). *Aspergillus* species are the second most frequent cause of nosocomial (i.e., acquired in hospital) fungal infections after *Candida* species (Perlroth et al., 2007). Invasive aspergillosis is mostly caused by *A. fumigatus*, the most important airborne human-pathogenic fungus (Dagenais and Keller, 2009). *A. fumigatus* is a thermotolerant species able to grow at up to 55 °C and surviving even up to 70 °C. In its natural environment, the fast growing saprophyte plays an important role in recycling carbon and nitrogen sources by decaying organic matter. All data obtained so far suggest that the virulence of *A. fumigatus* is a multifactorial process. Putative virulence determinants include conidial pigments or intermediates, surface proteins, toxins, allergens, enzymes, cell wall components, nutrient sensing, and adaptation to hypoxia and iron depletion (Brakhage, 2005).

Transcriptomics and proteomics

Two-channel DNA microarrays and two-dimensional difference gel electrophoresis (DIGE) are common techniques for analyzing changes in the regulation of transcripts and proteins of *A. fumigatus*. The following sections describe processing, analysis and interrelationship of such data. An analysis workflow (Fig. 1) was applied, for example, for the analysis of the heat shock response of *A. fumigatus* (Albrecht et al., 2010a).

Pre-processing

Raw data from microarrays and 2D gels are images that have to be transformed into numerical data for further analysis. This transformation is preferentially done by using commercial software packages, such as GenePix (MDS Analytical Technologies) for two-channel microarrays or DeCyder (GE Healthcare) for DIGE data. These numerical data can contain several technical errors that have to be removed before detection of differentially regulated genes or proteins.

The first step in such pre-processing is imputation of missing values. Missing values can occur, for example, when the background noise in a certain part of the image is higher than the spot intensity in this region. The analysis of 2D gels is additionally hampered by the fact that migration of individual proteins can differ between gels. Consequently, spots are not always located at exactly the same position in each gel. Similar spots in different gels have to be aligned, which can result in unmatched protein spots. Many methods have already been developed for the imputing of missing data. One of the most popular approaches is the k-nearest neighbor (KNN) imputing, which was introduced for microarray data (Troyanskaya et al., 2001) and which was also utilized for proteomics data (e.g. by Jung et al., 2006). KNN imputes missing data of a certain spot by the weighted average of the k most similar spots, but it is limited to imputing data of spots with few missing values. The authors of Albrecht et al. (2010b) combined KNN with the imputation by a minimal value for spots that show many missing values in replicates of one condition. They compared this approach for the above-mentioned heat shock DIGE dataset with several other imputing methods and found that it performs best.

The next step following imputing of missing values is normalization. Normalization methods account for technical bias in the data. Many normalization methods exist for microarray as well as 2D gel data. Whenever more than one microarray or gel is investigated, normalization must be conducted within and between arrays or gels. For DIGE data, it was shown that the combination of variance stabilization [VSN (Huber et al., 2002), between gel normalization] and local regression [LOESS (Yang et al., 2002), within gel normalization], two methods developed for microarray data analysis, performs better that each method alone and also better than normalization of commercially available software (Albrecht et al., 2008b).

Results of normalization are ratios for all genes or proteins, describing their change in expression or abundance, respectively, between the test conditions and a reference condition. For finding statistically significant differences, filtering methods are applied. Filtering of genes or proteins using a fixed fold change threshold followed by Student's *t*-test or analysis of variance (ANOVA) is most often utilized for determining differential expression when working with data from *A. fumigatus*. However, ratios can be transformed to *Z*-scores to be more specific to the particular dataset (Quackenbush, 2002). Genes or proteins with *Z*-scores outside the range of [−1.96, 1.96] are differentially expressed at 95% confidence level. When combining the *Z*-scores with *t*-test or ANOVA, *p*-values have to be adjusted for multiple testing. Several methods exist to do such adjustment, the false discovery rate (FDR) correction of Benjamini and Hochberg (1995) being the most popular approach.

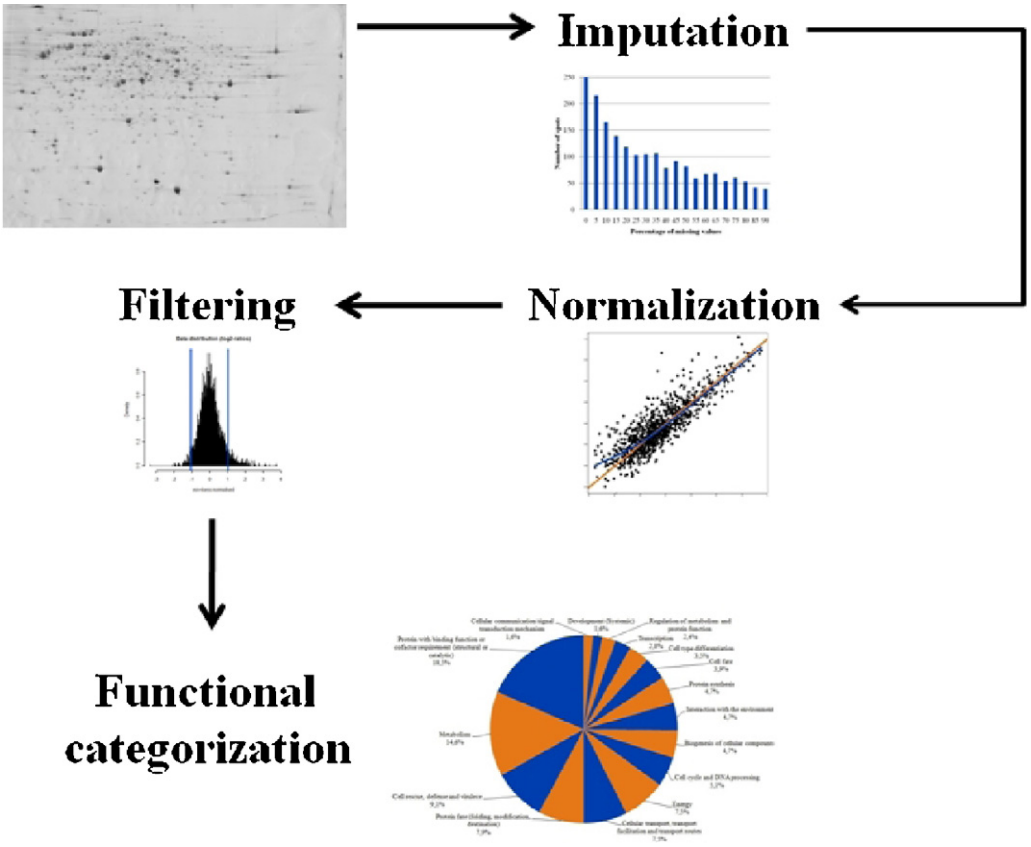

Fig. 1. Workflow for pre-processing and analysis of microarray and 2D gel data.

Functional categorization

Pre-processing of 2D gel- and microarray-based studies results in long lists of differentially regulated proteins or genes, the biological function of which has to be investigated. There are several possibilities to categorize genes/proteins according to their function. Three of them are of special importance for research on *A. fumigatus*: Kyoto Encyclopedia of Genes and Genomes (KEGG, Kanehisa and Goto, 2000), Gene Ontology (GO, Ashburner et al., 2000), and Functional Catalogue (FunCat, Ruepp et al., 2004). Another method, using protein domains, is described in Wolstencroft et al. (2006).

KEGG is a knowledge base for systematic analysis of gene functions, linking genomic information with other functional information. It is based on the concept of graphs for representation and manipulation of various objects from molecular to higher levels. All in all, KEGG pathway contains 327 functional categories with special emphasize on metabolism.

The GO project aims at standardizing the representation of genes and their products across species and databases. GO is structured in directed acyclic graphs using 3 subdomains: molecular function, biological process, and cellular localization. Nodes represent GO terms, and edges the relationships between the terms.

FunCat is an annotation scheme for the functional description of proteins from all living beings. It covers few manually annotated organisms like mouse or human. Additionally, all organisms are annotated, the sequences of which have been stored in NCBI RefSeq (Walter et al., 2009). FunCat exhibits 6 levels of increasing specificity and in total includes 1362 functional categories.

Besides the pure enumeration of genes/proteins in categories of whichever system, an enrichment analysis is advisable. Here, the number of genes or proteins from a dataset that fall into the encountered categories is compared to the number of genes or proteins from the whole genome/proteome that fall into the same categories. By using *p*-values, the most significant biological functions or processes in the dataset can be found.

Interrelationship

The whole-genome analysis of data from either transcriptome or proteome alone is still a reductionistic approach since both are strongly intertwined. Thus, to get a more holistic view, the use of information from both is advisable. Therefore, the authors of Albrecht et al. (2010a) obtained pre-processed transcriptomic data (Nierman et al., 2005) from the database ArrayExpress (Brazma et al., 2003). This dataset comprises normalized data of 1926 transcripts that were differentially regulated upon a temperature shift from 30 to 48 °C and, therefore, fitted their proteome dataset. For half of the 64 significant proteins, the respective transcripts were also differentially regulated.

A low correlation of transcripts and corresponding proteins has been reported in the literature (e.g. Greenbaum et al., 2003). For the temperature shift data of *A. fumigatus* (Albrecht et al., 2010a), relationships were examined by Pearson and Spearman correlation. Many proteins were negatively correlated with their respective transcript. Only few pairs showed good correlation. Fig. 2 shows some examples for high positive, low and high negative correlation of the time courses. Since a protein is most probably not regulated at exactly the same time point as its transcript, the authors introduced a time delay. Using this approach, all correlations became much better.

Low correlation between transcriptome and proteome was also found in a genome-wide analysis of heat shock in an *S. cerevisiae* mutant (Jones et al., 2004). This indicates that gene expression

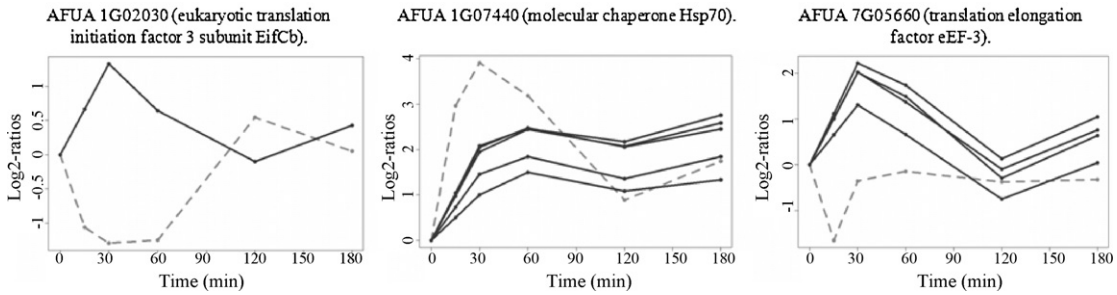

Fig. 2. Example time courses with high negative, high positive, and low correlation between transcript (dashed line) and protein (solid line). Several solid lines represent several spots for the same protein.

does not simply flow in a unidirectional manner from transcriptome to proteome. As cells have adopted elaborate regulatory mechanisms at the levels of transcription, post-transcription and post-translation, it is found that transcript and protein abundance measurements often are not concordant.

Network inference

Genes and proteins carry out their functions within a complex network of interactions, in particular gene regulatory networks, protein–protein interaction networks, signaling networks, and metabolic networks. There are different network model architectures possible (van Someren et al., 2002; Hecker et al., 2009). Due to the fact that infection is a dynamic process, dynamic modeling by differential equation systems is preferable as done to model the immune response towards bacterial infection (Guthke et al., 2005) using the network inference algorithm *NetGenerator* (Toepfer et al., 2007). The same approach and algorithm have been applied for modeling the transcriptomic response of *A. fumigatus* to a temperature shift from 30 to 48 °C (Guthke et al., 2007). The gene expression time series data were analyzed in 3 steps: First, clustering of time courses of differentially expressed genes; second, selection of cluster-representative genes; third, identification of the network model structure and parameters. The feature selection, i.e. the definition of cluster-representative genes, is the critical step within this approach of network modeling. In Guthke et al. (2007), this was done by text mining in gene annotations to identify strings overrepresented in the respective cluster. In an alternative approach, the feature selection was performed by gene enrichment analysis based on GO terms. The clusters for sub-strings of gene descriptions, GO-terms, and representative genes are shown in Table 1.

The expression profiles of the cluster-representative genes were simulated by a differential equation system, the structure and parameters of which were optimized. This procedure was repeated 10,000 times with randomly disturbed input data, and the most robust network structure was selected. The obtained consensus network structure is shown in Fig. 3. Genes coding for heat shock proteins are up-regulated transiently, whereas genes coding for ribosomal proteins are down-regulated transiently. According to the inferred network, the relaxation of the expression of genes coding for the heat shock proteins is caused by an influence of the

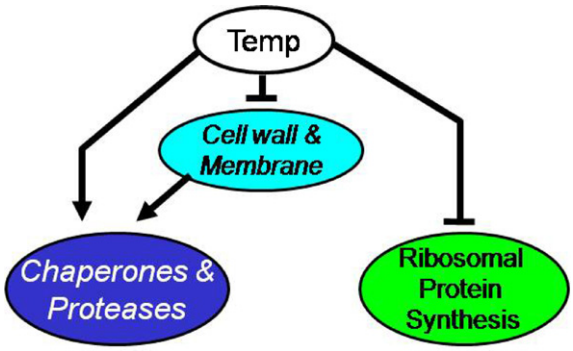

Fig. 3. Common generalized network structure inferred from gene expression profiles monitoring the heat shock.

down-regulated genes coding for membrane and cell wall proteins, such as erg11. Cell wall or membrane-bound proteins are known to be primary sensors of environmental stress proteins. The importance of these proteins in the virulence of pathogenic fungi and their role in the resistance to antifungal drugs is known. Interestingly, the gene erg11 (cytochrome P-450 sterol 14-alpha-demethylase) is a specific target of antifungal triazoles.

Data warehouse

Genome projects, DNA microarray, and proteomics technologies have increased the amount of generated data from biological systems. Standardized management of experimental data is a prerequisite for systems biological research, i.e. for collaboration between medicine, biology, mathematics, and computer science. Therefore, the data warehouse *OmniFung* (www.omnifung.hki-jena.de) was established to store microarray and 2D gel data from human-pathogenic fungi (Albrecht et al., 2007). *OmniFung* is used as central storage for data from many different research groups. By this, researchers working in related fields of fungal research have access to the work of others. This facilitates the sharing of results and enables researchers to reanalyze data of others with new techniques or under different points of view.

OmniFung presently contains data from 4 different fungal species: *A. fumigatus*, *Candida albicans*, and *C. glabrata* are important human-pathogenic fungi while *Aspergillus nidulans* has been

Table 1
Robust clusters of gene expression profiles of *A. fumigatus* after heat shock, their annotations and representatives.

Kinetics	String over-represented in the gene description	GO-term over-represented in the gene description	Cluster-representative genes	Generalized annotation
Increasing	'peroxi'	Peroxisomal part'	cat2, fadD35	Oxidative stress response
Decreasing	'sterol'	'Glucan metabolism'	erg11, matA	Cell wall and membrane
Maximum	'heat shock'	Protein folding	hsp30, clpB	Chaperones, proteases
Minimum	'ribosom'	Nuclear part	rpl3, nip7p	Ribosomal protein synthesis

included as model organism for filamentous fungi. Publicly available datasets in *OmniFung* are accessible without using a login and password. Non-public data are in different stages of development and are accessible to certain subsets of *OmniFung* users.

Each dataset in *OmniFung* contains at least one microarray or 2D gel. Each array or gel entry in the database includes an image and quantitative data of the spots in the image. *OmniFung* stores transcriptome and proteome data together with extensive experimental annotations, which follow MIAME (Brazma et al., 2001) and MIAPE (Taylor et al., 2007) guidelines. The parameters also serve as a quality measure of a project in *OmniFung*.

The aim of *OmniFung* is not only to store data, but also to provide a platform for analyzing them. Therefore, four data analysis tools are presently attached to the data warehouse: *DIGE analyzer* is the implementation of the pre-processing workflow described above. *Distance Scan* is a tool for prediction of potential functional combinations of transcription factor binding sites (TFBSs). *FungiFun* assigns functional annotations to fungal genes or proteins and conducts enrichment analysis as described above. *Survival Analysis* was designed to evaluate virulence studies conducted using murine, egg, or other animal infection models.

Images and movies

There are well established techniques to visualize fungal infections in hosts, such as confocal laser scanning microscopy and two-photon microscopy. The former is used in most microbiology laboratories, whereas the latter becomes increasingly important especially for immunologists (Niesner et al., 2008). The majority of produced data are images with fluorescently labeled immune cells and conidia. Two-photon microscopy has the advantage of imaging in living systems and a better spatial resolution especially in the three-dimensional case. The emphasis is placed on cell classification, cell counting, shape and extant measurements, cell–cell contacts, phagocytosis events, and, for movies, cell motility and velocity. However, the image analysis forms the bottleneck in these studies. Frequently, images are analyzed manually, including counting and tracking, which is very time consuming and error-prone (Genevesio et al., 2006). Automation of these processes is the goal of biological image analysis.

Many software packages are written for this purpose. An open source software is CellProfiler (Carpenter et al., 2006). It is built on modules, which can be sequentially applied. It follows a pixel-based approach, i.e. individual picture elements (pixel) are examined, classified, and potentially merged. CellProfiler was designed for common two-dimensional images and thus has limited support for time-lapse and three-dimensional images. Application of CellProfiler to fungal infections in human or mice has not been reported so far.

Another established commercial software with an entirely different approach is the Definiens Developer XD based on the Definines Cognition Network Technology®. In contrast to pixel-based approaches, Definiens does not simply identify the interesting objects, but all of the intermediate objects together with their interrelationships (context). In this way, a model is built, which stores all objects, sub-objects and their semantic relationships hierarchically. The contextual information contained in the network enables the automated extraction of information analogue to the way a human brain deciphers an image. The network is described and built via a set of rules which in turn are programmed using a high level script. An image analysis is designed in 3 main parts. The first step comprises image enhancement and noise reduction via different filters. The Gaussian filter and the median filter are the most used filtering algorithms in biological applications (Soille, 2003). The second and third step, segmentation and classification, are applied rather alternately than sequentially. Segmentation, a division of an image into constituent objects, forms the most vital and most difficult step in an image analysis task (Wählby et al., 2004). Multiresolution segmentation is a common method for dividing the background from the objects of interest (cells) as first step (Bouman and Liu, 1988). Afterwards, all cells have to be separated from each other. Using the interrelationships and already gained features, there are a lot of different possibilities to tackle this task, i.e., different fluorescently labeled objects could be divided and subsequently classified. Furthermore, a segmentation of the detected cells into cytoplasm and nucleus or cytoplasm and phagocytozed conidia is possible. Fig. 3 shows an example image. After a complete and satisfying image analysis, the resulting features, e.g. extent, shape, number of cell types, number of cells per type, cell velocities, and quantity of cell–cell contacts, are storable in different formats for modeling of the fungal infection with additional systems biology tools. The Definiens software is applied to screenings of different *A. fumigatus* mutants with respect to diverse phagocytosis behavior as well as cell-tracking of human immune cells, e.g. macrophages and neutrophiles, in time-lapse microscopy images (data not shown).

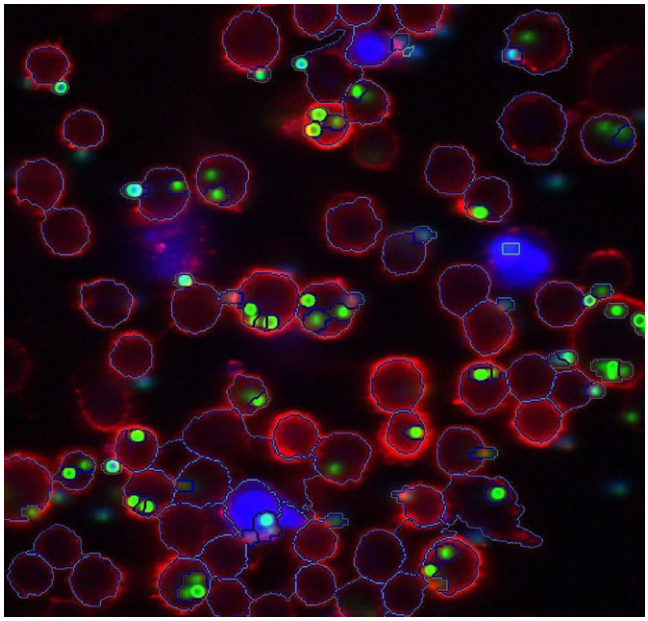

Fig. 4. Confocal laser microscope image. Red stained are macrophages, green are *Aspergillus fumigatus* conidia. Thin lines represent result of automatically segmented cells.

Conclusions and outlook

Recently, first steps towards a systems biology of *Aspergillus* species – in particular those of industrial interest – were made, including mathematical modeling in genomics, transcriptomics, proteomics, and metabolomics (Vongsangnak and Nielsen, 2010). Systems biology of *A. fumigatus* infection is still in its infancy. There are first systems biological studies in genomics, transcriptomics, and proteomics of *A. fumigatus*. Currently, mathematical modeling of networks in *A. fumigatus* is limited to small-scale systems with focused view. The next step will be the development of a large-scale genetic network modeling approach that: (a) integrates multiple distinct sources of measurement data (e.g. microarray, qRT-PCR, RNA-seq, proteome, protein–protein interaction, metabolome, microRNA, ChIP, imaging); (b) integrates multiple sources of functional data (Gene Ontology, literature mining) and sequence information (TFBS analysis).

Finally, systems biology of fungal pathogenicity has to include spatial aspects of the host–pathogen interaction. For that level, spatio-temporal data analysis is pivotal and has to be combined with genome-wide data analysis. This combination will be most conveniently realized within the agent-based modeling approach, where each biological entity is represented as an individual agent. The agents are equipped with dynamic properties characterizing their migration and interaction among each other. These properties are defined by molecular networks reconstructed from gene expression and proteome data (see Fig. 3) as well as by sets of rules and parameters extracted by context-based image analysis of spatio-temporal data (see Fig. 4), Next, spatio-temporal data are used for model validation. Finally, the dynamic model is applied to optimally design further experiments forming an iterative cycle of experimental and theoretical work. The strength of this agent-based modeling approach lies in the emergence of collective system properties that are not known a priori and, thus, provide this generic method with predictive power. The agent-based modeling approach is well-established in systems biology (Chavali et al., 2008). The increasing amount of diverse experimental data that are currently becoming available make it a promising tool to better understand the infection biology of *A. fumigatus* in the near future.

Ultimately, systems biology of fungal infection aims, first, at quantitative insights into key mechanisms underlying infection and immune responses, second, at developing reliable and predictive methods to increase efficiency and productivity for prediction of new therapeutic targets, of drug discovery, and, third, at development of vaccines and optimized strategies for individualized therapy (translational systems biology; Vodovotz et al., 2008). Future advances in understanding as well as prevention and therapy of infections caused by *A. fumigatus* are highly dependent on methodological advances and integration of the computational systems biology community with biologists and clinicians.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) Priority Program 1160 'Colonisation and infection by human-pathogenic fungi'. The authors thank the reviewers who helped to improve the manuscript.

References

Albrecht, D., Kniemeyer, O., Brakhage, A.A., Guthke, R., 2007. Integration of transcriptome and proteome data from human-pathogenic fungi by using a data warehouse. J. Integrative Bioinf. 4, 52.

Albrecht, D., Guthke, R., Kniemeyer, O., Brakhage, A.A., 2008a. Systems biology of human-pathogenic fungi. In: Daskalaki, A. (Ed.), Handbook of Research on Systems Biology Applications in Medicine, vol. 1. IGI Global, pp. 400–418.

Albrecht, D., Kniemeyer, O., Brakhage, A.A., Guthke, R., 2008b. Normalisation of 2D DIGE data – on the way to a Standard Operating Procedure. In: Küng, J., Schneider, K., Wagner, R. (Eds.), BIRD'08 2nd International Conference on Bioinformatics, Research and Development, Schriftenreihe Informatik 26. Trauner Verlag, Linz, pp. 55–64.

Albrecht, D., Guthke, R., Brakhage, A.A., Kniemeyer, O., 2010a. Integrative analysis of the heat shock response in *Aspergillus fumigatus*. BMC Genomics 11, 32.

Albrecht, D., Kniemeyer, O., Brakhage, A.A., Guthke, R., 2010b. Missing values in gel-based proteomics. Proteomics 6, 1202–1211.

Andersen, M.R., Nielsen, M.L., Nielsen, J., 2008. Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. Mol. Syst. Biol. 4, 178.

Andersen, M.R., Nielsen, J., 2009. Current status of systems biology in *Aspergilli*. Fungal Genet. Biol. 46 (Suppl. 1), S180–S190.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25, 25–29.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R Stat. Soc. Ser. B 57, 289–300.

Bouman, C., Liu, B., 1988. Segmentation of textured images using a multiple resolution approach. In: IEEE Int. Conf. Acoust., Speech, Signal Processing , pp. 1124–1127.

Brakhage, A.A., 2005. Systemic fungal infections caused by *Aspergillus* species: epidemiology, infection process and virulence determinants. Curr. Drug Targets 6, 875–886.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., Vingron, M., 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat. Genet. 29, 365–371.

Brazma, A., Parkinson, H., Sarkans, U., Shojatalab, M., Vilo, J., Abeygunawardena, N., Holloway, E., Kapushesky, M., Kemmeren, P., Lara, G.G., Oezcimen, A., Rocca-Serra, P., Sansone, S.A., 2003. ArrayExpress – a public repository for microarray gene expression data at the EBI. Nucleic Acids Res. 31, 68–71.

Chavali, A., Gianchandani, E., Tung, K., Lawrence, M., Peirce, S., Papin, J., 2008. Characterizing emergent properties of immunological systems with multi-cellular rule-based computational modeling. Trends Immunol. 29, 589–599.

Carpenter, A., Jones, T., Lamprecht, M., Clarke, C., Kang, I., Friman, O., Guertin, D., Chang, J., Lindquist, R., Moffat, J., Golland, P., Sabatini, D., 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. 7, R100.

Chan, E., 2006. Integrating transcriptomics and proteomics. Genom. Proteom. 6, 20–26.

Dagenais, T.R.T., Keller, N.P., 2009. Pathogenesis of *Aspergillus fumigatus* in invasive aspergillosis. Clin. Microbiol. Rev. 22, 447–465.

David, H., Hofmann, G., Oliveira, A.P., Jarmer, H., Nielsen, J., 2006. Metabolic network driven analysis of genome-wide transcription data from *Aspergillus nidulans*. Genome Biol. 7, R108.

Day, J., Schlesinger, L.S., Friedman, A., 2010. Tuberculosis research: going forward with a powerful "Translational Systems Biology" approach. Tuberculosis (Edinb.) 90, 7-8.

Franke, R., Müller, M., Wundrack, N., Gilles, E.D., Klamt, S., Kähne, T., Naumann, M., 2008. Host–pathogen systems biology: logical modelling of hepatocyte growth factor and *Helicobacter pylori* induced c-Met signal transduction. BMC Systems Biology 2, 4.

Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.J., Wortman, J.R., Batzoglou, S., Lee, S.I., Baştürkmen, M., Spevak, C.C., Clutterbuck, J., Kapitonov, V., Jurka, J., Scazzocchio, C., Farman, M., Butler, J., Purcell, S., Harris, S., Braus, G.H., Draht, O., Busch, S., D'Enfert, C., Bouchier, C., Goldman, G.H., Bell-Pedersen, D., Griffiths-Jones, S., Doonan, J.H., Yu, J., Vienken, K., Pain, A., Freitag, M., Selker, E.U., Archer, D.B., Peñalva, M.A., Oakley, B.R., Momany, M., Tanaka, T., Kumagai, T., Asai, K., Machida, M., Nierman, W.C., Denning, D.W., Caddick, M., Hynes, M., Paoletti, M., Fischer, R., Miller, B., Dyer, P., Sachs, M.S., Osmani, S.A., Birren, B.W., 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438 (7071), 1105–1115.

Genevesio, A., Liedl, T., Emiliani, V., Parak, W.J., Coppey-Moisan, M., Olivo-Marin, J.C., 2006. Multiple particle tracking in 3-d+t microscopy: method and application to tracking of endocytosed quantum dots. IEEE Trans Image Process 15, 1062–1070.

Greenbaum, D., Colangelo, C., Williams, K., Gerstein, M., 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol. 4, 117.

Guthke, R., Möller, U., Hoffmann, M., Thies, F., Töpfer, S., 2005. Dynamic network reconstruction from gene expression data applied to immune response during bacterial infection. Bioinformatics 21, 1626–1634.

Guthke, R., Kniemeyer, O., Albrecht, D., Brakhage, A.A., Moeller, U., 2007. Discovery of gene regulatory networks in *Aspergillus fumigatus*. Lect. Notes Bioinf. 4366, 22–41.

Harris, S.D., 2009. Special issue: physiology and systems biology of the fungal cell. Microbiology 155, 3797–3798.

Hecker, M., Lambeck, S., Toepfer, S., van Someren, E., Guthke, R., 2009. Gene regulatory network inference – data integration in dynamic models – a review. BioSystems 96, 86–103.

Huber, W., Heydebreck, v., Schultmann, A., Poustka, A.S., Vingron, A.M., 2002. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 18 (Suppl. 1), S96–S104.

Hummert, S., Hummert, C., Schröter, A., Hube, B., Schuster, S, 2010. Game theoretical modelling of survival strategies of *Candida albicans* inside macrophages. J. Theor. Biol. 264, 312–318.

Ideker, T., Galitski, T., Hood, L., 2001. A new approach to decoding life: systems biology. Annu. Rev. Genom. Hum. Genet. 2, 343–372.

Jansen, G., Lee, A.Y., Epp, E., Fredette, A., Surprenant, J., Harcus, D., Scott, M., Tan, E., Nishimura, T., Whiteway, M., Hallett, M., Thomas, D.Y., 2009. Chemogenomic profiling predicts antifungal synergies. Mol. Syst. Biol. 5, 338.

Jones, D.L., Petty, J., Hoyle, D.C., Hayes, A., Oliver, S.G., Riba-Garcia, I., Gaskell, S.J., Stateva, L., 2004. Genome-wide analysis of the effects of heat shock on a *Saccharomyces cerevisiae* mutant with a constitutively activated cAMP-dependent pathway. Comp. Funct. Genomics 5, 419–431.

Jung, K., Gannoun, A., Sitek, B., Apostolov, O., Schramm, A., Meyer, H.E., Stühler, K., Urfer, W., 2006. Statistical evaluation of methods for the analysis of dynamic protein expression data from a tumor study. REVSTAT 4, 67–80.

Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.

Klenk, M., Koczan, D., Guthke, R., Thiesen, H.J., Podbielski, A., Kreikemeyer, B., 2005. Global epithelial cell transcriptional responses reveal *Streptococcus pyogenes* Fas regulator activity association with bacterial aggressiveness. Cell Microbiol. 7, 1237–1250.

McNeil, M.M., Nash, S.L., Hajjeh, R.A., Phelan, M.A., Conn, L.A., Plikaytis, B.D., Warnocket, D.W., 2001. Trends in mortality due to invasive mycotic diseases in the United States, 1980–1997. Clin. Infect. Dis. 33,641–647.

Meyer, V., Arentshorst, M., Flitter, S.J., Nitsche, B.M., Kwon, M.J., Reynaga-Peña, C.G., Bartnicki-Garcia, S., van den Hondel, C.A., Ram, A.F., 2009. Reconstruction of signaling networks regulating fungal morphogenesis by transcriptomics. Eukaryot. Cell 8, 1677–1691.

Nierman, W.C., Pain, A., Anderson, M.J., Wortman, J.R., Kim, H.S., Arroyo, J., Berriman, M., Abe, K., Archer, D.B., Bermejo, C., Bennett, J., Bowyer, P., Chen, D., Collins, M., Coulsen, R., Davies, R., Dyer, P.S., Farman, M., Fedorova, N., Fedorova, N., Feldblyum, T.V., Fischer, R., Fosker, N., Fraser, A., García, J.L., García, M.J., Goble, A., Goldman, G.H., Gomi, K., Griffith-Jones, S., Gwilliam, R., Haas, B., Haas, H., Harris, D., Horiuchi, H., Huang, J., Humphray, S., Jiménez, J., Keller, N., Khouri, H., Kitamoto, K., Kobayashi, T., Konzack, S., Kulkarni, R., Kumagai, T., Lafon, A., Latgé, J.P., Li, W., Lord, A., Lu, C., Majoros, W.H., May, G.S., Miller, B.L., Mohamoud, Y., Molina, M., Monod, M., Mouyna, I., Mulligan, S., Murphy, L., O'Neil, S., Paulsen, I., Peñalva, M.A., Pertea, M., Price, C., Pritchard, B.L., Quail, M.A., Rabbinowitsch, E., Rawlins, N., Rajandream, M.A., Reichard, U., Renauld, H., Robson, G.D., Rodriguez de Córdoba, S., Rodríguez-Peña, J.M., Ronning, C.M., Rutter, S., Salzberg, S.L., Sanchez, M., Sánchez-Ferrero, J.C., Saunders, D., Seeger, K., Squares, R., Squares, S., Takeuchi, M., Tekaia, F., Turner, G., Vazquez de Aldana, C.R., Weidman, J., White, O., Woodward, J., Yu, J.H., Fraser, C., Galagan, J.E., Asai, K., Machida, M., Hall, N., Barrell, B., Denning, D.W., 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature 438 (7071), 1151–1156.

Niesner, R.A., Andresen, V., Gunzer, M., 2008. Intravital two-photon microscopy: focus on speed and time resolved imaging modalities. Immunol. Rev. 221, 7–25.

Oberhardt, M.A., Puchalka, J., Fryer, K.E., Dos Santos, V.A., Papin, J.A., 2008. Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. J. Bacteriol. 190, 2790–2803.

Pelkmans, L., 2009. Systems biology of virus infection in mammalian cells. Curr. Opin. Microbiol. 12, 429–431.

Perlroth, J., Choi, B., Spellberg, B., 2007. Nosocomial fungal infections: epidemiology, diagnosis, and treatment. Med. Mycol. 45, 321–346.

Quackenbush, J., 2002. Microarray data normalization and transformation. Nat. Genet. 32 (Suppl.), 496–501.

Raghunathan, A., Reed, J., Shin, S., Palsson, B., Daefler, S., 2009. Constraint-based analysis of metabolic capacity of *Salmonella typhimurium* during host–pathogen interaction. BMC Systems Biology 3, 38.

Raman, K., Yeturu, K., Chandra, N., 2008. targetTB: a target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis. BMC Systems Biology 2, 109.

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Güldener, U., Mannhaupt, G., Münsterkötter, M., Mewes, H.W., 2004. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res. 32, 5539–5545.

Sage, D., Neumann, F.R., Hediger, F., Gasser, S.M., Unser, M., 2005. Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics. IEEE Trans Image Process 14, 1372–1383.

Soille, P., 2003. Morphological Image Analysis. Springer-Verlag, Berlin, pp. 241–263.

Taylor, C.F., Paton, N.W., Lilley, K.S., Binz, P.A., Julian, R.K., Jones, A.R., Zhu, W., Apweiler, R., Aebersold, R., Deutsch, E.W., Dunn, M.J., Heck, A.J.R., Leitner, A., Macht, M., Mann, M., Martens, L., Neubert, T.A., Patterson, S.D., Ping, P., Seymour, S.L., Souda, P., Tsugita, A., Vandekerckhove, J., Vondriska, T.M., Whitelegge, J.P., Wilkins, M.R., Xenarios, I., Yates, J.R., Hermjakob, H., 2007. The minimum information about a proteomics experiment (MIAPE). Nat. Biotechnol. 25, 887–893.

Toepfer, S., Guthke, R., Driesch, D., Woetzel, D., Pfaff, M., 2007. The NetGenerator algorithm: reconstruction of gene regulatory networks. Lect. Notes Bioinf. 4366, 119–130.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibishirani, R., Botstein, D., Altman, R.B., 2001. Missing value estimation methods for DNA microarrays. Bioinformatics 17, 520–525.

van Someren, E.P., Wessels, L.F., Backer, E., Reinders, M.J., 2002. Genetic network modeling. Pharmacogenomics 3, 507–525.

Vodovotz, Y., Csete, M., Bartels, J., Chang, S., An, G., 2008. Translational systems biology of inflammation. PLoS Comput. Biol. 4 (4), e1000014.

Vongsangnak, W., Nielsen, J., 2010. Bioinformatics and systems biology of *Aspergillus*. In: Machida, M., Gomi, K. (Eds.), *Aspergillus*: Molecular Biology and Genomics. Caister Academic Press, Norfolk, pp. 61–84.

Wählby, C., Sintorn, I., Erlandsson, F., Borgefors, G., Bengtsson, E., 2004. Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. J. Microscop. 215, 67–76.

Walter, M.C., Rattei, T., Arnold, R., Güldener, U., Münsterkötter, M., Nenova, K., Kastenmüller, G., Tischler, P., Wölling, A., Volz, A., Pongratz, N., Jost, R., Mewes, H.W., Frishman, D., 2009. PEDANT covers all complete RefSeq genomes. Nucleic Acids Res. 37 (Database issue), D408–D411.

Wolstencroft, K., Lord, P., Tabernero, L., Brass, A., Stevens, R., 2006. Protein classification using ontology classification. Bioinformatics 22 (14), e530–e538.

Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., Speed, T.P., 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res. 30 (4), e15.