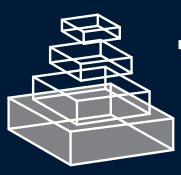
frontiers

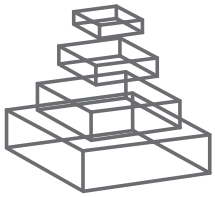
RESEARCH TOPICS

SYSTEMS BIOLOGY OF MICROBIAL INFECTION

Hosted by
Reinhard Guthke, Jörg Linde,
Marc Thilo Figge and Franziska Mech



frontiers in
MICROBIOLOGY



frontiers

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-060-7
DOI 10.3389/978-2-88919-060-7

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

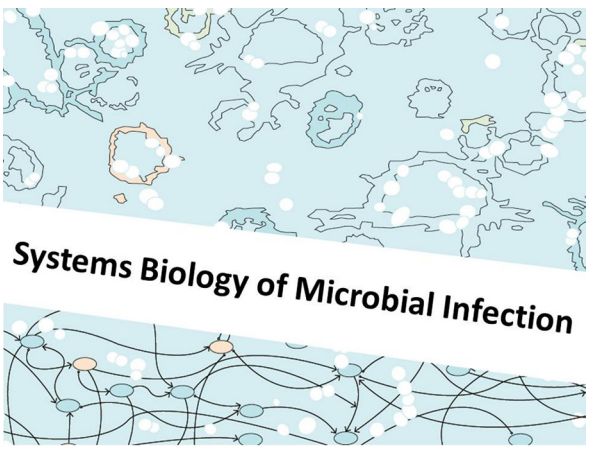
Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

SYSTEMS BIOLOGY OF MICROBIAL INFECTION

Hosted By:
Reinhard Guthke, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany
Jörg Linde, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany
Marc Thilo Figge, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany
Franziska Mech, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany



The systems biology of microbial infections aims at describing and analysing the confrontation of the host with bacterial and fungal pathogens. It intends to understand and to model the interaction of the host, in particular the immune system of humans or animals, with components of pathogens. This comprises experimental studies that provide spatio-temporal data from monitoring the response of host and pathogenic cells to perturbations or when interacting with each other, as well as the integrative analysis of genome-wide data from both the host and the pathogen. In perspective, the host-pathogen interaction should be described by a combination of spatio-temporal models with interacting molecular networks of the host and the pathogen. The aim is to unravel the main mechanisms of pathogenicity, to identify diagnostic biomarkers and potential drug targets, and to explore novel strategies for personalized therapy by computer simulations.

Some microorganisms are part of the normal microbial flora, existing either in a mutualistic or commensal relationship with the host. Microorganisms become pathogenic if they posses certain physiological characteristics and virulence determinants as well as capabilities for

immune evasion. Despite the different pathogenesis of infections, there are several common traits:

(1) Before infection, pathogens must be able to overcome (epithelial) barriers. The infection starts by adhesion and colonization and is followed by entering of the pathogen into the host through the mucosa or (injured) skin.
(2) Next, infection arises if the pathogen multiplies and overgrows the normal microbial flora, either at the place of entrance or in deeper tissue layers or organs.
(3) After the growth phase, the pathogen damages the host's cells, tissues and organs by producing toxins or destructive enzymes.

Thus, systems biology of microbial infection comprises all levels of the pathogen and the host's immune system. The investigation may start with the pathogen, its adhesion and colonization at the host, its interaction with host cell types e.g. epithelia cells, dendritic cells, macrophages, neutrophils, natural killer cells, etc. Because infection diseases are mainly found in patients with a weakened immune system, e.g. reduced activities of immune effector cells or defects in the epithelial barriers, systems biology of infection can also start with modelling of the immune defence including innate and adaptive immunity.

Systems biological studies comprise both experimental and theoretical approaches. The experimental studies may be dedicated to reveal the relevance of certain genes or proteins in the above mentioned processes on the side of the pathogen and/or the host by applying functional and biochemical analyses based on knock-out mutants and knock- down experiments.

At the theoretical, i.e. mathematical and computational, side systems biology of microbial infection comprises:

(1) modelling of molecular mechanisms of bacterial or fungal infections,
(2) modelling of non-protective and protective immune defences against microbial pathogens to generate information for possible immune therapy approaches,
(3) modelling of infection dynamics and identification of biomarkers for diagnosis and for individualized therapy,
(4) identifying essential virulence determinants and thereby predicting potential drug targets.

Table of Contents

frontiers in
MICROBIOLOGY

EDITORIAL
published: 12 September 2012
doi: 10.3389/fmicb.2012.00328

Systems biology of microbial infection

Reinhard Guthke*, Jörg Linde, Franziska Mech and Marc Thilo Figge

Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany
**Correspondence: reinhard.guthke@hki-jena.de*

Edited by:
Ian Marriott, University of North Carolina at Charlotte, USA

Reviewed by:
Ian Marriott, University of North Carolina at Charlotte, USA

This special Systems Biology of Microbial Infection research topic is dedicated to the dynamic modeling and model-driven analysis of microbial infection processes. It aims at describing and analyzing the confrontation of the host with bacterial and fungal pathogens, and at modeling and understanding the interactions of the immune system of humans and animals, with pathogens.

The early stages of the mathematical modeling of infectious diseases were initiated at the beginning of the twentieth century and were based on the population dynamics of epidemics, e.g., to design governmental vaccination policy (Ross, 1911, 1916).

Today, systems biology of microbial infection considers molecular, cellular, and even organismic levels of the pathogen and the host's immune system and is directed toward personalized medicine and theranostics.

The presented papers comprise theoretical and experimental studies. Some contributions present an integrative analysis of genome-wide data from both the host and the pathogen. Other contributions report on spatio-temporal determinations of host-pathogen interactions or the response of the host and pathogenic cells to defined perturbations that simulate infectious conditions.

This research topic is organized in two parts, with the first part consisting of four review articles and the second part consisting of seven original research papers.

PART I: REVIEW ARTICLES

Horn et al. (2012) review the systems biology of fungal infection, regarding virulence determinants of fungal pathogens, their adhesion and colonization of the host, their metabolism and gene expression patterns during infection, and their interaction with host cell types.

Haas (2012) reviews the regulation of iron uptake from the host and its storage within *Aspergillus fumigatus* cells as well as the regulation of iron homeostasis and transcriptional regulation under conditions of iron starvation within the pathogen.

Another example of adaptation to a pathogenic lifestyle to the conditions in the host is reviewed by Dandekar et al. (2012). The authors report on recent findings regarding the metabolic modeling of adaptation of *Salmonella* sp. to nutritional conditions within the host's vacuoles and membrane-bound compartments. Furthermore, the required metabolic conditions during *Salmonella* infection are compared with those for *Listeria* and *Legionella*.

Next, infection arises if the pathogen multiplies and overgrows the normal microbial flora, either at the place of entrance or in deeper tissue layers and organs. This is reviewed by Fuchs et al. (2012) focusing on the bacterial replication of metabolism in *Listeria*.

PART II: ORIGINAL RESEARCH PAPERS

The original research papers demonstrate different systems biology approaches to model: (i) host-pathogen interactions, (ii) molecular interaction networks of the pathogen, and (iii) host responses to the microbial infection.

HOST-PATHOGEN INTERACTION

In the past, most studies have focused on either the side of the pathogen or the host. Tierney et al. (2012) try to overcome this limitation using Next Generation Sequencing (RNA-Seq analysis) to monitor the gene expression of the host and the pathogen simultaneously. They analyze and model the interaction of mouse bone marrow-derived dendritic cells with the pathogenic yeast *Candida albicans*, and infer a gene regulatory network comprising genes and proteins of both host and pathogen.

Durmuş Tekir et al. (2012) study more than 23,000 protein–protein interactions that characterize the strategies of interaction of the human host with viral, bacterial, and some fungal pathogens utilizing the database pathogen-host-search-interaction-tool (PHISTO).

Binder et al. (2012) modeled the population dynamics of *Borrelia burgdorferi* using an ordinary differential equation approach taking into account the dissemination of bacteria to different tissues as well as cellular phagocytic activity during the innate immune response. By comparing different model scenarios with experimental data on infected mice, they aim at understanding why an almost cleared population of these bacteria can recover and reach a size that is even larger than the initial population before entering the chronic phase.

Applying an agent-based modeling approach that treats cells as discrete objects in space with dynamic properties, Tokarski et al. (2012) simulated the interactions of conidia of the human-pathogenic fungus *A. fumigatus* with migrating human neutrophils. Based on the analysis of time-lapse microscopy image data, different scenarios of phagocyte migration are explored and evaluated with regard to the clearance efficiency.

MOLECULAR INTERACTION NETWORK OF THE PATHOGEN

A genome-wide gene regulatory network has been inferred by Altwasser et al. (2012) for *C. albicans* using a compendium of microarray data by integration of prior knowledge from different

sources (e.g., transcription factor – target gene and protein–protein interaction data bases) as well as a large number of research papers screened for regulator/target gene interactions relations by an automatic text mining search.

HOST RESPONSE TO MICROBIAL INFECTION

Monitoring the immune response by flow cytometry for immune cells and cytokine production, Simon et al. (2012) identified and analyzed correlations and decision rules that formulated multidimensional and non-linear relations between the measured data on T-helper-cells and the cytokines GM-CSF, IFN-gamma, IL-2, IL-17, RANKL, and TNF-alpha. The authors found that, in particular, the production of GM-CSF and IL-17 turned out to be highly correlated.

Translational systems biology aims at the integration of omics data and clinical data from individual patients based on the investigation of animal models for therapeutical interventions. Lambeck et al. (2012) analyzed and compared the transcriptome from blood samples of a murine sepsis model together with that of patients from the pediatric intensive care unit.

SUMMARY

Systems biology of microbial infection aims at the development of testable mathematical and computational models of host-pathogen interactions that have predictive power for diagnosis and therapy by focusing on biomarkers and drug targets. This Research Topic presents a survey on the current state of this field. Some results were already presented at the first International Workshop on the "Systems Biology of Microbial Infection" that was held in 2011 in Jena, Germany, and will be organized in a biennial fashion by the Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute (HKI), Jena. Today, the field of systems biology of microbial infection is still in its infancy. However, we are convinced that the integration of "omics" approaches with image-based systems biology will strengthen and open new avenues to quantitative modeling of host-pathogen interactions in the future.

ACKNOWLEDGMENT

The authors acknowledge the excellence graduate school "Jena School for Microbial Communication (JSMC)" for support.

REFERENCES

Altwasser, R., Linde, J., Buyko, E., Hahn, U., and Guthke, R. (2012). Genome-wide scale-free network inference for *Candida albicans*. *Front. Microbiol.* 3:51. doi: 10.3389/fmicb.2012.00051

Binder, S. C., Telschow, A., and Meyer-Hermann, M. (2012). Population dynamics of *Borrelia burgdorferi* in Lyme disease. *Front. Microbiol.* 3:104. doi: 10.3389/fmicb.2012.00104

Dandekar, T., Fieselmann, A., Popp, J., and Hensel, M. (2012). *Salmonella enterica*: a surprisingly well-adapted intracellular lifestyle. *Front. Microbiol.* 3:164. doi: 10.3389/fmicb.2012.00164

Durmuş Tekir, S., Çakir, T., and Ülgen, K. Ö. (2012). Infection strategies of bacterial and viral pathogens through pathogen–human protein–protein interactions. *Front. Microbiol.* 3:46. doi: 10.3389/fmicb.2012.00046

Fuchs, T. M., Eisenreich, W., Kern, T., and Dandekar, T. (2012). Toward a systemic understanding of *Listeria monocytogenes* metabolism during infection. *Front. Microbiol.* 3:23. doi: 10.3389/fmicb.2012.00023

Haas, H. (2012). Iron – a key nexus in the virulence of *Aspergillus fumigatus*. *Front. Microbiol.* 3:28. doi: 10.3389/fmicb.2012.00028

Horn, F., Heinekamp, T., Kniemeyer, O., Pollmächer, J., Valiante, V., and Brakhage, A. A. (2012). Systems biology of fungal infection. *Front. Microbiol.* 3:108. doi: 10.3389/fmicb.2012.00108

Lambeck, S., Weber, M., Gonnert, F. A., Mrowka, R., and Bauer, M. (2012). Comparison of sepsis-induced transcriptomic changes in a murine model to clinical blood samples identifies common response patterns. *Front. Microbiol.* 3:284. doi: 10.3389/fmicb.2012.00284

Ross, R. (1911). *The Prevention of Malaria*. London: Murray.

Ross, R. (1916). An application of the theory of probabilities to the study of a priori pathometry. Part I. *Proc. R. Soc. Lond. A* 92, 204–230.

Simon, S., Guthke, R., Kamradt, T., and Frey, O. (2012). Multivariate analysis of flow cytometric data using decision trees. *Front. Microbiol.* 3:114. doi: 10.3389/fmicb.2012.00114

Tierney, L., Linde, J., Müller, S., Brunke, S., Molina, J. C., Hube, B., Schöck, U., Guthke, R., and Kuchler, K. (2012). An interspecies regulatory network inferred from simultaneous RNA-seq of *Candida albicans* invading innate immune cells. *Front. Microbiol.* 3:85. doi: 10.3389/fmicb.2012.00085

Tokarski, C., Hummert, S., Mech, F., Figge, M. T., Germerodt, S., Schroeter, A., and Schuster, S. (2012). Agent-based modeling approach of immune defense against spores of opportunistic human pathogenic fungi. *Front. Microbiol.* 3:129. doi: 10.3389/fmicb.2012.00129

Received: 20 July 2012; accepted: 24 August 2012; published online: 12 September 2012. Citation: Guthke R, Linde J, Mech F and Figge MT (2012) Systems biology of microbial infection. Front. Microbio. ***3****:328. doi: 10.3389/fmicb.2012.00328*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

REVIEW ARTICLE
published: 02 April 2012
doi: 10.3389/fmicb.2012.00108

Systems biology of fungal infection

Fabian Horn[1], Thorsten Heinekamp[2], Olaf Kniemeyer[2], Johannes Pollmächer[3], Vito Valiante[2] and Axel A. Brakhage[2,4]*

[1] *Systems Biology/Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany*
[2] *Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany*
[3] *Applied Systems Biology, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany*
[4] *Department of Microbiology and Molecular Biology, Institute of Microbiology, Friedrich Schiller University, Jena, Germany*

Edited by:
Jörg Linde, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Germany

Reviewed by:
Rene Alvarez, Naval Medical Research Unit-San Antonio, USA
Guoku Hu, Creighton University, USA
Scott G. Filler, Los Angeles Biomedical Research Institute, USA

***Correspondence:**
Axel A. Brakhage, Molecular and Applied Microbiology, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Beutenbergstrasse 11a, 07745 Jena, Germany.
e-mail: axel.brakhage@hki-jena.de

Elucidation of pathogenicity mechanisms of the most important human-pathogenic fungi, *Aspergillus fumigatus* and *Candida albicans*, has gained great interest in the light of the steadily increasing number of cases of invasive fungal infections. A key feature of these infections is the interaction of the different fungal morphotypes with epithelial and immune effector cells in the human host. Because of the high level of complexity, it is necessary to describe and understand invasive fungal infection by taking a systems biological approach, i.e., by a comprehensive quantitative analysis of the non-linear and selective interactions of a large number of functionally diverse, and frequently multifunctional, sets of elements, e.g., genes, proteins, metabolites, which produce coherent and emergent behaviors in time and space. The recent advances in systems biology will now make it possible to uncover the structure and dynamics of molecular and cellular cause-effect relationships within these pathogenic interactions. We review current efforts to integrate omics and image-based data of host-pathogen interactions into network and spatio-temporal models. The modeling will help to elucidate pathogenicity mechanisms and to identify diagnostic biomarkers and potential drug targets for therapy and could thus pave the way for novel intervention strategies based on novel antifungal drugs and cell therapy.

Keywords: *Aspergillus fumigatus, Candida albicans,* gene-regulatory network, network modeling, pathogenicity, pathogen-host interaction, spatio-temporal modeling, systems biology

1. BACKGROUND

1.1. HUMAN-PATHOGENIC FUNGI

It is estimated that the total number of fungal species exceeds 1.5 million (Hawksworth, 2001). However, only a small minority of approximately 100 species of fungi are associated with human diseases. Nevertheless, infections caused by fungal pathogens lead to a wide range of diseases including allergies, superficial infections, and invasive mycoses. The outcome of an infection with a human-pathogenic fungus often depends on the immune status of the host organism. Patients suffering from a weakened immune system are at high risk of developing a serious fungal infection. Continuous progress in medicine, e.g., in chemotherapy and organ or bone marrow transplantation, has led to an increasing number of patients with impaired immune status. In recent decades, the frequency of invasive fungal infections has increased steadily, resulting in considerable morbidity and mortality. In the USA, the incidence of sepsis caused by fungi has increased by more than 200% since 1991, whereas cases of bacterial sepsis have only increased moderate (Martin et al., 2003). Invasive mycoses are characterized by a high mortality rate. The increasing number of fungal infections has significantly contributed to health-related costs (Pfaller and Diekema, 2007).

The yeast *Candida albicans* and the filamentous fungus *Aspergillus fumigatus* are by far the most important causes of life-threatening invasive mycoses. Apart from *A. fumigatus*, around 10% of the more than 200 species of the genus *Aspergillus* are regarded as human pathogens or as having other adverse effects, e.g., *A. terreus*, *A. flavus*, and *A. niger* (Brakhage, 2005). The prevalence of *C. albicans* in clinical *Candida* samples is 50–70%, followed by infections with *Candida glabrata*, which is found in 20–25% of clinical *Candida* samples. Other pathogenic *Candida* species include *C. tropicalis*, *C. dubliniensis*, *C. krusei*, and *C. parapsilosis* (overview in Pfaller and Diekema, 2007). Another important human-pathogenic fungus of clinical relevance is the fungus *Cryptococcus neoformans*. The most common fungal infection among AIDS patients, cryptococcal meningitis, is caused by this basidiomycete. Furthermore, other fungal species, such as *Pneumocystis jiroveci*, *Zygomycetes*, *Fusarium* species, and *Scedosporium* species, have emerged as causal agents of invasive mycoses (Pfaller and Diekema, 2007).

Despite the different pathogenesis of infections caused by *C. albicans* and *A. fumigatus*, there are several common traits, particularly when the host response is considered: (i) the pathogens must be able to overcome epithelial barriers, (ii) innate immunity represents the major defense system, (iii) pathogenic fungi possess physiological characteristics, virulence determinants, and capabilities for immune evasion that make them aggressive pathogens, and (iv) invasive candidiasis and invasive aspergillosis are mainly found in patients with a weakened immune system either due to reduced activity of immune effector cells or defects in epithelial barriers.

Consequently, the aims of research on human-pathogenic fungi are (i) to unravel the pathogenic determinants specific to each

fungus, (ii) to investigate the distinct roles of epithelial barriers, the mechanisms of the innate immunity, and potential contributions of the adaptive immune system to the pathogenesis of fungal infections, and (iii) to elucidate the complex mechanisms of fungal infections and identify common principles of fungal pathogenesis.

1.2. PATHOBIOLOGY OF *ASPERGILLUS FUMIGATUS*: FROM ENVIRONMENTAL MICROORGANISM TO PATHOGEN

Within the last two decades, the filamentous fungus *A. fumigatus* has become one of the most important fungal pathogens. Conidia of this saprophytic fungus can be found almost everywhere, from the winds of the Sahara to the snow of the Antarctic. The most severe disease caused by *A. fumigatus* is invasive aspergillosis (IA), which occurs almost exclusively in immunocompromised patients (Brakhage, 2005). There is currently a lack of reliable diagnostic tools and effective treatment options for this condition, resulting in a high mortality rate despite therapy. Remarkably, *A. fumigatus* causes 90% of all systemic *Aspergillus* infections. This indicates that *A. fumigatus* possesses certain virulence determinants that favor this species becoming an opportunistic human pathogen. Because of their ubiquitous presence in the air, each person inhales several hundred *A. fumigatus* conidia daily. In immunosuppressed patients, the lung is the primary site of infection. In immunocompetent individuals, mucociliary clearance and phagocytic cells normally prevent the disease (Brakhage et al., 2010). However, there is a correlation between the degree of immunosuppression and the risk of contracting IA. Consequently, important risk factors include neutropenia, T cell depletion, CD34-selected stem cell products, corticosteroid therapy, and cytomegalovirus infections (Marr et al., 2002).

Since 2005, considerable progress has been made in the analysis of *A. fumigatus*. The genome sequence of *A. fumigatus* is available, and the transformation efficiency of the fungus was drastically increased by generation of *ku70* and *ku80* mutants of *A. fumigatus* (Nierman et al., 2005; da Silva Ferreira et al., 2006; Krappmann et al., 2006) making the generation of mutants by targeted gene deletion much easier. As a result of this improvement, the number of deletion mutants has increased from a handful, in the year 2000, to more than 400, today.

However, only a few virulence determinants of *A. fumigatus* have been characterized to date. These determinants include the siderophore-mediated iron uptake system (Schrettl et al., 2004) or the *pksP* gene, which is involved in the biosynthesis of the gray-green spore pigment (Langfelder et al., 1998; Thywißen et al., 2011; Volling et al., 2011). How these virulence determinants influence the infection is currently under investigation. DHN melanin was shown to inhibit both apoptosis and the acidification of conidia-containing phagolysosomes of macrophages (Thywißen et al., 2011; Volling et al., 2011). Because virulence is a multifactorial process, it can safely be expected that many more virulence-associated traits will be discovered, e.g., *A. fumigatus* is able to grow under hypoxic conditions. This ability is essential for pathogenicity (Willger et al., 2008). *A. fumigatus* also possesses immune-evasion mechanisms which reduce recognition, both by immune effector cells and the complement system (Behnsen et al., 2008, 2010; Aimanianda et al., 2009).

Innate immunity is of great importance in defense against *A. fumigatus*. Alveolar macrophages are the major resident cells of the lung alveoli and they phagocytose conidia. However, conidia have the ability to interfere with functions of the macrophages such as the maturation of phagolysosomes (Jahn et al., 2002; Ibrahim-Granet et al., 2003; Thywißen et al., 2011). Unphagocytosed conidia and outgrowing hyphae are killed by neutrophilic granulocytes, whose activity is essential for preventing IA (Feldmesser, 2006).

Also, the complement system appears to contribute to the defense mechanism (Moalli et al., 2010). Human pattern recognition receptors sensing fungal cell wall components include TLRs (Toll-like receptors), Galactin 3, DC-Sign (C-type lectin receptors), dendritic cell-specific intracellular adhesion molecule 3 (ICAM-3)-grabbing non-integrin, Dectin-1, SCARF1, and CD36. In line with the importance of these receptors for fungal recognition, a growing number of defined single nucleotide polymorphisms in the respective genes that appear to determine host susceptibility to *A. fumigatus* were identified (overview in Romani, 2011).

Furthermore, neutrophils possess recently discovered extracellular killing mechanisms: they degranulate, release DNA, and form neutrophil extracellular traps (NETs) both *in vitro* and *in vivo* against *A. fumigatus* (Bruns et al., 2010; McCormick et al., 2010). However, whether NET formation is detrimental for *A. fumigatus* overall is currently under investigation, and it remains unclear how neutrophils ultimately kill *A. fumigatus*. Reactive oxygen intermediates (ROI) most probably do not play a role as primary killing agents, but are required as signaling molecules (Lessing et al., 2007). Incubation of dendritic cells (DCs) with *A. fumigatus in vitro* resulted in the release of chemokine CXCL8, which attracts neutrophils (Gafa et al., 2007). Secretion of additional factors increased surface expression of CD11b and CD18 on neutrophils. Dectin-1, which is an important receptor on macrophages and neutrophils, is also expressed on the surface of immature DCs and is involved in the induction of a proinflammatory cytokine response (Mezger et al., 2008). DCs thus play an important role in defense against *A. fumigatus*. The pathogen recognition receptor Dectin-1 acts upstream of the Syk tyrosine kinase in response to an infection with *A. fumigatus*. Signaling via the Syk tyrosine kinase was recently found to be essential for the activation of NLRP3 inflammasome, another component of the innate immune system (Saïd-Sadier et al., 2010). Despite these findings, we are still in the early stages of understanding their role in organizing the immune defense mechanism.

Epithelial and endothelial cells in the lung can internalize conidia. It cannot, therefore, be excluded that these cells form sites of persistence and foci of infection (Latgé, 1999). The role of T cells has not yet been clarified. It appears most likely that they initiate the adaptive immune responses to *Aspergillus* species and directly influence the outcome of an infection (Dagenais and Keller, 2009). Phagocytosis of conidia by DCs leads to a protective Th1 response, whereas hyphal phagocytosis results in non-favorable Th2 responses and the generation of IL-10-producing CD4 cells (Romani, 2011).

1.3. PATHOBIOLOGY OF *CANDIDA ALBICANS*: FROM COMMENSAL TO PATHOGEN

Candida albicans normally exists as harmless commensal yeast on mucosal surfaces of the majority of the human population. Only under certain circumstances (imbalance of the normal microbial flora, immunosuppression, damage of tissue barriers), can *C. albicans* cause superficial (oral thrush in 90% of all untreated HIV patients, vaginal thrush in 75% of all women once in their lifetime) or life-threatening systemic infections (nosocomial candidiasis, candidemia; reviewed in Pfaller and Diekema, 2007; Martin et al., 2011). *C. albicans* is currently identified as the fourth most common blood isolate in US hospitals, accounting for around 10% of hospital-acquired bloodstream infections (Wisplinghoff et al., 2004). Cases of sepsis caused by this fungus lead to mortality rates of about 40% and are thus higher than observed for any bacterial sepsis (Gudlaugsson et al., 2003; Wisplinghoff et al., 2004; Picazo et al., 2008). Although *C. albicans* is an opportunistic pathogen, only a minority of cases (20%) of disseminated candidiasis occurs in patients with severe immunosuppression such as individuals with neutropenia, corticosteroid therapy, or HIV infection. Patients with severe illnesses who have prolonged periods of hospitalizations, a central venous catheter, gastrointestinal/cardiac surgery, or burns are at especially high risk of developing an invasive *Candida* infection (reviewed in Perlroth et al., 2007).

Almost all *C. albicans* infections are endogenous infections, caused by commensal strains of patients' own microflora. Despite numerous studies, it remains unclear how the transition from a harmless commensal to an aggressive pathogen is triggered. It seems certain that it is not only modifications of the microbial flora and host factors, but also specific attributes of the fungus that play an important role in this transition.

The ability to switch from yeast to pseudohyphal or hyphal growth is an important virulence trait for *C. albicans*. Several environmental conditions such as temperatures above 37˚C, pH values of 7.0 or higher, high exogenous CO_2 concentrations of more than 5%, or the presence of serum (Liu, 2002) trigger this transition. Numerous genes are involved in the regulation of the morphological switch, but molecular details are still poorly understood. The transcription factor Efg1p plays a central role in the control of morphogenesis (Stoldt et al., 1997; Doedt et al., 2004) and the loss of the *EFG1* gene led to mutants which were locked in the yeast form and showed reduced virulence in a murine model of candidiasis (Lo et al., 1997; Stoldt et al., 1997). In the regulation of the Efg1p pathway, cAMP-mediated signal transduction plays an essential role and mutants lacking the adenylyl cyclase Cdc35p do not form hyphae. As well as Efg1p, a large number of other transcription factors have also been reported to inhibit or trigger the yeast to hyphal transition, e.g., Efh1p, Mcm1p, Cph2p, and Tec1p (reviewed in Whiteway and Bachewich, 2007).

Besides the morphological plasticity, the ability of *C. albicans* to adhere to host cells and tissue and form biofilms is another important virulence factor. Amongst others, the ALS (agglutinin-like sequence) proteins are a well-studied group of proteins that form a family of peptide-binding proteins and which mediate adhesion (Salgado et al., 2011). They bind to extracellular matrix proteins, such as collagen, fibronectin and laminin (Als1p, Als3p, Als5p, Als6p, Als9p), endothelial and epithelial cells (Als1p, Als3p, Als5p), and also mediate cell-to-cell aggregation (Als5p) and iron acquisition (Als3p; Filler, 2006; Almeida et al., 2009). Hwp1p is another important adhesin, expressed only on hyphae, which binds tightly to oral epithelial cells and that is involved in biofilm formation (Nobile et al., 2006).

Secreted enzymes with proteolytic or lipolytic activity represent another group of proteins which contributes significantly to *C. albicans*' pathogenicity. A large proportion of the proteolytic activity is attributed to a multigene family of secreted aspartic proteinases (SAPs). Ten different SAPs have been described in *C. albicans*, eight of which are secreted extracellularly and two of which are anchored to the membrane via GPI linkage. Their contribution to the pathogenesis of *C. albicans* infections has been extensively investigated. *SAP1–SAP3* genes were considered to play a role in localized *C. albicans* infections and complement evasion, whereas *SAP4–SAP6* were postulated to play an important role in the pathogenesis of invasive candidiasis (Schaller et al., 2005; Gropp et al., 2009). However, in a recent study by Correia et al. (2010) the importance of *SAP1* to *SAP6* for virulence was reassessed in a murine model of candidemia. In contrast to previous findings, *SAP1* to *SAP6* were found to play no significant role in disseminated *C. albicans* infections. In addition to the virulence determinants described above, physiological fitness, in other words, high stress tolerance and metabolic flexibility, is another important factor that contributes to the pathogenicity of *C. albicans* (Brown et al., 2007; Fleck et al., 2011).

The fact that *C. albicans* is a diploid fungus and was long thought to be an obligate asexual organism (Alby et al., 2009), has long hindered the production of genetically defined mutants. In contrast to the well-studied yeast *Saccharomyces cerevisiae*, *C. albicans* does not have any natural DNA plasmids that could be used for transformation. In addition to this, *C. albicans* shows a non-standard codon usage and translates the CUG codon as serine instead of leucine (Lloyd and Sharp, 1992). Only after establishing protocols for targeted gene disruptions of both alleles, creating conditional null mutants based on tetracycline-regulatable systems, sequencing of the entire genome, establishment of a genome database, production of genome-wide microarrays, and production of reporter strains and other molecular tools as well as infection models, has *C. albicans* reached the status of a model organism for yeast infections (Theiss et al., 2002; Fradin et al., 2003; Jones et al., 2004; Braun et al., 2005; Samaranayake and Hanes, 2011; Szabo and MacCallum, 2011).

Due to this technical progress, interactions of *C. albicans* with host cells and the immune system have been the focus of many studies in recent years. Fungal recognition is the first step in the antifungal immune response and is mediated by pattern recognition receptors (PRR). The mannan cell wall component is recognized by the mannose receptor, the C-type lectin-like receptor Dectin-2, and the TLR4. Furthermore, TLR2 triggers an immune response by binding to phospholipomannans, as the Dectin-1 receptor does by binding to β-glucan (Netea et al., 2008). TLR2 and Dectin-1 regulate also the gene transcription of proinflammatory cytokines such as the pro-IL-1β. This interleukin is further processed into its active mature form via the NLRP3 inflammasome, a multiprotein complex. It has recently been shown that it is crucial for antifungal host defense (Gross et al., 2009; Hise

et al., 2009). After recognition, phagocytes, like macrophages or neutrophilic granulocytes, kill *C. albicans* cells by phagocytosis or secretion of antimicrobials. One mechanism discovered recently is the formation of neutrophil extracellular traps as mentioned above in the section on *A. fumigatus* (Urban et al., 2006). To link these interactions to the highly complex setting of clinical infections, complex *in vitro*, *ex vivo*, and *in vivo* infection models have been established and genome-wide transcriptional profiles, including direct *C. albicans* transcriptomes from patient samples, have been produced (Wilson et al., 2009; Cairns et al., 2010).

1.4. SYSTEMS BIOLOGY OF INFECTION

Due to their high complexity, it is conceivable that invasive infections caused by human-pathogenic fungi can be described and understood in a comprehensive manner by taking a systems biological approach. There are two complementary strategies in systems biology: (i) Starting from smaller, even minimal models capturing the essential and abstract interactions in the system under study and (ii) using experimental measurements such as large omics datasets in combination with large-scale models. Hybrid approaches, which integrate these top-down and bottom-up perspectives, contribute to our understanding of the multiple interdependencies of different hierarchical levels in biological systems (Forst, 2006). The global dynamics of a system can only be understood and quantified if the functionality of modular subsystems is elucidated, while considering the most important interactions on different levels in the biological system. With the aim of revealing molecular and cellular cause-effect relationships within the host-pathogen interaction in a non-ambiguous and efficient way, the setup of experiments and the design of experimental series can be optimized on the basis of established mathematical and computational models. Data exchange proceeds in an iterative cycle between model and experiment, with a constant refinement and validation of the models and the model-based planning of experiments (Ideker et al., 2001).

Current experimental and modeling techniques focus on specific perspectives at different scales. At present, experimental data from high-throughput experiments are increasingly and routinely used as the basis for mathematical modeling. This wealth of information has become available recently. Nevertheless, the history of the mathematical modeling of infectious diseases can be traced back to the eighteenth century, when today's basic concepts of evolution, genetics, and molecular biology were still unknown. Back in 1760, Daniel Bernoulli predicted the life expectancy of a population which has been immunized with cowpox (Bernoulli, 1760). In the early twentieth century, mathematical models were developed that mainly focused on the spread of diseases such as measles and malaria (Ross, 1911; Bailey, 1975). Most models would be related to today's research field of *population biology*, meaning that they dealt with fluctuations in population size under different modes of disease transmission. Since the second half of the twentieth century, these models have become more sophisticated. Amongst other concepts, they incorporated new aspects (e.g., population variables, May and Anderson, 1979; transmission rates, Real and Biek, 2007; pathogen life cycles, and host specificities, Woolhouse et al., 2001; Barrett et al., 2008) and extended the model to allow multi-level modeling (Roux and Aiello, 2005) or concepts of evolution (reviewed in Tong and Ng, 2011).

Naturally, each infection process is unique. Nevertheless, the modeling makes it possible to reveal fundamental similarities and differences in the underlying processes. The influence of single model parameters and their interdependency can thus be deduced. These parameters are also assessed if they serve as effective control options for the implementation of governmental public health risk management programs (Tong and Ng, 2011).

With the progress in molecular biology, infection biology, and biotechnology, it is now possible to study species-specific host-pathogen interactions at the molecular level in order to search directly for biomarkers with diagnostic potential and drug targets for novel therapeutic treatment strategies. The focus of research has diversified, resulting in specialized databases and research groups. Only early steps toward the computational systems biology of *A. fumigatus* and *C. albicans* have been made, including genome-scale data mining and mathematical modeling of infection processes by these fungi (reviewed for human-pathogenic fungi in Albrecht et al., 2008, 2011 and Rizzetto and Cavalieri, 2011).

2. DATA BASIS AND DATA ANALYSIS

The aim of understanding the complexity of host-pathogen interactions can be achieved by exploiting the increasing amount of experimental data, including high-throughput data and information available from public repositories as well as from biomolecular databases (see **Table 1**). Typically, experimental series comprise knock-down experiments along with global and specific screening using knockout mutants of pathogenic fungi. Procedures can be designed to analyze the complex structured data obtained

Table 1 | Bioinformatic resources of special interest for fungal systems biology.

Resource	Website	Description
AsperCyc	www.aspercyc.org	*Aspergillus* metabolic pathways
Aspergillus genome DB	www.aspgd.org	*Aspergillus* genomics
BROAD	www.broadinstitute.org	Genomics
Candida genome DB	www.candidagenome.org	*Candida* genomics
CFGB	http://cfgp.riceblast.snu.ac.kr	Comparative genomics platform
Ensembl	http://fungi.ensembl.org	Genomics
FunCatDB	www.helmholtz-muenchen.de/en/mips/projects/funcat	Gene-annotations
FungiDB	www.fungidb.org	Genomics
FungiFun	https://sbi.hki-jena.de/FungiFun/	Gene set enrichment analysis
JGI	www.jgi.doe.gov	Genomics
Omnifung	www.omnifung.hki-jena.de	Data warehouse for omics data
PhiBase	www.phibase.org	Database of virulence genes
SysMo-DB	www.sysmo-db.org	Collaborative platform

from these experimental series. Standardized pre-processing of the raw data and integrative analyses allow the identification of key regulators involved in pathogenicity. The management and integrative analysis of experimental data is challenging and constitutes a research field itself within bioinformatics (Albrecht et al., 2008). It was presented exemplary for genome, transcriptome, and proteome analysis of the heat shock response of *A. fumigatus* (Albrecht et al., 2010). Each technique applied generates different types of data, which have been acquired at the various levels of information. For each level, all of the available data types and their accompanying computational methods are usually referred to as "*omics*." New technologies, such as high-throughput sequencing, impose new challenges on efficient data storage, data retrieval, and statistical analysis. In order to handle the tremendous amount of data required for systems biology, the heterologous data has to be linked between existing databases. Automatic computer access has to be provided, and different user perspectives have to be taken into account. The standardization of biological terms (GO, FunCat, SBO, Ruepp et al., 2004; Arnaud et al., 2009; Courtot et al., 2011), data formats (SBML, Hucka et al., 2003), experimental metadata (MIBBI, Taylor et al., 2008), and operating procedures (Taverna, Hull et al., 2006) contribute to this objective.

Despite the wealth of information gained from analyzing high-throughput data, research into biological systems is mostly focused on a single data level. To overcome these limitations, data warehousing approaches (Omnifung, IntegromeDB, Albrecht et al., 2007; Kozhenkov et al., 2011) aim to integrate the different data layers. At the same time, most integration tools offer different visualization techniques as an additional key method for querying and understanding large datasets (Köhler et al., 2006; Smoot et al., 2011). Even though efforts have been made in this respect, an integrated view of different omics data levels is still a major challenge. Regulatory processes, different time scales, and non-linear processes in the biological system as well as technical limitations of omics technologies hamper the detection of causality or even correlation between different data layers (Albrecht et al., 2011).

2.1. "OMICS" BASED DATA

2.1.1. Genome

The importance of fungal-derived infections is reflected impressively by the number of fungal genomes that have been sequenced in recent years. The genome of the diploid *C. albicans* was published in 2004 (Jones et al., 2004; Braun et al., 2005), followed by the *C. neoformans* genome (Loftus et al., 2005) and by the *A. fumigatus* genome (Nierman et al., 2005). Since then, other *Candida* and *Aspergilli* genomes have been sequenced (Arnaud et al., 2007; Fedorova et al., 2008). Also, the first genomes of dermatophytes have recently been published (Burmester et al., 2011). **Table 2** lists all human-pathogenic fungi for which the full genome sequence is available. This list does not include genome projects currently in progress such as the sequencing of several *C. albicans* strains to evaluate *Candida* genome plasticity. New sequencing technologies now allow even greater numbers of genomes and transcriptomes to be deciphered, which will contribute to a better understanding of fungal pathogenesis. This information has been collected and released in suitable and easy to use web tools (see **Table 1**). For *Candida* sp., the main websites used are the CandidaDB (d'Enfert et al., 2005) and the *Candida* Genome Database (CGD; Arnaud et al., 2007). For *A. fumigatus*, the community normally refers to the Central *Aspergillus* Data REpository (CADRE, Gilsenan et al., 2012) and the new *Aspergillus* Genome Database (AspGD, Arnaud et al., 2012).

The information available in the different genome databases allowed us to apply studies of comparative genomics to human-pathogenic fungi focusing on evolutionary aspects of virulence genes. Studies of different *Candida* sp. highlighted that cell

Table 2 | List of human-pathogenic fungi of which the genomes have been sequenced.

Scientific classification	Species
ASCOMYCOTA	
Ascomycetes	*Ajellomyces dermatitidis, Aspergillus clavatus, Aspergillus fischeri, Aspergillus flavus, Aspergillus fumigatus, Aspergillus terreus, Blastomyces dermatitidis, Exophiala dermatitidis, Histoplasma capsulatum*
Euascomycetes	*Coccidioides immitis, Coccidioides posadasii, Penicillium marneffei*
Eurotiomycetes	*Arthroderma benhamiae, Arthroderma canis, Arthroderma gypseum, Arthroderma otae, Lacazia loboi, Paracoccidioides brasiliensis, Penicillium chrysogenum, Trichophyton equinum, Trichophyton rubrum, Trichophyton tonsurans, Trichophyton verrucosum*
Saccharomycetes	*Candida albicans, Candida dubliniensis, Candida glabrata, Candida guilliermondii, Candida lusitaniae, Candida parapsilosis, Candida tropicalis*
Sordariomycetes	*Chaetomium globosum, Fusarium oxysporum, Fusarium verticilloides, Nectria haematococca*
Pneumocystidomycetes	*Pneumocystis jirovecii*
BASIDIOMYCOTA	
Agaricomycetes	*Cryptococcus gattii, Cryptococcus neoformans*
Ustilaginomycetes	*Malassezia globosa, Malassezia restricta*
ZYGOMYCOTA	
Mucorales	*Rhizopus oryzae*
MICROSPORIDIA	
	Encephalitozoon cuniculi, Encephalitozoon intestinalis

wall-associated genes, important for host recognition and virulence of *Candida* sp., have been subject to gene duplications (Butler et al., 2009). Furthermore, the high number of lipases or GPI anchored proteins (normally clustered) is important for the virulence of *C. albicans* and *C. glabrata* (van het Hoog et al., 2007; Dujon, 2010). This phenomenon is known as gene family expansion and results in an increase of enzymatic power during infection processes or, where possible, easier rearrangement of the genome. In both cases, there is an increase in the competitiveness of the pathogens during infection processes (Moran et al., 2011).

Many clustered genes present in *Aspergillus* sp. encode proteins for secondary metabolite production, such as mycotoxins and antibiotics, and do not normally represent repetition of genes with similar enzymatic activity. A higher degree of genome rearrangement was observed in telomeric regions, where many of these clusters are located. The origins of these gene clusters have always been associated with the possibility of vertical gene transfer from other microbes. However, extensive comparative studies on *Aspergilli* genomes suggest that the presence of paralogs in the different species could be ascribed to gene duplication within the genus and subsequent translocation to telomere-proximal locations (Fedorova et al., 2008).

In addition to the importance for taxonomic studies, comparative genomics also highlighted peculiarities and similarities between different fungal species. One example was given by comparing signaling pathways in different pathogenic fungi. The mitogen activated protein kinase (MAPK) signaling pathways and the calcineurin pathway have been extensively studied in pathogenic fungi because of their involvement in pathogenesis. Global sequence similarity analysis indicated that, on the one hand, core structures involved in signaling are highly conserved while, on the other hand, upstream (e.g., receptors) and downstream factors (e.g., transcription factors) are far more species-specific (Rispail et al., 2009). Knowledge obtained in these studies can be used to identify suitable intraspecific or interspecific targets for therapeutic intervention.

2.1.2. Transcriptome

A genome, apart from its arrangement and complexity, can always be regarded as a static feature. In contrast to this, transcriptome analysis provides information about the dynamics of a genome's expression. Due to the availability of fungal genome data, it is now possible to design microarrays for genome-wide expression analysis.

For more than 10 years, scientists have had access to various *C. albicans* transcriptome studies. Many experiments have focused on studying gene expression in response to antifungal agents (e.g., azole derivatives, amphotericin B, echinocandins; Backer et al., 2001; Barker et al., 2004; Liu et al., 2005) but also at different stages of development and during biofilm formation (Doedt et al., 2004; Murillo et al., 2005). For *C. albicans*, already in 2005 a comparative gene expression analysis was published (Ihmels et al., 2005).

Aspergillus fumigatus' transcriptome history is comparatively recent. The first entire global transcriptome analysis was published together with the release of the first genome sequence (Nierman et al., 2005). After that, scientists had access to various transcriptome studies of developmental stages, during iron starvation, of biofilm formation, and response to antifungals (e.g., da Silva Ferreira et al., 2006; Schrettl et al., 2008; Bruns et al., 2010; Cagas et al., 2011; Jain et al., 2011). Fewer transcriptome studies have been carried out for *A. fumigatus* than for *C. albicans*. Furthermore, the comparability of transcriptome data is low. One reason for this is the development of different microarray platforms that made it harder to compare different gene expression data. In addition to this, the experimental design of transcriptome studies has been quite heterogenous with respect to media, strains, and general growth conditions. At present, we can only reliably compare different transcriptome data on the basis of genes that have high fold-changes in expression levels.

The main challenge in infection biology is to understand gene responses during infection. Such studies have been performed in several ways (co-culture of fungi and immune cells or direct tissue infection). The main problem remains the enrichment of the RNA from the pathogen prior to the hybridization step in order to avoid a decrease in data quality caused by cross-hybridization. This step is normally performed by separation of the different RNA species that can eventually be amplified to increase the nucleic acid quantity (Nygaard and Hovig, 2006). Previous studies identified genes involved in nutrient acquisition, oxidative stress response, and metal homeostasis, which are differentially regulated in *C. albicans* when co-cultured with macrophages and neutrophils (Lorenz and Fink, 2001; Lorenz et al., 2004; Wilson et al., 2009). Similar results have also been found in *A. fumigatus* when co-cultured with neutrophils and dendritic cells (Lessing et al., 2007; Sugui et al., 2008; Morton et al., 2011). A recent review described a comparison of transcriptome data obtained from pathogenic fungi during organ or tissue infections (Cairns et al., 2010). Many genes involved in primary metabolism appeared to be differentially expressed during infection in both human and plant pathogens. This data suggested that physiological reprogramming during infection remains relatively well conserved among various pathogens. On the other hand, these types of studies highlighted the limitations of hybridization-based techniques such as microarrays.

Recently, new techniques based on deep RNA sequencing have been introduced, i.e., the RNA-seq technique to analyze transcriptomes (Wang et al., 2009). Recent work on *C. albicans* and *A. fumigatus* indicated that with this technology it is possible to identify misannotated genes and differences in the level of low expressed genes which, for example, is relevant for many secondary metabolite gene clusters (Bruno et al., 2010; Gibbons et al., 2012). Furthermore, in infection biology it is important to study the gene expression profiles during infection, not only from the pathogen side, but also from the host side. To date, microarray analysis was limited in this respect, because separation of organism-specific RNA prior the hybridization is hard to achieve. Theoretically, RNA-seq analysis is capable of handling this problem. Technically, RNA from different species can be pooled, and then the data obtained can be separated during the analysis by aligning raw sequence data to different genomes. This approach seems possible because the sequences not matching a genome are normally discarded. The RNA-seq technique could potentially give us a way of monitoring gene expression profiles from the pathogen and the host simultaneously.

2.1.3. Proteome

To gain global insights into the biology of fungal pathogens and their interactions, genome-wide studies should focus not only on the transcript level, but also on the protein level. Proteins are the molecules that are catalytically active, build up the cellular structure, and mediate signal transduction and gene regulation. The release of the genome sequences of *C. albicans* and *A. fumigatus* paved the way for studies on the fungal proteome, i.e., the entire set of proteins which is synthesized and modified at a given time under defined conditions. Two dimensional-gel electrophoresis, a method invented in the mid-1970s, was the first technique used to study the presence of proteins on a global scale (O'Farrell, 1975; Klose and Kobalz, 1995). Meanwhile, mass spectrometry (MS)-based methods have become more and more popular. Here, the separation of tryptically digested peptides by liquid chromatography is coupled to mass spectrometry (reviewed in Aebersold and Mann, 2003). Despite the significant technical progress made in recent years in the field of MS-based proteomics, no technique currently available allows us to entirely profile the highly dynamic range and complexity of the protein set of a eukaryotic organism. However, a lot of knowledge about the proteome of *C. albicans* and *A. fumigatus* has already been obtained, which is summarized in several recent reviews (Rupp, 2004; Thomas et al., 2006; Kim et al., 2008; Kniemeyer and Brakhage, 2008; Andersen and Nielsen, 2009; Kniemeyer, 2011; Kniemeyer et al., 2011). Here, a brief overview will be given and current trends will be highlighted.

For *C. albicans*, the first studies on the extent of changes in cytoplasmic proteins during yeast-mycelial transition were conducted in the early 1980s, 25 years before the start of the post-genomic era in the field of fungal pathogenicity (Manning and Mitchell, 1980). More 2D-PAGE studies followed (Niimi et al., 1996), and differences in protein expression between the two morphotypes were also analyzed using MS-based techniques (Melanson et al., 2006). Several proteins that also play a role in the virulence of *C. albicans* were found to have increased expression levels in hyphal cells (reviewed in Kniemeyer and Brakhage, 2008). Recently, a study by Monteoliva et al. (2011) indicated that the primary metabolism undergoes a reorganization during morphotype-switching. Also, the composition of the *C. albicans* cell surface proteome undergoes changes during the yeast-mycelial transition. Several extraction techniques have been established in recent years to study this phenomenon (Pitarch et al., 2002; de Groot et al., 2004; Castillo et al., 2008; Hernáez et al., 2010). Heilmann et al. (2011) gave a first quantitative proteomic snapshot of the changes occurring in the cell wall proteome of *C. albicans* during the transition from yeast to hyphal cells. Several proteins were identified as indicators of hyphal growth, including the adhesin Als3p.

Due to the fact that *C. albicans* faces a multitude of diverse stresses, e.g., oxidative stress, higher temperatures, hypoxia, and low pH during infection, the proteomic response to these and other adverse conditions have been investigated by many groups. Kusch et al. (2007) and Yin et al. (2009) showed that the levels of many proteins with antioxidative functions were significantly increased during the oxidative stress response. Sosinska et al. (2008) and Sosinska et al. (2011) investigated the variability of the cell wall proteome during iron depletion, hypoxia, and at different pH values (pH 4 and 7).

Proteomics is also a suitable approach for gaining a deeper insight into the response of *C. albicans* toward antifungal compounds. Bruneau et al. (2003) characterized the proteome changes in *C. albicans* induced by triazoles (fluconazole and itraconazole) and an echinocandin-like lipopeptide (mulundocandin). The different modes of action of triazoles and echinocandins, two different classes of antifungal agents, were also reflected at the protein level. Similar results were obtained by Hoehamer et al. (2010), who additionally included the polyene amphotericin B in their study. Results of a recent study using a liquid chromatography–mass spectrometry (LC-MS) based approach suggested a cell wall destabilizing effect of the triazole fluconazole (Sorgo et al., 2011). Other studies addressed the mechanism of drug resistance by comparing the proteome of a drug-resistant mutant strain with a drug-susceptible wild-type strain (Hooshdaran et al., 2004; Yan et al., 2007).

A lot of progress has also been made in the field of *A. fumigatus* proteomics. In contrast to *C. albicans*, this pathogenic mold shows not only a mycelial growth form, but also produces spores for dispersal. Proteome maps of both morphotypes and the composition of the secretome have been established (Vödisch et al., 2009; Teutschbein et al., 2010; Cagas et al., 2011; Wartenberg et al., 2011). Also, the stress response to oxidative and heat stress was characterized (Lessing et al., 2007; Albrecht et al., 2010). Chaperones and antioxidative enzymes were produced under both conditions. Additionally, the results showed that the thioredoxin system seems to play an important role in maintaining the cellular redox balance for *A. fumigatus*. Recent findings revealed that *A. fumigatus* is exposed to oxygen-depleted microenvironments during infection (Grahl et al., 2011). The response to hypoxia was also studied at the protein level using an oxygen-controlled chemostat (Vödisch et al., 2011). Under hypoxic conditions, *A. fumigatus* cells developed a higher respiratory capacity, induced the synthesis of enzymes of the nitrosative stress response, and activated a secondary metabolite gene cluster (pseurotin A). Also, other growth conditions triggered the production of secondary metabolites with biological activity. When *A. fumigatus* attaches to surfaces, it can grow into a biofilm-like structure including the formation of an extracellular matrix. Under these conditions, higher levels of the immunosuppressive secondary metabolite gliotoxin were produced (Bruns et al., 2010).

Several groups profiled the *A. fumigatus* proteome in response to the antifungal compounds caspofungin (Cagas et al., 2011) and amphotericin B (Gautam et al., 2008). Amphotericin B influenced various metabolic processes including the ergosterol pathway, whereas caspofungin induced a strong increase in levels of ribosomal proteins.

The study of the interplay between fungal pathogens and human cells at the level of the proteome remains a challenging task due to the complexity and the limited number of methods available for the separation of fungal cells from the human effector cells. Several studies characterized the proteome of murine macrophages, which had been exposed to living or heat-inactivated *C. albicans* yeast cells (Shin et al., 2005; Martínez-Solano et al., 2006, 2009). To complement these studies, proteomic changes of *C. albicans* yeast cells due to macrophage confrontation were investigated and revealed an increase of the level of chaperones and

other stress-related proteins. No proteomic data is available on the interaction with other cells of the immune system or on the interaction of *A. fumigatus* with cells of the immune system. Nevertheless, proteomic technologies, such as targeted MS methods with high-performance instruments, have the potential to determine the mode of interaction between pathogenic fungi and their host. Targeted proteomic approaches are based on the selection of specific peptides of a protein for mass analysis. This process is termed *multiple reaction monitoring* (MRM). It allows the identification of very low-abundance proteins (Domon and Aebersold, 2010). However, this technique is still limited to the analysis of several hundred proteins in a single LC-MS/MS run, but technical progress in this field can be expected and may help to make a leap forward to the systematic investigation of host-pathogen interplay.

2.2. IMAGE-BASED DATA

Investigation and elucidation of the pathobiology of fungal infections strictly requires analysis of the interaction of the pathogen with the host immune effector cells (Brakhage et al., 2010). Image analysis is an indispensable tool for doing this. In combination with advances in computer performance and computing resources, the use of several imaging technologies led to the generation of large amounts of data (Behnsen et al., 2007; Hickey and Read, 2009; Brock, 2012) awaiting integration via a systems biological approach for analysis and interpretation. In general, image data can be obtained by different experimental approaches, e.g., microscopy, positron emission tomography (PET)/computer tomography (CT), or bioluminescence imaging. Due to the fact that the latter two techniques for monitoring fungal infections in living hosts are still in the early stages of development (Avet et al., 2009; Ibrahim-Granet et al., 2010), we focus here on recent studies using microscopy-generated image data. A good example of such data is represented by monitoring the interaction of labeled fungal cells, i.e., conidia, germlings, or hyphae, with immune effector cells using microscopy, especially fluorescence microscopy or confocal laser scanning microscopy (CLSM). Interaction of *A. fumigatus* and phagocytes based on fluorescence microscopy and manual image analysis was described, for example, by Ibrahim-Granet et al. (2003) and Jahn et al. (2002). In theses studies, interaction of *A. fumigatus* with phagocytes was monitored in detail, which revealed that *A. fumigatus* is able to inhibit acidification of phagolysosomes. More recently, Thywißen et al. (2011) were able to assign this ability to the presence of an active PksP, the polyketide synthase involved in dihydroxynaphthalene melanin biosynthesis. All of these studies were based on differentially labeled conidia, germlings, phagocytes, and their structures and compartments. It is even possible to monitor different pH values within distinct compartments of the phagocytes, allowing us to monitor phagocytosis rates in general and the fate of conidia after confrontation with immune cells in detail (Thywißen et al., 2011). In the future, a large number of mutants can be screened using this assay to identify further pathogen-derived components interfering with phagocytosis. Although phagocytosis of *A. fumigatus* by different phagocytes from both human and mouse has been analyzed in detail in several studies (Jahn et al., 2002; Ibrahim-Granet et al., 2003), only a limited amount of data is available with regard to direct observation of the phagocytosis process itself by live cell imaging. This applies to live cell imaging of phagocytosis of *Candida* species. Cell motility, however, is an essential requirement for the function of phagocytes. A detailed spatio-temporal analysis of the dynamics of the interaction of phagocytes with *A. fumigatus* and *C. albicans* using time-lapse microscopy and single-cell tracking was performed by Behnsen et al. (2007). In this study, the natural environments of different phagocytes were simulated by 2D liquid cultures and by generation of a 3D collagen environment. Live imaging showed that the interaction of phagocytes with *A. fumigatus* conidia or *C. albicans* cells in both 2D and 3D environments is a highly dynamic process that includes touching, dragging, and phagocytosis of fungal structures. Interestingly, the different immune cells, i.e., neutrophils, macrophages, and dendritic cells, exhibited different behavior with regard to the dependence on environmental dimensionality and as well as to the processing of *A. fumigatus* and *C. albicans*. Whereas neutrophils and alveolar macrophages efficiently phagocytosed or dragged *A. fumigatus* conidia in a 2D environment, their function was severely impaired in a 3D matrix. The opposite was found for processing of *C. albicans* cells. Phagocytosis was reduced in 2D environments, while in 3D environments most neutrophils internalized multiple yeast cells. These differences were also found in competitive assays, when both *C. albicans* and *A. fumigatus* were confronted with immune cells in the respective environment. Despite frequent touching of the other pathogen, neutrophils primarily incorporated *A. fumigatus* conidia in 2D and *C. albicans* yeast cells in a 3D environment. It is therefore conceivable that the activity and efficacy of the different phagocytes is best in the environment where a pathogen is naturally encountered.

Analysis of image data with regard to host-pathogen interaction was performed almost exclusively manually, with all of the inherent drawbacks and disadvantages. Manual data analysis is very time-consuming, error-prone, and last but not least, dependent on subjective criteria of the person performing the analysis. A first approach to automatize image analysis was performed by Mech et al. (2011). As a proof of principle, the interaction of *A. fumigatus* conidia with macrophages was monitored (see **Figure 1**). Data was collected by CLSM using cells labeled with different fluorescent dyes. The ruleset developed for processing microscopic raw data allows fully automated and context-based analysis of image data. By applying this method, discrimination between different cell types, i.e., phagocytes and conidia, was facilitated. Furthermore, cell counting based on discrimination between phagocytosed, adherent and non-adherent exterior conidia was performed. This is of particular importance since the different steps in conidia-macrophage interaction, i.e., recognition, adherence, ingestion, and intracellular processing of inhaled spores, define important pathogenesis-related processes. A prerequisite for automated image analysis, e.g., to determine phagocytosis rates, is the digitization of images. However, due to the fact that phagocytosis of *A. fumigatus* conidia by macrophages is a complex process, current image analysis tools that require precisely defined and homogeneous objects (overview in Shamir et al., 2010; Sysko and Davis, 2010) cannot be applied here. During co-incubation, conidia and macrophages tend to attach and form clusters. In addition to this, the relevant structures vary in their intensities as a result of different labeling efficacies and, last but not least,

the background differs considerably. These parameters were taken into consideration by the ruleset developed by Mech et al. (2011) and thereby allowed fully automated and context-based analysis of spatially resolved biological data, i.e., phagocytosis rates of *A. fumigatus* conidia. As this ruleset can be easily transferred to high-throughput microscopy measurements of other pathogens, e.g., *Candida* sp. or *Cryptococcus neoformans*, it will contribute to the further elucidation of host-pathogen interactions.

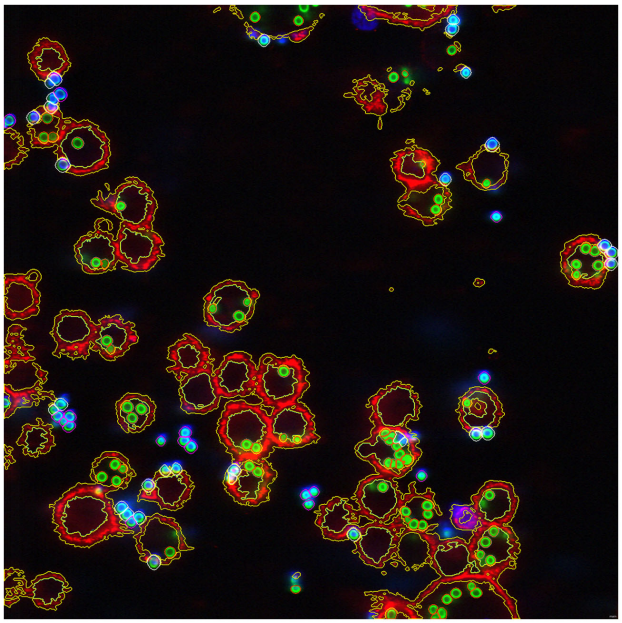


FIGURE 1 | Image of the phagocytosis assay showing all conidia and macrophages after segmentation and classification. Exterior non-adherent conidia outlined in magenta, adherent conidia outlined in white, interior conidia outlined in orange, and macrophages outlined in yellow. (© Taken from Mech et al., 2011.)

3. MODELING

The identification of pathogenic traits will increasingly be supported by mathematical and computational modeling and by the integration of experimental data and prior knowledge into complex models that describe the underlying virulence mechanisms. The generation of models and their subsequent application aim to support the optimal and standardized design of experiments and to generate and validate hypotheses. This allows that new knowledge is gained and predictions for novel strategies for diagnostics (biomarker design) and therapy (drug discovery, drug administration, therapy decision support) are made.

Systems biology of fungal infections describes and analyses various aspects of the confrontation of the host and its pathogen under defined conditions. The interactions and co-evolution of host and pathogen can be described with the help of evolutionary models (May and Anderson, 1979). Currently, molecular modeling of the host-pathogen interaction generally takes a reductionistic approach.

In general, models either describe a biological perspective on a single scale, or they span several different orders of magnitude (see **Figure 2**). Molecular mechanisms of host-pathogen interaction were identified with the help of statistical and integrated analysis of experimental data (see section 2). Single interactions represent only a reduced level of complexity, and a combination of several different mechanisms is capable of reflecting the global behavior at a cellular level. Such multi-systems interactions will become more complex when several species are considered. More recent approaches incorporate different biological levels to

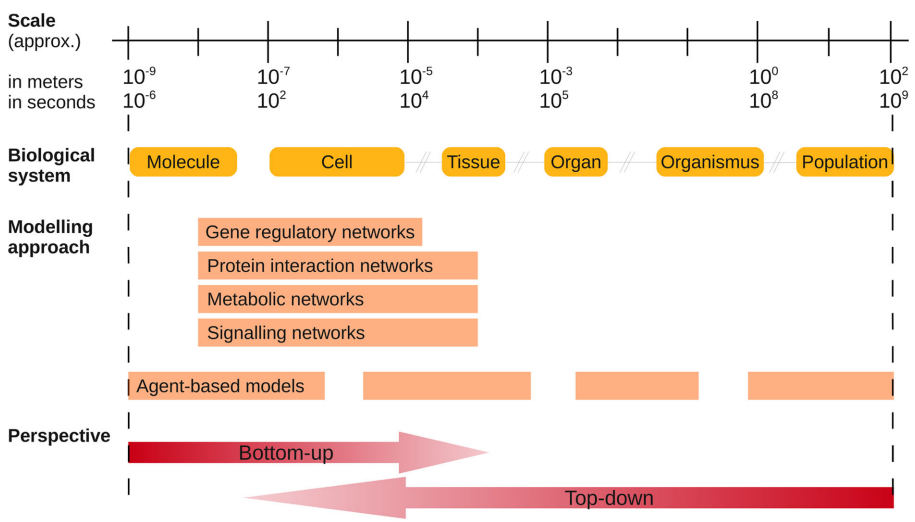


FIGURE 2 | Schematic diagram of different biological and modeling levels of systems biology of infectious diseases. Biological systems span several orders of magnitude. Mathematical methods, which are presented in this paper, focus on different biological levels of a fungal or similar infection. Modeling approaches can be applied to host and pathogen systems and their interplay. The flexibility and adaptability of agent-based modeling allows analysis at multiple levels without any restrictions with respect to the biological system. In order to generate new hypotheses, the advantages of bottom-up and top-down models are usually incorporated into the analysis. (Figure adapted from Forst, 2006.)

get an interrelated view of fungal infections. These multi-scale approaches need to integrate data from experiments on two or more scales (Walker and Southgate, 2009). The greatest challenge arises from the fact that an effective computational framework has to deal with the complexity of different length and time scales spanning several orders of magnitude. This leads to quantifiable differences of crucial constituents (e.g., one cell and its sizeable number of molecules) and to newly considered events (e.g., movement of the cell and the corresponding events at the molecular scale).

To date, only a few models of fungal infections of humans have been studied, which is mainly due to the complex modeling challenges and the previous lack of measurements of model parameters. In perspective, the host-pathogen interaction should be described by a combination of spatio-temporal models with interacting molecular network models.

3.1. NETWORK MODELING

Nodes in networks stand for interacting molecular entities (e.g., genes, proteins, metabolites) whose concentration or activity can be quantified by discrete or continuous variables. Edges, which stand for the relationships between the nodes, can be modeled in different ways, e.g., by directed or undirected edges and labeled by linear or non-linear functionality (Hecker et al., 2009). The cellular behavior of a system is usually represented by gene-regulatory networks, signaling networks, protein-protein interaction networks (PPI), and metabolic networks (see **Figure 3**). In addition to this, a confrontation between pathogen and host can be viewed as two interacting molecular networks, for example one within the host epithelial or immune cells and the other in the colonizing, persisting, or invading pathogens. Network models are capable of reflecting the non-linear dynamic behavior of the systems. Despite the fact that network visualization and handling is not scalable, its representation is intuitive and, as an example, there are ongoing projects which aim to standardize the graphical representation (BioPax, PSI-MI, SBGN, Hermjakob et al., 2004; Strömbäck and Lambrix, 2005; Le Novère et al., 2009). Basic networks only model the coordinated behavior of biological entities, and experimental data can be mapped to the network in order to confirm and annotate experimental results. The integration of general and specific knowledge transforms those "influence networks" into "mechanistic networks" of higher quality (Hecker et al., 2009; Santamaría et al., 2011) which themselves already represent a molecular interaction model that can be tested with the help of perturbation experiments. The dynamic nature of a system is partly reflected in the network topology. The investigation of the interactivity, the distribution and regulation of hubs, network motifs, and cross-talk of functional modules and reaction sets (e.g., coupled reaction sets and elementary flux patterns in metabolic networks) contribute to our understanding of the robustness and flexibility of a system (Barabási and Oltvai, 2004).

The major challenge of a systems biology-based approach to understanding the host-pathogen interaction arises from the robustness of the pathogenic system. The robustness originates from the network structure of the biological system that makes it unlikely that it will be possible to develop a single biomarker or drug against fungal infection. In the interests of clinical success, a system-oriented drug design with multiple antifungal strategies needs to affect sufficient points in the infection process (Kitano, 2007).

3.1.1. *Gene-regulatory networks*

Gene expression is mainly regulated by transcription factors and co-factors and additionally by post-transcriptional modification as well as mRNA and protein degradation. The reverse engineering of genome-wide interdependencies between these molecular entities relies on comprehensive datasets. Since there are only a few datasets on infectious processes available, one of the major tasks is to collect and process data and prior knowledge required for the development of novel parsimonious network models, describing essential fungus-host interactions. The inference process is mathematically challenging because the search space (number of possible gene regulations) increases exponentially with the number of nodes (genes). The modeling, on the other hand, relies on a small amount of data, which is usually obtained from microarray time series experiments. The experimental design is a trade-off that minimizes the cost and effort of the experiment, while ensuring that the data reliably reflects the underlying processes of the perturbation experiment. Many different network model architectures and inference methods have been proposed (reviewed in Hecker et al., 2009; Marbach et al., 2010). As part of the DREAM5-initiative[1] (Dialog for Reverse Engineering Assessments and Methods), gene expression data of the bacterial pathogen *Staphylococcus aureus* was presented to infer large-scale gene-regulatory networks. The inferred networks have been used to create a predictive community network of 1054 genes and 1688 edges, which need to be validated experimentally (Stolovitzky,

[1] http://wiki.c2b2.columbia.edu/dream/

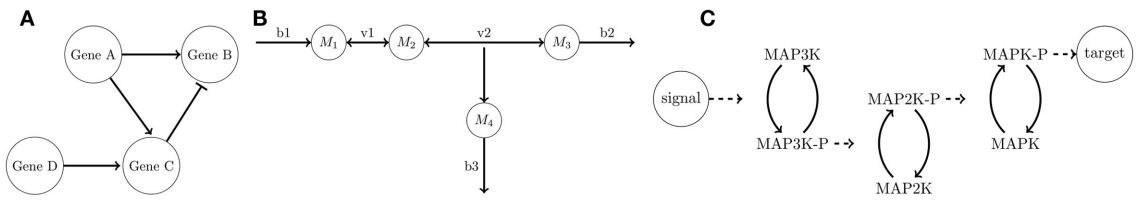


FIGURE 3 | Graphical representation of (A) gene-regulatory network, (B) metabolic network, and (C) signaling network. (A) The activation of genes is shown by arrows whereas repression is marked by bars. **(B)** Metabolites (M_i) are connected with internal (v_i) or external (b_i) metabolic fluxes. **(C)** A signal can lead to a phosphorylation ("-P") cascade of MAP kinases, and eventually the target is activated or repressed.

2011). The infection process of the fungi *C. albicans* and *A. fumigatus* were modeled using tools such as NetGenerator (Guthke et al., 2005; Toepfer et al., 2007), which are based on ordinary differential equation systems and linear regression methods that are capable of describing the dynamic behavior of systems, even on a global scale (Altwasser et al., 2012). The high dimensionality of the mathematical problem can be reduced at several stages in the network reconstruction process (Hecker et al., 2009). For example, Guthke et al. (2007) reconstructed the underlying gene-regulatory events when *A. fumigatus* was exposed to a temperature shift. In order to select features for the network modeling, expression profiles with similar significantly regulated time courses were clustered to functional entities. Cluster representatives, indicating main biological functions, were assigned to these entities using gene annotation (Priebe et al., 2011).

The dimensionality of the problem can be further reduced when known global characteristics of a gene-regulatory network are taken into account. The observed high modularity, hierarchical structure, and over-representation of network motifs can guide the optimization of the structure of the model. A widely used criterion is the network sparseness, meaning that each gene is regulated only by a small number of regulatory genes. This feature of gene networks was incorporated in studies by Linde et al. (2010, 2012), who analyzed transcriptomic data focused on the iron homeostasis of *C. albicans* during infection and of *A. fumigatus* after a change in available iron concentration. The authors also used prior knowledge in order to reduce the search space. Interactions between genes and transcription factors found in literature were used to build a network template, and this prior knowledge was integrated into the network reconstruction procedures with the help of a weighting function. Currently, prior knowledge about fungi is scarcely available in databases and needs to be extracted from literature or genome sequences, e.g., by predicting transcription factor binding sites (Fazius et al., 2011). The prior knowledge also forms the basis for the analysis of the validity of the inferred network and the performance of the applied reconstruction method.

Recently, the inference of gene-regulatory network was applied to predict host-pathogen interactions (Tierney et al., 2012). RNA-seq data from a *C. albicans* infection of *Mus musculus* were used to predict subnetworks which were subsequently combined into an interspecies network. This model contained predicted regulations between the two species during the infection process. The results were supported by experimental findings which, overall, demonstrated that network inference can be used to support the deciphering of complex infectious processes.

3.1.2. Protein-protein interaction network

The biological function of genes originates from encoded proteins which form molecular structures, catalyze metabolic reactions, and are involved in signaling and regulation. With the help of technologies such as yeast two-hybrid or tandem affinity purification combined with mass spectrometry, it is possible to identify interacting protein partners which form an "interactome" network. The interactome shifts the focus from the detection of single protein interactions to the decoding of the global organization of proteomes (Barabási and Oltvai, 2004). Protein networks are increasingly used to identify host immune molecules and pathogenic effector proteins associated with host infection and drug targets (reviewed in Ideker and Sharan, 2008). For fungal infections, the main challenge is the lack of experimental protein interaction data, which is required to (semi)automatically construct such a network. The bioinformatic approaches either rely on text-mining (Zhou and He, 2008; Rao et al., 2010), or the interaction is predicted on the basis of sequence analyses (Dyer et al., 2007; Skrabanek et al., 2008). For the modeling of infectious diseases, Dyer et al. (2008) mapped all available host-pathogen protein interactions to a single protein network. Many pathogens were found to target the same process and therefore supported the hypothesis that topological properties of the protein-protein interaction network can be used to identify pathogenetic traits (Mukhtar et al., 2011) and drug targets (Hase et al., 2009; Zhu et al., 2009). In addition to this, the functional annotation of single gene products is supported by the interaction context of proteins within the global protein-protein interaction network (Xu and Li, 2006; Sharan et al., 2007).

3.1.3. Signaling Networks

The cell's response to an internal and external stimulus is triggered by a signaling network whose regulation is the key control for cellular behavior. The understanding and modeling of this network holds great promise for the development of new therapeutic strategies and, consequently, many different modeling techniques have been developed (reviewed in Aldridge et al., 2006). The elements of the signal transduction networks help to identify the effect of positive or negative feedback loops (Blüthgen et al., 2009) and also shed light on the cross-talk between different signaling pathways (Borisov et al., 2009).

Successful application of signaling models depends on careful validation of the underlying data, especially since molecular signals are hard to measure experimentally. This is one reason why many applications in this research area concentrate on the perspective of the human host where, for example, the JAK/Stat pathway (Vera et al., 2011) and macrophage activation (Raza et al., 2008) have been modeled. Currently, the yeast *Saccharomyces cerevisiae* is a model organism for the analysis of fungal signaling networks (Waltermann and Klipp, 2010). The modeling of fungal pathogen signaling is still in its early stages. Efforts are being made to reconstruct the underlying signaling network, e.g., for *C. albicans*, with the help of sequence analysis and molecular biological experiments (Rispail et al., 2009). Additionally, the effects of host-pathogen interaction on the signaling pathways of the host are the subject of several studies (reviewed in Brodsky and Medzhitov, 2009; Hajishengallis and Lambris, 2011). For example, Franke et al. (2008) modeled the c-Met signaling network of hepatocytes after infection with the pathogenic bacterium *Helicobacter pylori* and predicted the effects of gene knock-outs, which were subsequently confirmed experimentally.

3.1.4. Metabolic networks

Supported by the increasing number of sequenced fungal genomes, the modeling of host-pathogen interactions with the help of genome-scale metabolic networks is feasible. The functionality of thousands of genes can be associated with a set of metabolic

reactions, because they either encode enzymes or they regulate related reactions (Cavalieri and Filippo, 2005). The steady state of metabolic fluxes through a metabolic network structure can be understood as a phenotypic state of an organism. For example, the metabolic capacity of the pathogen needs to be able to produce a variety of secretory metabolites and proteins which correspond to the pathogenic traits of the species and which are interesting drug targets (Fang et al., 2009). On the other hand, fungi also rely on the uptake of essential minerals which are bound to storage proteins in the host, e.g., iron bound to hemoglobin or ferritin during *C. albicans* infection (Almeida et al., 2009). The intertwined nature of the host-pathogen interaction can thus be modeled by two interacting metabolic networks (Raghunathan et al., 2009).

Despite the increasing number of sequencing projects for pathogenic fungi (see **Table 2**), the number of reconstructed networks is limited to species which are in the focus of metabolic engineering. To the best of our knowledge, only two large-scale, manually curated metabolic network reconstructions of fungal pathogens exist, namely for *A. fumigatus* (Tuckwell et al., 2011) and *C. albicans*[2]. Despite the recent efforts made at automation, a high quality network reconstruction process needs manual curation, a step which acts as the bottleneck in this modeling technique (Pitkänen et al., 2010). During the curation of specific metabolic pathways, genes that encode missing enzymes are annotated with the help of bioinformatic tools or experimental validation (Pitkänen et al., 2010). The reconstruction process itself thus already contributes to the elucidation of the underlying molecular processes.

The mathematical modeling and simulation of metabolic networks allow us to address questions such as (i) the influence of single enzymes (and corresponding genes) within the network, (ii) the search for invariant steady states, (iii) the prediction of elementary flux modes, and even (iv) the consideration of different optimization strategies for the organism (Ruppin et al., 2010). With these objectives in mind, specific methods and concepts of metabolic modeling, which do not rely on mostly unknown specific kinetic parameters, can be applied to fungi.

Nevertheless, it is promising to compare metabolic networks of closely related pathogenic and non-pathogenic species in order to understand the physical features of pathogenicity (Lee et al., 2009). Recently, Oberhardt et al. (2011) examined differences and similarities between a pathogenic and non-pathogenic *Pseudomonas* species. Prior to analysis, the authors reconciled both metabolic networks in order to minimize the influence of different reconstruction approaches on the results. The subsequent analysis confirmed the multifactorial aspect of pathogenicity, but differences in the flexibility of sulfur related pathways were also found.

3.2. SPATIO-TEMPORAL MODELING

In contrast to the network models discussed in section 3.1, which are based on "omics"-data that does not contain any spatial information, spatio-temporal models are developed to include this type of information. For example, modeling the affinity maturation of antibodies in germinal centers in response to the recognition of antigens has been successfully performed in recent years in joint experimental and theoretical studies (Figge, 2005; Figge et al., 2008; Garin et al., 2010). Similarly, this modeling approach can be applied to the field of fungal infections to model the innate immune response (see **Figure 4**). With the application of special imaging techniques to generate spatially and spatio-temporally resolved data and their automated analysis (see section 2.2), the way has been paved to model, simulate, and study interactions between the host and pathogenic fungi from this point of view.

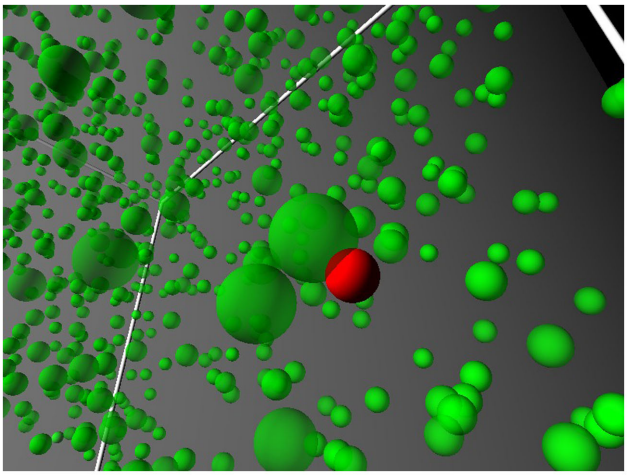


FIGURE 4 | Schematic illustration of a continuous 3D spatial environment *in silico* within an agent-based model. Spherical objects represent cells, the different colors depict different cell types.

Two common approaches to model spatio-temporally resolved systems are *partial differential equations* (PDEs) and *agent-based models* (ABMs). The former treats systemic constituents (e.g., molecules or cells) as concentrations (or populations), whereas the latter deals with them as discrete objects. When it comes to the number of constituents and their quantities, scalability is one of the computational strengths of PDEs. Apart from this advantage, crucial limitations such as the identification of single objects, their specific local interactions and their particular internal states exist (Van Dyke Parunak et al., 1998). ABMs can cope with all of the constraints mentioned above, but come at the price of more expensive computer resources in terms of computer memory and computing time (Chavali et al., 2008).

Each modeling approach has its own right to exist. ABM systems proved to have the necessary properties when higher levels of granularity for single constituents are considered. PDEs are adequate when the general behavior, such as Brownian motion, for a whole population of constituents is considered (Guo and Tay, 2005). Another criterion for making the right choice is the amount of data and knowledge available for the scale under consideration. Cells can be differentiated in various states. They may either move in a straight line or show random migration behavior and interact individually, depending on the interaction partners involved. These characteristics favor the choice of an ABM. However, modeling molecules at the level of individual molecules also has several advantages. Nevertheless, describing innate immune responses against fungal infections at all levels within ABMs renders the

[2]http://www.candidagenome.org and commercially available at Insilico Biotechnology AG and ERGO

multi-scale system intractable because of the overwhelmingly large number of cells and molecules. A promising approach toward modeling multi-scale systems was proposed by Guo et al. (2008), where, within a hybrid model, cells are described in continuous space as individual agents by an ABM approach, while molecular interaction is described by PDEs via reaction-diffusion equations for molecular populations. Such a hybrid multi-scale approach has the potential to model fungal infections since it is adaptable and extensible, thus providing the required flexibility.

At the beginning of the modeling process of this multi-scale system, one major issue is the definition of the necessary features of cells at the cellular scale. Examples of macroscopic characteristics are the cellular migration type as well as their morphology and interaction with other cells. A major source of these essential parameters are images and time-lapse videos, which are recorded during microscopy experiments. With the analysis of single frames, cellular sizes, morphological properties, and population quantities can be computationally extracted in an automated fashion (see **Figure 1**). An automated analysis of the interaction between macrophages and *A. fumigatus* was recently presented by Mech et al. (2011). Furthermore, the analysis of cell tracks extracted from videos plays an important role in the calculation of characteristic properties of the cellular movement such as speed, motility coefficient, and diffusion constant. The intercellular interactions can be defined at both the cellular and molecular scale. The initial setup of the interaction model can be arranged by implementing simple interaction rules of cellular behavior based on phenomenological knowledge. An extension toward a complex rule-system is also conceivable as well as the implementation of specific strategic behavior of the pathogenic fungi, as recently unraveled by Hummert et al. (2010), by applying methods of game theory. It would also be conceivable to include cell-cell interactions by considering interacting molecules at the cellular surface and their subsequent impact on internal and external signaling. These processes, taking place at the molecular level, require the data basis that underlies the network models (see section 3.1) and can be integrated together with the microscopy data basis into one and the same multi-scale model.

In addition to the integration of data, there are also other challenging tasks, such as the estimation of unknown parameters and the establishment of efficient algorithms and adequate data structures, in order to make simulations computational feasible. The former issue is strongly related to the generation of hypotheses. The efficiency of the applied algorithms needs to be verified experimentally in further studies. We are still at the beginning of the systems biology cycle regarding the spatio-temporal modeling of fungal infections, but the first promising steps have successfully been made.

4. OUTLOOK

Although great advances have been made in the understanding of molecular and cellular mechanisms of fungal infections, they do not currently provide a genome-wide view on the pathogenic processes of both fungal pathogens and hosts. New technologies such as RNA-seq, single-cell measurements, and PET/CT imaging open up new opportunities to unravel molecular and cellular mechanisms in greater detail and complexity. However, integrative analysis of high-throughput and spatio-temporal data on several molecular and cellular levels is increasingly becoming the limiting step when identifying molecular key regulators and mechanisms involved in fungal pathogenicity. Efforts to standardize data management and annotation promise advances for computational approaches to the mathematical modeling of host-pathogen interactions. Bioinformatic tools also need to be further adapted to the specific biological context of fungi. The current focus of systems biology is on the application of network and agent-based modeling techniques to genome-wide and spatio-temporally resolved dynamic systems thus exploiting the full range of high-throughput and image data available. In order to fully facilitate existing computational methods, it is necessary to deepen our understanding of molecular mechanisms thus improving gene annotation and establishing wider knowledge bases.

Genomic knowledge, coupled with high-throughput gene knock-out methodologies, are already advancing. The availability of mutant strains has grown rapidly over the last 10 years (McCluskey et al., 2010). We now have the possibility to screen more than 3,000 *C. albicans* single deletion mutants, but we still lack complete *A. fumigatus* or *C. albicans* knock-out libraries. At present, many high-throughput investigations are performed using complete gene knock-out libraries from *S. cerevisiae* or *Neurospora crassa*. This approach can be exploited to discover wide spectra of antimycotic drugs, but is not helpful for virulence studies. However, complete sets of viable genome-wide mutants for many pathogenic fungi are expected to become available in the next few years. In parallel, many efforts are being made to create high-throughput screening infection model systems. Recently, alternative systems using insects, nematodes, or embryonated eggs have been used (Ferrandon et al., 2007; Mylonakis, 2008; Moy et al., 2009; Jacobsen et al., 2010).

In the meantime, the scientific community continues to investigate the role of innate immunity during fungal infections. It has yet to be explored how neutrophil granulocytes actually kill outgrowing hyphae of *A. fumigatus*. The cellular mechanisms can be explored with the help of images which are a promising data source for driving the systems biology cycle. *Image-based Systems Biology* focuses on spatial properties such as cellular morphology or the mechanical features of interactions, and allows further insights which contribute to a better understanding of biological systems.

In clinical practice, differential diagnosis of infectious diseases and sepsis is based primarily on clinical criteria. These criteria lack the required sensitivity and specificity (Alberti et al., 2003) making the identification of novel biomarkers essential to identify the pathogen causing the infection on time. In clinical decision-making, fast and reliable diagnosis of specific pathogens is required. Theranostics (Chen, 2011) is a novel concept of combining novel platforms and technologies in clinical diagnostics and therapy such as image-guided therapy by, e.g., PET (Walther et al., 2011) or Raman spectroscopy (Neugebauer et al., 2006). There is currently a great deal of effort being made to push commercialization and collaboration strategies to establish the so-called companion diagnostics (CDx) marketplace. So-called *Translational Systems Biology* describes the clinical application of mathematical or computational models by enhancing the understanding of the complex dynamics of biomedical processes in an integrative, genome-wide way (Vodovotz et al., 2008). It integrates the scientific background of these trends and is based on current

experimental findings. Apart from integrating different "omics" data levels, models should not exclusively focus on the host's and the pathogen's aspects of the infection process, but also on the interaction between both biological systems. This interconnected perspective supports the elucidation of mechanisms underlying the complex process of fungal pathogenicity. Translational Systems Biology is primarily directed at drug target identification and validation as well as rational drug design, supported by analysis of the inferred molecular network models (Klipp et al., 2010). Furthermore, Translational Systems Biology of infection aims (i) to recognize pathogens by their molecular signatures, (ii) to make the outcome and responsiveness of therapeutic interventions more predictable, and (iii) to identify more effective therapies using mathematical and computational models. The first steps in Translational Systems Biology were taken in the area of sepsis control (Vodovotz et al., 2008) and tuberculosis research (Day et al., 2010). Translational Systems Biology of fungal infections with applications in personalized medicine (Willard and Ginsburg, 2009) can be expected to be developed in the near future.

ACKNOWLEDGMENTS

Part of this work was supported by the excellence graduate school Jena School for Microbial Communication (JSMC), funded by the Deutsche Forschungsgemeinschaft, as part o fit, by the International Leibniz Research School for Microbial and Biomolecular Interactions (ILRS Jena) and by a grant from the Federal Ministry of Education and Research (BMBF) in the frame of ERA-NetPathoGenoMics.

REFERENCES

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* 422, 198–207.

Aimanianda, V., Bayry, J., Bozza, S., Kniemeyer, O., Perruccio, K., Elluru, S. R., Clavaud, C., Paris, S., Brakhage, A. A., Kaveri, S. V., Romani, L., and Latgé, J.-P. (2009). Surface hydrophobin prevents immune recognition of airborne fungal spores. *Nature* 460, 1117–1121.

Alberti, C., Brun-Buisson, C., Goodman, S. V., Guidici, D., Granton, J., Moreno, R., Smithies, M., Thomas, O., Artigas, A., Gall, J. R. L., and Group, E. S. (2003). Influence of systemic inflammatory response syndrome and sepsis on outcome of critically ill infected patients. *Am. J. Respir. Crit. Care Med.* 168, 77–84.

Albrecht, D., Guthke, R., Brakhage, A. A., and Kniemeyer, O. (2010). Integrative analysis of the heat shock response in *Aspergillus fumigatus*. *BMC Genomics* 11, 32. doi:10.1186/1471-2164-11-32

Albrecht, D., Guthke, R., Kniemeyer, O., and Brakhage, A. A. (2008). "Systems biology of human-pathogenic fungi," in *Handbook of Research on Systems Biology Applications in Medicine*, Vol. 1, ed. A. Daskalaki (Hershey: IGI Global), 400–418.

Albrecht, D., Kniemeyer, O., Brakhage, A. A., Berth, M., and Guthke, R. (2007). Integration of transcriptome and proteome data from human-pathogenic fungi by using a data warehouse. *J. Integr. Bioinform.* 4, 52.

Albrecht, D., Kniemeyer, O., Mech, F., Gunzer, M., Brakhage, A. A., and Guthke, R. (2011). On the way toward systems biology of *Aspergillus fumigatus* infection. *Int. J. Med. Microbiol.* 301, 453–459.

Alby, K., Schaefer, D., and Bennett, R. J. (2009). Homothallic and heterothallic mating in the opportunistic pathogen *Candida albicans*. *Nature* 460, 890–893.

Aldridge, B. B., Burke, J. M., Lauffenburger, D. A., and Sorger, P. K. (2006). Physicochemical modelling of cell signalling pathways. *Nat. Cell Biol.* 8, 1195–1203.

Almeida, R. S., Wilson, D., and Hube, B. (2009). *Candida albicans* iron acquisition within the host. *FEMS Yeast Res.* 9, 1000–1012.

Altwasser, R., Linde, J., Buyko, E., Hahn, U., and Guthke, R. (2012). Genome-wide scale-free network inference for *Candida albicans*. *Front. Microbiol.* 3:51. doi:10.3389/fmicb.2012.00051

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

Arnaud, M. B., Cerqueira, G. C., Inglis, D. O., Skrzypek, M. S., Binkley, J., Chibucos, M. C., Crabtree, J., Howarth, C., Orvis, J., Shah, P., Wymore, F., Binkley, G., Miyasato, S. R., Simison, M., Sherlock, G., and Wortman, J. R. (2012). The Aspergillus Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. *Nucleic Acids Res.* 40, D653–D659.

Arnaud, M. B., Costanzo, M. C., Shah, P., Skrzypek, M. S., and Sherlock, G. (2009). Gene ontology and the annotation of pathogen genomes: the case of *Candida albicans*. *Trends Microbiol.* 17, 295–303.

Arnaud, M. B., Costanzo, M. C., Skrzypek, M. S., Shah, P., Binkley, G., Lane, C., Miyasato, S. R., and Sherlock, G. (2007). Sequence resources at the Candida Genome Database. *Nucleic Acids Res.* 35, D452–D456. [Database issue].

Avet, J., Granjon, D., Prevot-Bitot, N., Isnardi, V., Berger, C., Stephan, J. L., and Dubois, F. (2009). Monitoring of systemic candidiasis by 18F-FDG PET/CT. *Eur. J. Nucl. Med. Mol. Imaging* 36, 1900.

Backer, M. D. D., Ilyina, T., Ma, X. J., Vandoninck, S., Luyten, W. H., and Bossche, H. V. (2001). Genomic profiling of the response of *Candida albicans* to itraconazole treatment using a DNA microarray. *Antimicrob. Agents Chemother.* 45, 1660–1670.

Bailey, N. J. T. (1975). *The Mathematical Theory of Infectious Diseases and its Application*. London: Griffin.

Barabási, A.-L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113.

Barker, K. S., Crisp, S., Wiederhold, N., Lewis, R. E., Bareither, B., Eckstein, J., Barbuch, R., Bard, M., and Rogers, P. D. (2004). Genome-wide expression profiling reveals genes associated with amphotericin B and fluconazole resistance in experimentally induced antifungal resistant isolates of *Candida albicans*. *J. Antimicrob. Chemother.* 54, 376–385.

Barrett, L. G., Thrall, P. H., Burdon, J. J., and Linde, C. C. (2008). Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends Ecol. Evol. (Amst.)* 23, 678–685.

Behnsen, J., Hartmann, A., Schmaler, J., Gehrke, A., Brakhage, A. A., and Zipfel, P. F. (2008). The opportunistic human pathogenic fungus *Aspergillus fumigatus* evades the host complement system. *Infect. Immun.* 76, 820–827.

Behnsen, J., Lessing, F., Schindler, S., Wartenberg, D., Jacobsen, I. D., Thoen, M., Zipfel, P. F., and Brakhage, A. A. (2010). Secreted *Aspergillus fumigatus* protease Alp1 degrades human complement proteins C3, C4, and C5. *Infect. Immun.* 78, 3585–3594.

Behnsen, J., Narang, P., Hasenberg, M., Gunzer, F., Bilitewski, U., Klippel, N., Rohde, M., Brock, M., Brakhage, A. A., and Gunzer, M. (2007). Environmental dimensionality controls the interaction of phagocytes with the pathogenic fungi *Aspergillus fumigatus* and *Candida albicans*. *PLoS Pathog.* 3, e13. doi:10.1371/journal.ppat.0030013

Bernoulli, D. (1760). Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir. *Mémoires de Mathématiques et de Physique Paris Académie Royale des Sciences* 1, 1–45.

Blüthgen, N., Legewie, S., Kielbasa, S. M., Schramme, A., Tchernitsa, O., Keil, J., Solf, A., Vingron, M., Schäfer, R., Herzel, H., and Sers, C. (2009). A systems biological approach suggests that transcriptional feedback regulation by dual-specificity phosphatase 6 shapes extracellular signal-related kinase activity in RAS-transformed fibroblasts. *FEBS J.* 276, 1024–1035.

Borisov, N., Aksamitiene, E., Kiyatkin, A., Legewie, S., Berkhout, J., Maiwald, T., Kaimachnikov, N. P., Timmer, J., Hoek, J. B., and Kholodenko, B. N. (2009). Systems-level interactions between insulin-EGF networks amplify mitogenic signaling. *Mol. Syst. Biol.* 5, 256.

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

Brakhage, A. A., Bruns, S., Thywissen, A., Zipfel, P. F., and Behnsen, J. (2010). Interaction of phagocytes with filamentous fungi. *Curr. Opin. Microbiol.* 13, 409–415.

Braun, B. R., van Het Hoog, M., d'Enfert, C., Martchenko, M., Dungan, J., Kuo, A., Inglis, D. O., Uhl, M. A., Hogues, H., Berriman, M., Lorenz, M., Levitin, A., Oberholzer, U., Bachewich, C., Harcus, D., Marcil, A., Dignard, D., Iouk, T., Zito, R., Frangeul, L., Tekaia, F., Rutherford, K., Wang, E., Munro, C. A., Bates, S., Gow, N. A., Hoyer, L. L., Köhler, G., Morschhäuser, J., Newport, G., Znaidi, S., Raymond, M., Turcotte, B., Sherlock, G., Costanzo, M., Ihmels, J., Berman, J., Sanglard, D., Agabian, N., Mitchell, A. P., Johnson, A. D., Whiteway, M., and Nantel, A. (2005). A human-curated annotation of the *Candida albicans* genome. *PLoS Genet.* 1, 36–57.

Brock, M. (2012). Application of bioluminescence imaging for in vivo monitoring of fungal infections. *Int. J. Microbiol.* 2012, 956794.

Brodsky, I. E., and Medzhitov, R. (2009). Targeting of immune signalling networks by bacterial pathogens. *Nat. Cell Biol.* 11, 521–526.

Brown, A. J. P., Odds, F. C., and Gow, N. A. R. (2007). Infection-related gene expression in *Candida albicans*. *Curr. Opin. Microbiol.* 10, 307–313.

Bruneau, J. -M., Maillet, I., Tagat, E., Legrand, R., Supatto, F., Fudali, C., Caer, J.-P. L., Labas, V., Lecaque, D., and Hodgson, J. (2003). Drug induced proteome changes in *Candida albicans*: comparison of the effect of beta(1,3) glucan synthase inhibitors and two triazoles, fluconazole and itraconazole. *Proteomics* 3, 325–336.

Bruno, V. M., Wang, Z., Marjani, S. L., Euskirchen, G. M., Martin, J., Sherlock, G., and Snyder, M. (2010). Comprehensive annotation of the transcriptome of the human fungal pathogen *Candida albicans* using RNA-seq. *Genome Res.* 20, 1451–1458.

Bruns, S., Seidler, M., Albrecht, D., Salvenmoser, S., Remme, N., Hertweck, C., Brakhage, A. A., Kniemeyer, O., and Müller, F.-M. C. (2010). Functional genomic profiling of *Aspergillus fumigatus* biofilm reveals enhanced production of the mycotoxin gliotoxin. *Proteomics* 10, 3097–3107.

Burmester, A., Shelest, E., Glöckner, G., Heddergott, C., Schindler, S., Staib, P., Heidel, A., Felder, M., Petzold, A., Szafranski, K., Feuermann, M., Pedruzzi, I., Priebe, S., Groth, M., Winkler, R., Li, W., Kniemeyer, O., Schroeckh, V., Hertweck, C., Hube, B., White, T. C., Platzer, M., Guthke, R., Heitman, J., Wöstemeyer, J., Zipfel, P. F., Monod, M., and Brakhage, A. A. (2011). Comparative and functional genomics provide insights into the pathogenicity of dermatophytic fungi. *Genome Biol.* 12, R7.

Butler, G., Rasmussen, M. D., Lin, M. F., Santos, M. A. S., Sakthikumar, S., Munro, C. A., Rheinbay, E., Grabherr, M., Forche, A., Reedy, J. L., Agrafioti, I., Arnaud, M. B., Bates, S., Brown, A. J. P., Brunke, S., Costanzo, M. C., Fitzpatrick, D. A., de Groot, P. W. J., Harris, D., Hoyer, L. L., Hube, B., Klis, F. M., Kodira, C., Lennard, N., Logue, M. E., Martin, R., Neiman, A. M., Nikolaou, E., Quail, M. A., Quinn, J., Santos, M. C., Schmitzberger, F. F., Sherlock, G., Shah, P., Silverstein, K. A. T., Skrzypek, M. S., Soll, D., Staggs, R., Stansfield, I., Stumpf, M. P. H., Sudbery, P. E., Srikantha, T., Zeng, Q., Berman, J., Berriman, M., Heitman, J., Gow, N. A. R., Lorenz, M. C., Birren, B. W., Kellis, M., and Cuomo, C. A. (2009). Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. *Nature* 459, 657–662.

Cagas, S. E., Jain, M. R., Li, H., and Perlin, D. S. (2011). Profiling the *Aspergillus fumigatus* proteome in response to caspofungin. *Antimicrob. Agents Chemother.* 55, 146–154.

Cairns, T., Minuzzi, F., and Bignell, E. (2010). The host-infecting fungal transcriptome. *FEMS Microbiol. Lett.* 307, 1–11.

Castillo, L., Calvo, E., Martínez, A. I., Ruiz-Herrera, J., Valentín, E., Lopez, J. A., and Sentandreu, R. (2008). A study of the *Candida albicans* cell wall proteome. *Proteomics* 8, 3871–3881.

Cavalieri, D., and Filippo, C. D. (2005). Bioinformatic methods for integrating whole-genome expression results into cellular networks. *Drug Discov. Today* 10, 727–734.

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

Chen, X. S. (2011). Introducing Theranostics journal – from the editor-in-chief. *Theranostics* 1, 1–2.

Correia, A., Lermann, U., Teixeira, L., Cerca, F., Botelho, S., da Costa, R. M. G., Sampaio, P., Gärtner, F., Morschhäuser, J., Vilanova, M., and Pais, C. (2010). Limited role of secreted aspartyl proteinases Sap1 to Sap6 in *Candida albicans* virulence and host immune response in murine hematogenously disseminated candidiasis. *Infect. Immun.* 78, 4839–4849.

Courtot, M., Juty, N., Knüpfer, C., Waltemath, D., Zhukova, A., Dräger, A., Dumontier, M., Finney, A., Golebiewski, M., Hastings, J., Hoops, S., Keating, S., Kell, D. B., Kerrien, S., Lawson, J., Lister, A., Lu, J., Machne, R., Mendes, P., Pocock, M., Rodriguez, N., Villeger, A., Wilkinson, D. J., Wimalaratne, S., Laibe, C., Hucka, M., and Novère, N. L. (2011). Controlled vocabularies and semantics in systems biology. *Mol. Syst. Biol.* 7, 543.

da Silva Ferreira, M. E., Malavazi, I., Savoldi, M., Brakhage, A. A., Goldman, M. H. S., Kim, H. S., Nierman, W. C., and Goldman, G. H. (2006). Transcriptome analysis of *Aspergillus fumigatus* exposed to voriconazole. *Curr. Genet.* 50, 32–44.

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

Day, J., Schlesinger, L. S., and Friedman, A. (2010). Tuberculosis research: going forward with a powerful "translational systems biology" approach. *Tuberculosis (Edinb.)* 90, 7–8.

de Groot, P. W. J., de Boer, A. D., Cunningham, J., Dekker, H. L., de Jong, L., Hellingwerf, K. J., de Koster, C., and Klis, F. M. (2004). Proteomic analysis of *Candida albicans* cell walls reveals covalently bound carbohydrate-active enzymes and adhesins. *Eukaryotic Cell* 3, 955–965.

d'Enfert, C., Goyard, S., Rodriguez-Arnaveilhe, S., Frangeul, L., Jones, L., Tekaia, F., Bader, O., Albrecht, A., Castillo, L., Dominguez, A., Ernst, J. F., Fradin, C., Gaillardin, C., Garcia-Sanchez, S., de Groot, P., Hube, B., Klis, F. M., Krishnamurthy, S., Kunze, D., Lopez, M.-C., Mavor, A., Martin, N., Moszer, I., Onésime, D., Martin, J. P., Sentandreu, R., Valentin, E., and Brown, A. J. P. (2005). CandidaDB: a genome database for *Candida albicans* pathogenomics. *Nucleic Acids Res.* 33, D353–D357. [Database issue].

Doedt, T., Krishnamurthy, S., Bockmühl, D. P., Tebarth, B., Stempel, C., Russell, C. L., Brown, A. J. P., and Ernst, J. F. (2004). APSES proteins regulate morphogenesis and metabolism in *Candida albicans*. *Mol. Biol. Cell* 15, 3167–3180.

Domon, B., and Aebersold, R. (2010). Options and considerations when selecting a quantitative proteomics strategy. *Nat. Biotechnol.* 28, 710–721.

Dujon, B. (2010). Yeast evolutionary genomics. *Nat. Rev. Genet.* 11, 512–524.

Dyer, M. D., Murali, T. M., and Sobral, B. W. (2007). Computational prediction of host-pathogen protein-protein interactions. *Bioinformatics* 23, i159–i166.

Dyer, M. D., Murali, T. M., and Sobral, B. W. (2008). The landscape of human proteins interacting with viruses and other pathogens. *PLoS Pathog.* 4, e32. doi:10.1371/journal.ppat.0040032

Fang, X., Wallqvist, A., and Reifman, J. (2009). A systems biology framework for modeling metabolic enzyme inhibition of *Mycobacterium tuberculosis*. *BMC Syst. Biol.* 3, 92.

Fazius, E., Shelest, V., and Shelest, E. (2011). SiTaR: a novel tool for transcription factor binding site prediction. *Bioinformatics* 27, 2806–2811.

Fedorova, N. D., Khaldi, N., Joardar, V. S., Maiti, R., Amedeo, P., Anderson, M. J., Crabtree, J., Silva, J. C., Badger, J. H., Albarraq, A., Angiuoli, S., Bussey, H., Bowyer, P., Cotty, P. J., Dyer, P. S., Egan, A., Galens, K., Fraser-Liggett, C. M., Haas, B. J., Inman, J. M., Kent, R., Lemieux, S., Malavazi, I., Orvis, J., Roemer, T., Ronning, C. M., Sundaram, J. P., Sutton, G., Turner, G., Venter, J. C., White, O. R., Whitty, B. R., Youngman, P., Wolfe, K. H., Goldman, G. H., Wortman, J. R., Jiang, B., Denning, D. W., and Nierman, W. C. (2008). Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. *PLoS Genet.* 4, e1000046. doi:10.1371/journal.pgen.1000046

Feldmesser, M. (2006). Role of neutrophils in invasive aspergillosis. *Infect. Immun.* 74, 6514–6516.

Ferrandon, D., Imler, J.-L., Hetru, C., and Hoffmann, J. A. (2007). The *Drosophila* systemic immune response: sensing and signalling during bacterial and fungal infections. *Nat. Rev. Immunol.* 7, 862–874.

Figge, M. T. (2005). Stochastic discrete event simulation of germinal center reactions. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 71, 051907.

Figge, M. T., Garin, A., Gunzer, M., Kosco-Vilbois, M., Toellner, K.-M., and Meyer-Hermann, M. (2008). Deriving a germinal center lymphocyte migration model from two-photon data. *J. Exp. Med.* 205, 3019–3029.

Filler, S. G. (2006). *Candida*-host cell receptor-ligand interactions. *Curr. Opin. Microbiol.* 9, 333–339.

Fleck, C. B., Schöbel, F., and Brock, M. (2011). Nutrient acquisition by pathogenic fungi: nutrient availability, pathway regulation, and differences in substrate utilization. *Int. J. Med. Microbiol.* 301, 400–407.

Forst, C. V. (2006). Host-pathogen systems biology. *Drug Discov. Today* 11, 220–227.

Fradin, C., Kretschmar, M., Nichterlein, T., Gaillardin, C., d'Enfert, C., and Hube, B. (2003). Stage-specific gene expression of *Candida albicans* in human blood. *Mol. Microbiol.* 47, 1523–1543.

Franke, R., Müller, M., Wundrack, N., Gilles, E.-D., Klamt, S., Kähne, T., and Naumann, M. (2008). Host-pathogen systems biology: logical modelling of hepatocyte growth factor and *Helicobacter pylori* induced c-Met signal transduction. *BMC Syst. Biol.* 2, 4. doi:10.1186/1752-0509-2-4

Gafa, V., Remoli, M. E., Giacomini, E., Gagliardi, M. C., Lande, R., Severa, M., Grillot, R., and Coccia, E. M. (2007). In vitro infection of human dendritic cells by *Aspergillus fumigatus* conidia triggers the secretion of chemokines for neutrophil and Th1 lymphocyte recruitment. *Microbes Infect.* 9, 971–980.

Garin, A., Meyer-Hermann, M., Contie, M., Figge, M. T., Buatois, V., Gunzer, M., Toellner, K. -M., Elson, G., and Kosco-Vilbois, M. H. (2010). Toll-like receptor 4 signaling by follicular dendritic cells is pivotal for germinal center onset and affinity maturation. *Immunity* 33, 84–95.

Gautam, P., Shankar, J., Madan, T., Sirdeshmukh, R., Sundaram, C. S., Gade, W. N., Basir, S. F., and Sarma, P. U. (2008). Proteomic and transcriptomic analysis of *Aspergillus fumigatus* on exposure to amphotericin B. *Antimicrob. Agents Chemother.* 52, 4220–4227.

Gibbons, J. G., Beauvais, A., Beau, R., McGary, K. L., Latge, J.-P., and Rokas, A. (2012). Global transcriptome changes underlying colony growth in the opportunistic human pathogen *Aspergillus fumigatus*. *Eukaryotic Cell.* 11, 68–78.

Gilsenan, J. M., Cooley, J., and Bowyer, P. (2012). CADRE: The Central Aspergillus Data Repository 2012. *Nucleic Acids Res.* 40, D660–D666.

Grahl, N., Puttikamonkul, S., MacDonald, J. M., Gamcsik, M. P., Ngo, L. Y., Hohl, T. M., and Cramer, R. A. (2011). In vivo hypoxia and a fungal alcohol dehydrogenase influence the pathogenesis of invasive pulmonary aspergillosis. *PLoS Pathog.* 7, e1002145. doi:10.1371/journal.ppat.1002145

Gropp, K., Schild, L., Schindler, S., Hube, B., Zipfel, P. F., and Skerka, C. (2009). The yeast *Candida albicans* evades human complement attack by secretion of aspartic proteases. *Mol. Immunol.* 47, 465–475.

Gross, O., Poeck, H., Bscheider, M., Dostert, C., Hannesschläger, N., Endres, S., Hartmann, G., Tardivel, A., Schweighoffer, E., Tybulewicz, V., Mocsai, A., Tschopp, J., and Ruland, J. (2009). Syk kinase signalling couples to the Nlrp3 inflammasome for anti-fungal host defence. *Nature* 459, 433–436.

Gudlaugsson, O., Gillespie, S., Lee, K., Berg, J. V., Hu, J., Messer, S., Herwaldt, L., Pfaller, M., and Diekema, D. (2003). Attributable mortality of nosocomial candidemia, revisited. *Clin. Infect. Dis.* 37, 1172–1177.

Guo, Z., Sloot, P. M., and Tay, J. C. (2008). A hybrid agent-based approach for modeling microbiological systems. *J. Theor. Biol.* 255, 163–175.

Guo, Z., and Tay, J. (2005). A comparative study on modeling strategies for immune system dynamics under HIV-1 infection. *Artif. Immune Syst.* 3627, 220–233.

Guthke, R., Kniemeyer, O., Albrecht, D., Brakhage, A. A., and Möller, U. (2007). Discovery of gene regulatory networks in *Aspergillus fumigatus*. *Lect. Notes Bioinform.* 4366, 22–41.

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

Hajishengallis, G., and Lambris, J. D. (2011). Microbial manipulation of receptor crosstalk in innate immunity. *Nat. Rev. Immunol.* 11, 187–200.

Hase, T., Tanaka, H., Suzuki, Y., Nakagawa, S., and Kitano, H. (2009). Structure of protein interaction networks and their implications on drug design. *PLoS Comput. Biol.* 5, e1000550. doi:10.1371/journal.pcbi.1000550

Hawksworth, D. (2001). The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol. Res.* 105, 1422–1432.

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

Heilmann, C. J., Sorgo, A. G., Siliakus, A. R., Dekker, H. L., Brul, S., de Koster, C. G., de Koning, L. J., and Klis, F. M. (2011). Hyphal induction in the human fungal pathogen *Candida albicans* reveals a characteristic wall protein profile. *Microbiology* 157, 2297–2307.

Hermjakob, H., Montecchi-Palazzi, L., Bader, G., Wojcik, J., Salwinski, L., Ceol, A., Moore, S., Orchard, S., Sarkans, U., von Mering, C., Roechert, B., Poux, S., Jung, E., Mersch, H., Kersey, P., Lappe, M., Li, Y., Zeng, R., Rana, D., Nikolski, M., Husi, H., Brun, C., Shanker, K., Grant, S. G. N., Sander, C., Bork, P., Zhu, W., Pandey, A., Brazma, A., Jacq, B., Vidal, M., Sherman, D., Legrain, P., Cesareni, G., Xenarios, I., Eisenberg, D., Steipe, B., Hogue, C., and Apweiler, R. (2004). The HUPO PSI's molecular interaction format – a community standard for the representation of protein interaction data. *Nat. Biotechnol.* 22, 177–183.

Hernáez, M. L., Ximénez-Embún, P., Martínez-Gomariz, M., Gutiérrez-Blázquez, M. D., Nombela, C., and Gil, C. (2010). Identification of *Candida albicans* exposed surface proteins in vivo by a rapid proteomic approach. *J. Proteomics* 73, 1404–1409.

Hickey, P. C., and Read, N. D. (2009). Imaging living cells of *Aspergillus* in vitro. *Med. Mycol.* 47(Suppl. 1), S110–S119.

Hise, A. G., Tomalka, J., Ganesan, S., Patel, K., Hall, B. A., Brown, G. D., and Fitzgerald, K. A. (2009). An essential role for the NLRP3 inflammasome in host defense against the human fungal pathogen *Candida albicans*. *Cell Host Microbe* 5, 487–497.

Hoehamer, C. F., Cummings, E. D., Hilliard, G. M., and Rogers, P. D. (2010). Changes in the proteome of *Candida albicans* in response to azole, polyene, and echinocandin antifungal agents. *Antimicrob. Agents Chemother.* 54, 1655–1664.

Hooshdaran, M. Z., Barker, K. S., Hilliard, G. M., Kusch, H., Morschhäuser, J., and Rogers, P. D. (2004). Proteomic analysis of azole resistance in *Candida albicans* clinical isolates. *Antimicrob. Agents Chemother.* 48, 2733–2735.

Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., Cuellar, A. A., Dronov, S., Gilles, E. D., Ginkel, M., Gor, V., Goryanin, I. I., Hedley, W. J., Hodgman, T. C., Hofmeyr, J.-H., Hunter, P. J., Juty, N. S., Kasberger, J. L., Kremling, A., Kummer, U., Novère, N. L., Loew, L. M., Lucio, D., Mendes, P., Minch, E., Mjolsness, E. D., Nakayama, Y., Nelson, M. R., Nielsen, P. F., Sakurada, T., Schaff, J. C., Shapiro, B. E., Shimizu, T. S., Spence, H. D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J., Wang, J., and Forum, S. B. M. L. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19, 524–531.

Hull, D., Wolstencroft, K., Stevens, R., Goble, C., Pocock, M. R., Li, P., and Oinn, T. (2006). Taverna: a tool for building and running workflows of services. *Nucleic Acids Res.* 34, W729–W732. [Web Server issue].

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

Ibrahim-Granet, O., Jouvion, G., Hohl, T. M., Droin-Bergère, S., Philippart, F., Kim, O. Y., Adib-Conquy, M., Schwendener, R., Cavaillon, J.-M., and Brock, M. (2010). In vivo bioluminescence imaging and histopathopathologic analysis reveal distinct roles for resident and recruited immune effector cells in defense against invasive aspergillosis. *BMC Microbiol.* 10, 105. doi:10.1186/1471-2180-10-105

Ibrahim-Granet, O., Philippe, B., Boleti, H., Boisvieux-Ulrich, E., Grenet, D., Stern, M., and Latgé, J. P. (2003). Phagocytosis and intracellular fate of *Aspergillus fumigatus* conidia in alveolar macrophages. *Infect. Immun.* 71, 891–903.

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

Ideker, T., and Sharan, R. (2008). Protein networks in disease. *Genome Res.* 18, 644–652.

Ihmels, J., Bergmann, S., Berman, J., and Barkai, N. (2005). Comparative gene expression analysis by differential clustering approach: application to the *Candida albicans* transcription program. *PLoS Genet.* 1, e39. doi:10.1371/journal.pgen.0010039

Jacobsen, I. D., Grosse, K., Slesiona, S., Hube, B., Berndt, A., and Brock, M. (2010). Embryonated eggs as an alternative infection model to investigate *Aspergillus fumigatus* virulence. *Infect. Immun.* 78, 2995–3006.

Jahn, B., Langfelder, K., Schneider, U., Schindel, C., and Brakhage, A. A. (2002). PKSP-dependent reduction

of phagolysosome fusion and intracellular kill of *Aspergillus fumigatus* conidia by human monocyte-derived macrophages. *Cell. Microbiol.* 4, 793–803.

Jain, R., Valiante, V., Remme, N., Docimo, T., Heinekamp, T., Hertweck, C., Gershenzon, J., Haas, H., and Brakhage, A. A. (2011). The MAP kinase MpkA controls cell wall integrity, oxidative stress response, gliotoxin production and iron adaptation in *Aspergillus fumigatus*. *Mol. Microbiol.* 82, 39–53.

Jones, T., Federspiel, N. A., Chibana, H., Dungan, J., Kalman, S., Magee, B. B., Newport, G., Thorstenson, Y. R., Agabian, N., Magee, P. T., Davis, R. W., and Scherer, S. (2004). The diploid genome sequence of *Candida albicans*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7329–7334.

Kim, Y., Nandakumar, M. P., and Marten, M. R. (2008). The state of proteome profiling in the fungal genus *Aspergillus*. *Brief. Funct. Genomic. Proteomic.* 7, 87–94.

Kitano, H. (2007). A robustness-based approach to systems-oriented drug design. *Nat. Rev. Drug Discov.* 6, 202–210.

Klipp, E., Wade, R. C., and Kummer, U. (2010). Biochemical network-based drug-target prediction. *Curr. Opin. Biotechnol.* 21, 511–516.

Klose, J., and Kobalz, U. (1995). Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* 16, 1034–1059.

Kniemeyer, O. (2011). Proteomics of eukaryotic microorganisms: the medically and biotechnologically important fungal genus *Aspergillus*. *Proteomics* 11, 3232–3243.

Kniemeyer, O., and Brakhage, A. A. (2008). "Proteomics and its application to the human-pathogenic fungi," in *Human and Animal Relationships: The Mycota VI*, eds A. A. Brakhage and P. F. Zipfel (Berlin: Springer), 154–186.

Kniemeyer, O., Schmidt, A. D., Vödisch, M., Wartenberg, D., and Brakhage, A. A. (2011). Identification of virulence determinants of the human pathogenic fungi *Aspergillus fumigatus* and *Candida albicans* by proteomics. *Int. J. Med. Microbiol.* 301, 368–377.

Köhler, J., Baumbach, J., Taubert, J., Specht, M., Skusa, A., Rüegg, A., Rawlings, C., Verrier, P., and Philippi, S. (2006). Graph-based analysis and visualization of experimental results with ONDEX. *Bioinformatics* 22, 1383–1390.

Kozhenkov, S., Sedova, M., Dubinina, Y., Gupta, A., Ray, A., Ponomarenko, J., and Baitaluk, M. (2011). Biological networks – tools enabling the integration of multi-scale data for the host-pathogen studies. *BMC Syst. Biol.* 5, 7. doi:10.1186/1752-0509-5-7

Krappmann, S., Sasse, C., and Braus, G. H. (2006). Gene targeting in *Aspergillus fumigatus* by homologous recombination is facilitated in a nonhomologous end-joining-deficient genetic background. *Eukaryotic Cell* 5, 212–215.

Kusch, H., Engelmann, S., Albrecht, D., Morschhäuser, J., and Hecker, M. (2007). Proteomic analysis of the oxidative stress response in *Candida albicans*. *Proteomics* 7, 686–697.

Langfelder, K., Jahn, B., Gehringer, H., Schmidt, A., Wanner, G., and Brakhage, A. A. (1998). Identification of a polyketide synthase gene (pksP) of *Aspergillus fumigatus* involved in conidial pigment biosynthesis and virulence. *Med. Microbiol. Immunol.* 187, 79–89.

Latgé, J. P. (1999). *Aspergillus fumigatus* and aspergillosis. *Clin. Microbiol. Rev.* 12, 310–350.

Le Novère, N., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E., Wegner, K., Aladjem, M. I., Wimalaratne, S. M., Bergman, F. T., Gauges, R., Ghazal, P., Kawaji, H., Li, L., Matsuoka, Y., Villéger, A., Boyd, S. E., Calzone, L., Courtot, M., Dogrusoz, U., Freeman, T. C., Funahashi, A., Ghosh, S., Jouraku, A., Kim, S., Kolpakov, F., Luna, A., Sahle, S., Schmidt, E., Watterson, S., Wu, G., Goryanin, I., Kell, D. B., Sander, C., Sauro, H., Snoep, J. L., Kohn, K., and Kitano, H. (2009). The systems biology graphical notation. *Nat. Biotechnol.* 27, 735–741.

Lee, D.-S., Burd, H., Liu, J., Almaas, E., Wiest, O., Barabási, A.-L., Oltvai, Z. N., and Kapatral, V. (2009). Comparative genome-scale metabolic reconstruction and flux balance analysis of multiple *Staphylococcus aureus* genomes identify novel antimicrobial drug targets. *J. Bacteriol.* 191, 4015–4024.

Lessing, F., Kniemeyer, O., Wozniok, I., Loeffler, J., Kurzai, O., Haertl, A., and Brakhage, A. A. (2007). The *Aspergillus fumigatus* transcriptional regulator AfYap1 represents the major regulator for defense against reactive oxygen intermediates but is dispensable for pathogenicity in an intranasal mouse infection model. *Eukaryotic Cell* 6, 2290–2302.

Linde, J., Hortschansky, P., Fazius, E., Brakhage, A. A., Guthke, R., and Haas, H. (2012). Regulatory interactions for iron homoeostatis in *Aspergillus fumigatus* inferred by a systems biology approach. *BMC Syst. Biol.* 6, 6. doi:10.1186/1752-0509-6-6

Linde, J., Wilson, D., Hube, B., and Guthke, R. (2010). Regulatory network modelling of iron acquisition by a fungal pathogen in contact with epithelial cells. *BMC Syst. Biol.* 4, 148. doi:10.1186/1752-0509-4-148

Liu, H. (2002). Co-regulation of pathogenesis with dimorphism and phenotypic switching in *Candida albicans*, a commensal and a pathogen. *Int. J. Med. Microbiol.* 292, 299–311.

Liu, T. T., Lee, R. E. B., Barker, K. S., Lee, R. E., Wei, L., Homayouni, R., and Rogers, P. D. (2005). Genome-wide expression profiling of the response to azole, polyene, echinocandin, and pyrimidine antifungal agents in *Candida albicans*. *Antimicrob. Agents Chemother.* 49, 2226–2236.

Lloyd, A. T., and Sharp, P. M. (1992). Evolution of codon usage patterns: the extent and nature of divergence between *Candida albicans* and *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 20, 5289–5295.

Lo, H. J., Köhler, J. R., DiDomenico, B., Loebenberg, D., Cacciapuoti, A., and Fink, G. R. (1997). Nonfilamentous *C. albicans* mutants are avirulent. *Cell* 90, 939–949.

Loftus, B. J., Fung, E., Roncaglia, P., Rowley, D., Amedeo, P., Bruno, D., Vamathevan, J., Miranda, M., Anderson, I. J., Fraser, J. A., Allen, J. E., Bosdet, I. E., Brent, M. R., Chiu, R., Doering, T. L., Donlin, M. J., D'Souza, C. A., Fox, D. S., Grinberg, V., Fu, J., Fukushima, M., Haas, B. J., Huang, J. C., Janbon, G., Jones, S. J. M., Koo, H. L., Krzywinski, M. I., Kwon-Chung, J. K., Lengeler, K. B., Maiti, R., Marra, M. A., Marra, R. E., Mathewson, C. A., Mitchell, T. G., Pertea, M., Riggs, F. R., Salzberg, S. L., Schein, J. E., Shvartsbeyn, A., Shin, H., Shumway, M., Specht, C. A., Suh, B. B., Tenney, A., Utterback, T. R., Wickes, B. L., Wortman, J. R., Wye, N. H., Kronstad, J. W., Lodge, J. K., Heitman, J., Davis, R. W., Fraser, C. M., and Hyman, R. W. (2005). The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* 307, 1321–1324.

Lorenz, M. C., Bender, J. A., and Fink, G. R. (2004). Transcriptional response of *Candida albicans* upon internalization by macrophages. *Eukaryotic Cell* 3, 1076–1087.

Lorenz, M. C., and Fink, G. R. (2001). The glyoxylate cycle is required for fungal virulence. *Nature* 412, 83–86.

Manning, M., and Mitchell, T. G. (1980). Morphogenesis of *Candida albicans* and cytoplasmic proteins associated with differences in morphology, strain, or temperature. *J. Bacteriol.* 144, 258–273.

Marbach, D., Prill, R. J., Schaffter, T., Mattiussi, C., Floreano, D., and Stolovitzky, G. (2010). Revealing strengths and weaknesses of methods for gene network inference. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6286–6291.

Marr, K. A., Carter, R. A., Boeckh, M., Martin, P., and Corey, L. (2002). Invasive aspergillosis in allogeneic stem cell transplant recipients: changes in epidemiology and risk factors. *Blood* 100, 4358–4366.

Martin, G. S., Mannino, D. M., Eaton, S., and Moss, M. (2003). The epidemiology of sepsis in the United States from 1979 through 2000. *N. Engl. J. Med.* 348, 1546–1554.

Martin, R., Wächtler, B., Schaller, M., Wilson, D., and Hube, B. (2011). Host-pathogen interactions and virulence-associated genes during *Candida albicans* oral infections. *Int. J. Med. Microbiol.* 301, 417–422.

Martínez-Solano, L., Nombela, C., Molero, G., and Gil, C. (2006). Differential protein expression of murine macrophages upon interaction with *Candida albicans*. *Proteomics* 6(Suppl. 1), S133–S144.

Martínez-Solano, L., Reales-Calderón, J. A., Nombela, C., Molero, G., and Gil, C. (2009). Proteomics of RAW 264.7 macrophages upon interaction with heat-inactivated *Candida albicans* cells unravel an anti-inflammatory response. *Proteomics* 9, 2995–3010.

May, R. M., and Anderson, R. M. (1979). Population biology of infectious diseases. Part II. *Nature* 280, 455–461.

McCluskey, K., Wiest, A., and Plamann, M. (2010). The Fungal Genetics Stock Center: a repository for 50 years of fungal genetics research. *J. Biosci.* 35, 119–126.

McCormick, A., Heesemann, L., Wagener, J., Marcos, V., Hartl, D., Loeffler, J., Heesemann, J., and Ebel, F. (2010). NETs formed by human neutrophils inhibit growth of the pathogenic mold *Aspergillus fumigatus*. *Microbes Infect.* 12, 928–936.

Mech, F., Thywissen, A., Guthke, R., Brakhage, A. A., and Figge, M. T. (2011). Automated image analysis

of the host-pathogen interaction between phagocytes and *Aspergillus fumigatus*. *PLoS ONE* 6, e19591. doi:10.1371/journal.pone.0019591

Melanson, J. E., Chisholm, K. A., and Pinto, D. M. (2006). Targeted comparative proteomics by liquid chromatography/matrix-assisted laser desorption/ionization triple-quadrupole mass spectrometry. *Rapid Commun. Mass Spectrom.* 20, 904–910.

Mezger, M., Kneitz, S., Wozniok, I., Kurzai, O., Einsele, H., and Loeffler, J. (2008). Proinflammatory response of immature human dendritic cells is mediated by dectin-1 after exposure to *Aspergillus fumigatus* germ tubes. *J. Infect. Dis.* 197, 924–931.

Moalli, F., Doni, A., Deban, L., Zelante, T., Zagarella, S., Bottazzi, B., Romani, L., Mantovani, A., and Garlanda, C. (2010). Role of complement and fcgamma receptors in the protective activity of the long pentraxin PTX3 against *Aspergillus fumigatus*. *Blood* 116, 5170–5180.

Monteoliva, L., Martinez-Lopez, R., Pitarch, A., Hernaez, M. L., Serna, A., Nombela, C., Albar, J. P., and Gil, C. (2011). Quantitative proteome and acidic subproteome profiling of *Candida albicans* yeast-to-hypha transition. *J. Proteome Res.* 10, 502–517.

Moran, G. P., Coleman, D. C., and Sullivan, D. J. (2011). Comparative genomics and the evolution of pathogenicity in human pathogenic fungi. *Eukaryotic Cell* 10, 34–42.

Morton, C. O., Varga, J. J., Hornbach, A., Mezger, M., Sennefelder, H., Kneitz, S., Kurzai, O., Krappmann, S., Einsele, H., Nierman, W. C., Rogers, T. R., and Loeffler, J. (2011). The temporal dynamics of differential gene expression in *Aspergillus fumigatus* interacting with human immature dendritic cells in vitro. *PLoS ONE* 6, e16016. doi:10.1371/journal.pone.0016016

Moy, T. I., Conery, A. L., Larkins-Ford, J., Wu, G., Mazitschek, R., Casadei, G., Lewis, K., Carpenter, A. E., and Ausubel, F. M. (2009). High-throughput screen for novel antimicrobials using a whole animal infection model. *ACS Chem. Biol.* 4, 527–533.

Mukhtar, M. S., Carvunis, A.-R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., Tasan, M., Galli, M., Hao, T., Nishimura, M. T., Pevzner, S. J., Donovan, S. E., Ghamsari, L., Santhanam, B., Romero, V., Poulin, M. M., Gebreab, F., Gutierrez, B. J., Tam, S., Monachello, D., Boxem, M., Harbort, C. J., McDonald, N., Gai, L., Chen, H., He, Y.„ E. U. E. C., Vandenhaute, J., Roth, F. P., Hill, D. E., Ecker, J. R., Vidal, M., Beynon, J., Braun, P., and Dangl, J. L. (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science* 333, 596–601.

Murillo, L. A., Newport, G., Lan, C.-Y., Habelitz, S., Dungan, J., and Agabian, N. M. (2005). Genome-wide transcription profiling of the early phase of biofilm formation by *Candida albicans*. *Eukaryotic Cell* 4, 1562–1573.

Mylonakis, E. (2008). *Galleria mellonella* and the study of fungal pathogenesis: making the case for another genetically tractable model host. *Mycopathologia* 165, 1–3.

Netea, M. G., Brown, G. D., Kullberg, B. J., and Gow, N. A. R. (2008). An integrated model of the recognition of *Candida albicans* by the innate immune system. *Nat. Rev. Microbiol.* 6, 67–78.

Neugebauer, U., Schmid, U., Baumann, K., Holzgrabe, U., Ziebuhr, W., Kozitskaya, S., Kiefer, W., Schmitt, M., and Popp, J. (2006). Characterization of bacterial growth and the influence of antibiotics by means of UV resonance Raman spectroscopy. *Biopolymers* 82, 306–311.

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., Lafton, 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., de Córdoba, S. R., and 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., de Aldana, C. R. V., Weidman, J., White, O., Woodward, J., Yu, J. H., Fraser, C., Galagan, J. E., Asai, K., Machida, M., Hall, N., Barrell, B., and Denning, D. W. (2005). Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* 438, 1151–1156.

Niimi, M., Shepherd, M. G., and Monk, B. C. (1996). Differential profiles of soluble proteins during the initiation of morphogenesis in *Candida albicans*. *Arch. Microbiol.* 166, 260–268.

Nobile, C. J., Nett, J. E., Andes, D. R., and Mitchell, A. P. (2006). Function of *Candida albicans* adhesin Hwp1 in biofilm formation. *Eukaryotic Cell* 5, 1604–1610.

Nygaard, V., and Hovig, E. (2006). Options available for profiling small samples: a review of sample amplification technology when combined with microarray profiling. *Nucleic Acids Res.* 34, 996–1014.

Oberhardt, M. A., Puchalka, J., Martins dos Santos, V. A. P., and Papin, J. A. (2011). Reconciliation of genome-scale metabolic reconstructions for comparative systems analysis. *PLoS Comput. Biol.* 7, e1001116. doi:10.1371/journal.pcbi.1001116

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007–4021.

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

Pfaller, M. A., and Diekema, D. J. (2007). Epidemiology of invasive candidiasis: a persistent public health problem. *Clin. Microbiol. Rev.* 20, 133–163.

Picazo, J. J., González-Romo, F., and Candel, F. J. (2008). Candidemia in the critically ill patient. *Int. J. Antimicrob. Agents* 32(Suppl. 2), S83–S85.

Pitarch, A., Sánchez, M., Nombela, C., and Gil, C. (2002). Sequential fractionation and two-dimensional gel analysis unravels the complexity of the dimorphic fungus *Candida albicans* cell wall proteome. *Mol. Cell Proteomics* 1, 967–982.

Pitkänen, E., Rousu, J., and Ukkonen, E. (2010). Computational methods for metabolic reconstruction. *Curr. Opin. Biotechnol.* 21, 70–77.

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

Raghunathan, A., Reed, J., Shin, S., Palsson, B., and Daefler, S. (2009). Constraint-based analysis of metabolic capacity of *Salmonella typhimurium* during host-pathogen interaction. *BMC Syst. Biol.* 3, 38. doi:10.1186/1752-0509-3-38

Rao, A., Kumar, M. K., Joseph, T., and Bulusu, G. (2010). Cerebral malaria: insights from host-parasite protein-protein interactions. *Malar. J.* 9, 155.

Raza, S., Robertson, K. A., Lacaze, P. A., Page, D., Enright, A. J., Ghazal, P., and Freeman, T. C. (2008). A logic-based diagram of signalling pathways central to macrophage activation. *BMC Syst. Biol.* 2, 36. doi:10.1186/1752-0509-2-36

Real, L. A., and Biek, R. (2007). Infectious disease modeling and the dynamics of transmission. *Curr. Top. Microbiol. Immunol.* 315, 33–49.

Rispail, N., Soanes, D. M., Ant, C., Czajkowski, R., Grünler, A., Huguet, R., Perez-Nadales, E., Poli, A., Sartorel, E., Valiante, V., Yang, M., Beffa, R., Brakhage, A. A., Gow, N. A. R., Kahmann, R., Lebrun, M.-H., Lenasi, H., Perez-Martin, J., Talbot, N. J., Wendland, J., and Di Pietro, A. (2009). Comparative genomics of MAP kinase and calcium-calcineurin signalling components in plant and human pathogenic fungi. *Fungal Genet. Biol.* 46, 287–298.

Rizzetto, L., and Cavalieri, D. (2011). Friend or foe: using systems biology to elucidate interactions between fungi and their hosts. *Trends Microbiol.* 19, 509–515.

Romani, L. (2011). Immunity to fungal infections. *Nat. Rev. Immunol.* 11, 275–288.

Ross, R. (1911). *The Prevention of Malaria, with Addendum on the Theory of Happenings*. London: Murray.

Roux, A. V. D., and Aiello, A. E. (2005). Multilevel analysis of infectious diseases. *J. Infect. Dis.* 191(Suppl. 1), S25–S33.

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

Rupp, S. (2004). Proteomics on its way to study host-pathogen interaction in *Candida albicans*. *Curr. Opin. Microbiol.* 7, 330–335.

Ruppin, E., Papin, J. A., de Figueiredo, L. F., and Schuster, S. (2010). Metabolic reconstruction, constraint-based analysis and game theory to probe genome-scale metabolic networks. *Curr. Opin. Biotechnol.* 21, 502–510.

Saïd-Sadier, N., Padilla, E., Langsley, G., and Ojcius, D. M. (2010). *Aspergillus fumigatus* stimulates the NLRP3 inflammasome through a pathway requiring ROS production and the Syk tyrosine kinase. *PLoS ONE* 5, e10008. doi:10.1371/journal.pone.0010008

Salgado, P. S., Yan, R., Taylor, J. D., Burchell, L., Jones, R., Hoyer, L. L., Matthews, S. J., Simpson, P. J., and Cota, E. (2011). Structural basis for the broad specificity to host-cell ligands by the pathogenic fungus *Candida albicans*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15775–15779.

Samaranayake, D. P., and Hanes, S. D. (2011). Milestones in *Candida albicans* gene manipulation. *Fungal Genet. Biol.* 48, 858–865.

Santamaría, R., Rizzetto, L., Bromley, M., Zelante, T., Lee, W., Cavalieri, D., Romani, L., Miller, B., Gut, I., Santos, M., Pierre, P., Bowyer, P., and Kapushesky, M. (2011). Systems biology of infectious diseases: a focus on fungal infections. *Immunobiology* 216, 1212–1227.

Schaller, M., Korting, H. C., Borelli, C., Hamm, G., and Hube, B. (2005). *Candida albicans*-secreted aspartic proteinases modify the epithelial cytokine response in an in vitro model of vaginal candidiasis. *Infect. Immun.* 73, 2758–2765.

Schrettl, M., Bignell, E., Kragl, C., Joechl, C., Rogers, T., Arst, H. N., Haynes, K., and Haas, H. (2004). Siderophore biosynthesis but not reductive iron assimilation is essential for *Aspergillus fumigatus* virulence. *J. Exp. Med.* 200, 1213–1219.

Schrettl, M., Kim, H. S., Eisendle, M., Kragl, C., Nierman, W. C., Heinekamp, T., Werner, E. R., Jacobsen, I., Illmer, P., Yi, H., Brakhage, A. A., and Haas, H. (2008). SreA-mediated iron regulation in *Aspergillus fumigatus*. *Mol. Microbiol.* 70, 27–43.

Shamir, L., Delaney, J. D., Orlov, N., Eckley, D. M., and Goldberg, I. G. (2010). Pattern recognition software and techniques for biological image analysis. *PLoS Comput. Biol.* 6, e1000974. doi:10.1371/journal.pcbi.1000974

Sharan, R., Ulitsky, I., and Shamir, R. (2007). Network-based prediction of protein function. *Mol. Syst. Biol.* 3, 88.

Shin, Y.-K., Kim, K.-Y., and Paik, Y.-K. (2005). Alterations of protein expression in macrophages in response to *Candida albicans* infection. *Mol. Cells* 20, 271–279.

Skrabanek, L., Saini, H. K., Bader, G. D., and Enright, A. J. (2008). Computational prediction of protein-protein interactions. *Mol. Biotechnol.* 38, 1–17.

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432.

Sorgo, A. G., Heilmann, C. J., Dekker, H. L., Bekker, M., Brul, S., de Koster, C. G., de Koning, L. J., and Klis, F. M. (2011). Effects of fluconazole on the secretome, the wall proteome, and wall integrity of the clinical fungus *Candida albicans*. *Eukaryotic Cell* 10, 1071–1081.

Sosinska, G. J., de Groot, P. W. J., de Mattos, M. J. T., Dekker, H. L., de Koster, C. G., Hellingwerf, K. J., and Klis, F. M. (2008). Hypoxic conditions and iron restriction affect the cell-wall proteome of *Candida albicans* grown under vagina-simulative conditions. *Microbiology* 154(Pt 2), 510–520.

Sosinska, G. J., de Koning, L. J., de Groot, P. W. J., Manders, E. M. M., Dekker, H. L., Hellingwerf, K. J., de Koster, C. G., and Klis, F. M. (2011). Mass spectrometric quantification of the adaptations in the wall proteome of *Candida albicans* in response to ambient pH. *Microbiology* 157(Pt 1), 136–146.

Stoldt, V. R., Sonneborn, A., Leuker, C. E., and Ernst, J. F. (1997). Efg1p, an essential regulator of morphogenesis of the human pathogen *Candida albicans*, is a member of a conserved class of bHLH proteins regulating morphogenetic processes in fungi. *EMBO J.* 16, 1982–1991.

Stolovitzky, G. (2011). "Wisdom of crowds for gene network inference," in *12th International Conference on Systems Biology*, August 28 to September 1st, 2011, Heidelberg.

Strömbäck, L., and Lambrix, P. (2005). Representations of molecular pathways: an evaluation of SBML, PSI MI and BioPAX. *Bioinformatics* 21, 4401–4407.

Sugui, J. A., Kim, H. S., Zarember, K. A., Chang, Y. C., Gallin, J. I., Nierman, W. C., and Kwon-Chung, K. J. (2008). Genes differentially expressed in conidia and hyphae of *Aspergillus fumigatus* upon exposure to human neutrophils. *PLoS ONE* 3, e2655. doi:10.1371/journal.pone.0002655

Sysko, L. R., and Davis, M. A. (2010). From image to data using common image-processing techniques. *Curr. Protoc. Cytom.* Chapter 12, Unit 12.21.

Szabo, E. K., and MacCallum, D. M. (2011). The contribution of mouse models to our understanding of systemic candidiasis. *FEMS Microbiol. Lett.* 320, 1–8.

Taylor, C. F., Field, D., Sansone, S.-A., Aerts, J., Apweiler, R., Ashburner, M., Ball, C. A., Binz, P.-A., Bogue, M., Booth, T., Brazma, A., Brinkman, R. R., Clark, A. M., Deutsch, E. W., Fiehn, O., Fostel, J., Ghazal, P., Gibson, F., Gray, T., Grimes, G., Hancock, J. M., Hardy, N. W., Hermjakob, H., Julian, R. K., Kane, M., Kettner, C., Kinsinger, C., Kolker, E., Kuiper, M., Novère, N. L., Leebens-Mack, J., Lewis, S. E., Lord, P., Mallon, A.-M., Marthandan, N., Masuya, H., McNally, R., Mehrle, A., Morrison, N., Orchard, S., Quackenbush, J., Reecy, J. M., Robertson, D. G., Rocca-Serra, P., Rodriguez, H., Rosenfelder, H., Santoyo-Lopez, J., Scheuermann, R. H., Schober, D., Smith, B., Snape, J., Stoeckert, C. J., Tipton, K., Sterk, P., Untergasser, A., Vandesompele, J., and Wiemann, S. (2008). Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nat. Biotechnol.* 26, 889–896.

Teutschbein, J., Albrecht, D., Pötsch, M., Guthke, R., Aimanianda, V., Clavaud, C., Latgé, J.-P., Brakhage, A. A., and Kniemeyer, O. (2010). Proteome profiling and functional classification of intracellular proteins from conidia of the human-pathogenic mold *Aspergillus fumigatus*. *J. Proteome Res.* 9, 3427–3442.

Theiss, S., Köhler, G. A., Kretschmar, M., Nichterlein, T., and Hacker, J. (2002). New molecular methods to study gene functions in *Candida* infections. *Mycoses* 45, 345–350.

Thomas, D. P., Bachmann, S. P., and Lopez-Ribot, J. L. (2006). Proteomics for the analysis of the *Candida albicans* biofilm lifestyle. *Proteomics* 6, 5795–5804.

Thywißen, A., Heinekamp, T., Dahse, H.-M., Schmaler-Ripcke, J., Nietzsche, S., Zipfel, P. F., and Brakhage, A. A. (2011). Conidial dihydroxynaphthalene melanin of the human pathogenic fungus *Aspergillus fumigatus* interferes with the host endocytosis pathway. *Front. Microbiol.* 2, 96. doi:10.3389/fmicb.2011.00096

Tierney, L., Linde, J., Müller, S., Brunke, S., Molina, J., Hube, B., Schöck, U., Guthke, R., and Kuchler, K. (2012). An interspecies regulatory network inferred from simultaneous RNA-seq of *Candida albicans* invading innate immune cells. *Front. Microbiol.* 3, 85. doi:10.3389/fmicb.2012.00085

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

Tong, J. C., and Ng, L. F. P. (2011). Understanding infectious agents from an in silico perspective. *Drug Discov. Today* 16, 42–49.

Tuckwell, D., Denning, D. W., and Bowyer, P. (2011). A public resource for metabolic pathway mapping of *Aspergillus fumigatus* Af293. *Med. Mycol.* 49(Suppl 1), S114–S119.

Urban, C. F., Reichard, U., Brinkmann, V., and Zychlinsky, A. (2006). Neutrophil extracellular traps capture and kill *Candida albicans* yeast and hyphal forms. *Cell. Microbiol.* 8, 668–676.

Van Dyke Parunak, H., Savit, R., and Riolo, R. (1998). "Agent-based modeling vs. equation-based modeling: a case study and users' guide," in *Multi-Agent Systems and Agent-Based Simulation*, eds J. Sichman, R. Conte, and N. Gilbert, *Lecture Notes in Computer Science*, Vol. 1534 (Berlin: Springer), 277–283.

van het Hoog, M., Rast, T. J., Martchenko, M., Grindle, S., Dignard, D., Hogues, H., Cuomo, C., Berriman, M., Scherer, S., Magee, B. B., Whiteway, M., Chibana, H., Nantel, A., and Magee, P. T. (2007). Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. *Genome Biol.* 8, R52.

Vera, J., Rateitschak, K., Lange, F., Kossow, C., Wolkenhauer, O., and Jaster, R. (2011). Systems biology of JAK-STAT signalling in human malignancies. *Prog. Biophys. Mol. Biol.* 106, 426–434.

Vödisch, M., Albrecht, D., Lessing, F., Schmidt, A. D., Winkler, R., Guthke, R., Brakhage, A. A., and Kniemeyer, O. (2009). Two-dimensional proteome reference maps for the human pathogenic filamentous fungus *Aspergillus fumigatus*. *Proteomics* 9, 1407–1415.

Vödisch, M., Scherlach, K., Winkler, R., Hertweck, C., Braun, H.-P., Roth, M., Haas, H., Werner, E. R., Brakhage, A. A., and Kniemeyer, O. (2011). Analysis of the *Aspergillus fumigatus* proteome reveals metabolic changes and the activation of the pseurotin A biosynthesis gene cluster in response to hypoxia. *J. Proteome Res.* 10, 2508–2524.

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

4, e1000014. doi:10.1371/journal.pcbi.1000014

Volling, K., Thywissen, A., Brakhage, A. A., and Saluz, H. P. (2011). Phagocytosis of melanized *Aspergillus conidia* by macrophages exerts cytoprotective effects by sustained PI3K/Akt signalling. *Cell. Microbiol.* 13, 1130–1148.

Walker, D. C., and Southgate, J. (2009). The virtual cell – a candidate coordinator for 'middle-out' modelling of biological systems. *Brief. Bioinformatics* 10, 450–461.

Waltermann, C., and Klipp, E. (2010). Signal integration in budding yeast. *Biochem. Soc. Trans.* 38, 1257–1264.

Walther, M., Gebhardt, P., Grosse-Gehling, P., Würbach, L., Irmler, I., Preusche, S., Khalid, M., Opfermann, T., Kamradt, T., Steinbach, J., and Saluz, H.-P. (2011). Implementation of 89Zr production and in vivo imaging of B-cells in mice with 89Zr-labeled anti-B-cell antibodies by small animal PET/CT. *Appl. Radiat. Isot.* 69, 852–857.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63.

Wartenberg, D., Lapp, K., Jacobsen, I. D., Dahse, H.-M., Kniemeyer, O., Heinekamp, T., and Brakhage, A. A. (2011). Secretome analysis of *Aspergillus fumigatus* reveals Asp-hemolysin as a major secreted protein. *Int. J. Med. Microbiol.* 301, 602–611.

Whiteway, M., and Bachewich, C. (2007). Morphogenesis in *Candida albicans*. *Annu. Rev. Microbiol.* 61, 529–553.

Willard, H., and Ginsburg, G. (2009). *Genomic and Personalized Medicine. Number Vol. 2 in Genomic and Personalized Medicine.* Amsterdam: Academic Press.

Willger, S. D., Puttikamonkul, S., Kim, K.-H., Burritt, J. B., Grahl, N., Metzler, L. J., Barbuch, R., Bard, M., Lawrence, C. B., and Cramer, R. A. (2008). A sterol-regulatory element binding protein is required for cell polarity, hypoxia adaptation, azole drug resistance, and virulence in *Aspergillus fumigatus*. *PLoS Pathog.* 4, e1000200. doi:10.1371/journal.ppat.1000200

Wilson, D., Thewes, S., Zakikhany, K., Fradin, C., Albrecht, A., Almeida, R., Brunke, S., Grosse, K., Martin, R., Mayer, F., Leonhardt, I., Schild, L., Seider, K., Skibbe, M., Slesiona, S., Waechtler, B., Jacobsen, I., and Hube, B. (2009). Identifying infection-associated genes of *Candida albicans* in the postgenomic era. *FEMS Yeast Res.* 9, 688–700.

Wisplinghoff, H., Bischoff, T., Tallent, S. M., Seifert, H., Wenzel, R. P., and Edmond, M. B. (2004). Nosocomial bloodstream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study. *Clin. Infect. Dis.* 39, 309–317.

Woolhouse, M. E., Taylor, L. H., and Haydon, D. T. (2001). Population biology of multihost pathogens. *Science* 292, 1109–1112.

Xu, J., and Li, Y. (2006). Discovering disease-genes by topological features in human protein-protein interaction network. *Bioinformatics* 22, 2800–2805.

Yan, L., Zhang, J.-D., Cao, Y.-B., Gao, P.-H., and Jiang, Y.-Y. (2007). Proteomic analysis reveals a metabolism shift in a laboratory fluconazole-resistant *Candida albicans* strain. *J. Proteome Res.* 6, 2248–2256.

Yin, Z., Stead, D., Walker, J., Selway, L., Smith, D. A., Brown, A. J. P., and Quinn, J. (2009). A proteomic analysis of the salt, cadmium and peroxide stress responses in *Candida albicans* and the role of the Hog1 stress-activated MAPK in regulating the stress-induced proteome. *Proteomics* 9, 4686–4703.

Zhou, D., and He, Y. (2008). Extracting interactions between proteins from the literature. *J. Biomed. Inform.* 41, 393–407.

Zhu, M., Gao, L., Li, X., Liu, Z., Xu, C., Yan, Y., Walker, E., Jiang, W., Su, B., Chen, X., and Lin, H. (2009). The analysis of the drug-targets based on the topological properties in the human protein-protein interaction network. *J. Drug Target* 17, 524–532.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 December 2011; accepted: 05 March 2012; published online: 02 April 2012.

*Citation: Horn F, Heinekamp T, Kniemeyer O, Pollmächer J, Valiante V and Brakhage AA (2012) Systems biology of fungal infection. Front. Microbio. **3**:108. doi: 10.3389/fmicb.2012.00108*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

REVIEW ARTICLE
published: 06 February 2012
doi: 10.3389/fmicb.2012.00028

Iron – a key nexus in the virulence of *Aspergillus fumigatus*

Hubertus Haas*

Division of Molecular Biology/Biocenter, Innsbruck Medical University, Innsbruck, Austria

Edited by:
Reinhard Guthke, Hans Knöll Institute, Germany

Reviewed by:
Sven Krappmann, Research Center for Infectious Diseases, Germany
Sean Doyle, National University of Ireland Maynooth, Ireland

****Correspondence:***
Hubertus Haas, Division of Molecular Biology/Biocenter, Innsbruck Medical University, Fritz-Pregl-Str. 3, A-6020 Innsbruck, Austria.
e-mail: hubertus.haas@i-med.ac.at

Iron is an essential but, in excess, toxic nutrient. Therefore, fungi evolved fine-tuned mechanisms for uptake and storage of iron, such as the production of siderophores (low-molecular mass iron-specific chelators). In *Aspergillus fumigatus*, iron starvation causes extensive transcriptional remodeling involving two central transcription factors, which are interconnected in a negative transcriptional feed-back loop: the GATA-factor SreA and the bZip-factor HapX. During iron sufficiency, SreA represses iron uptake, including reductive iron assimilation and siderophore-mediated iron uptake, to avoid toxic effects. During iron starvation, HapX represses iron-consuming pathways, including heme biosynthesis and respiration, to spare iron and activates synthesis of ribotoxin AspF1 and siderophores, the latter partly by ensuring supply of the precursor, ornithine. In accordance with the expression pattern and mode of action, detrimental effects of inactivation of SreA and HapX are confined to growth during iron sufficiency and iron starvation, respectively. Deficiency in HapX, but not SreA, attenuates virulence of *A. fumigatus* in a murine model of aspergillosis, which underlines the crucial role of adaptation to iron limitation in virulence. Consistently, production of both extra and intracellular siderophores is crucial for virulence of *A. fumigatus*. Recently, the sterol regulatory element binding protein SrbA was found to be essential for adaptation to iron starvation, thereby linking regulation of iron metabolism, ergosterol biosynthesis, azole drug resistance, and hypoxia adaptation.

Keywords: iron, virulence, fungi, siderophore, isoprenoid, ergosterol, mevalonate, ornithine

INTRODUCTION

Iron is an essential nutrient for all eukaryotes and nearly all prokaryotes (Kaplan and Kaplan, 2009). As mono or diiron center as well as incorporated into heme or iron–sulfur clusters, this metal is an indispensable cofactor for a variety of cellular processes including electron transport, amino acid metabolism, and biosynthesis of DNA and sterols. Nevertheless, iron excess has the potential to catalyze the formation of cell-damaging reactive oxygen species (Halliwell and Gutteridge, 1984). The complex intertwining of iron metabolism and oxidative stress is emphasized by the iron-dependence of detoxification of oxidative stress as, e.g., catalases and peroxidases require heme as cofactor. Despite its high abundance in the Earth's crust, the bioavailability of iron is low owing to its oxidation into sparingly soluble ferric (Fe^{3+}) hydroxides by atmospheric oxygen. To ensure iron supply but to avoid iron toxicity, all organisms evolved sophisticated mechanisms to balance acquisition, storage, and consumption of iron. The control over access to iron is one of the central battlefields during infection as pathogens have to "steal" the iron from the host. Moreover, the mammalian innate immune system restricts access to iron by pathogens via a variety of mechanisms (Ganz, 2009; Weinberg, 2009).

Aspergillus fumigatus is a ubiquitous saprophytic fungus, which has become the most common air-borne fungal pathogen of humans (Tekaia and Latge, 2005). Clinical manifestations range from allergic reactions to life-threatening invasive disease, termed aspergillosis, particularly in immuno-compromised patients. The identification and functional characterization of 24 genes that are involved in iron homeostasis in *A. fumigatus* and/or *Aspergillus nidulans*, respectively, revealed significant insights into iron metabolism and its regulation (**Table 1**). *A. nidulans* is a less virulent *A. fumigatus* relative and longstanding genetic model organism. Inactivation of 10 of the 19 *A. fumigatus* genes caused defects in virulence. All of the virulence-associated genes are transcriptionally upregulated during iron starvation and encode functions that are important for survival during iron starvation, which emphasizes the crucial role of adaptation to iron starvation in virulence. This review summarizes the current knowledge on iron homeostasis and its role in virulence in *Aspergillus* spp.

IRON ACQUISITION

As microorganisms are believed to lack mechanisms for iron excretion, control of iron uptake is considered the major iron homeostatic mechanism (Haas et al., 2008). In contrast to various bacterial and fungal pathogens (Ratledge and Dover, 2000; Almeida et al., 2009), both *A. fumigatus* and *A. nidulans* lack systems for direct uptake of host iron sources such as heme, ferritin, or transferrin (Eisendle et al., 2003; Schrettl et al., 2004a).

Abbreviations: CBC, CCAAT-binding complex; FC, ferricrocin; FsC, fusarinine C; RIA, reductive iron assimilation; SB, siderophore biosynthesis; SIT, siderophore-iron transporter; TAFC, triacetylfusarinine C; TF, transcription factor.

Table 1 | Functionally analyzed proteins involved in iron homeostasis in *A. fumigatus* and *A. nidulans*.

Protein[1]	Gene	Function	Gene deletion-caused defect in iron metabolism[4]	Expression[2]	Virulence[3]	Reference
ENZYMES/TRANSPORTERS INVOLVED IN RIA (3)						
FetC	AFUA_5G03790	Ferroxidase	RIA	−Fe	+	Schrettl et al. (2004a)
FreB	AFUA_1G17270	Ferric reductase	RIA	−Fe	na	Blatzer et al. (2011b)
FtrA	AFUA_5G03800	Iron permease	RIA	−Fe	+	Schrettl et al. (2004a)
ENZYMES INVOLVED IN SB (10)						
EstA	AFUA_3G03660	TAFC esterase	TAFC hydrolysis after uptake (partial)	−Fe	na	Kragl et al. (2007)
NpgA/ PptA§	AFUA_2G08590	Phosphopantetheinyl transferase	Entire SB (all NRPS and polyketide synthetases)	–	na	Oberegger et al. (2003)
SidA§	AFUA_2G07680	Ornithine monooxygenase	Entire SB	−Fe	–	Schrettl et al. (2004a)
SidC§	AFUA_1G17200	FC NRPS	FC biosynthesis	−Fe	±	Schrettl et al. (2007)
SidD	AFUA_3G03420	FSC NRPS	FSC and TAFC biosynthesis	−Fe	±	Schrettl et al. (2007)
SidF	AFUA_3G03400	Transacylase	FSC and TAFC biosynthesis	−Fe	±	Schrettl et al. (2007)
SidG	AFUA_3G03650	Transacetylase	TAFC biosynthesis (but increased FSC biosynthesis)	−Fe	±	Schrettl et al. (2007)
SidH	AFUA_3G03410	Mevalonyl-CoA hydratase	FSC and TAFC biosynthesis	−Fe	±	Yasmin et al. (2011)
SidI	AFUA_1G17190	Mevalonyl-CoA ligase	FSC and TAFC biosynthesis	−Fe	±	Yasmin et al. (2011)
SidL	AFUA_1G04450	Transacetylase	FC biosynthesis (partial)	–	na	Blatzer et al. (2011c)
SITs (2)						
MirA	AN7800	Enterobactin transporter	na	−Fe	na	Haas et al. (2003)
MirB	AN8540	TAFC transporter	na	−Fe	na	Haas et al. (2003)
REGULATORY PROTEINS (9)						
AcuM°	AFUA_2G12330	Zn_2Cys_6 TF	Repression of iron uptake including SB and RIA	–	±	Liu et al. (2010)
HapB*	AN7545	Subunit of the CBC	See HapX	–	na	Hortschansky et al. (2007)
HapC*	AN4034	Subunit of the CBC	See HapX	–	na	Hortschansky et al. (2007)
HapE*	AN6492	Subunit of the CBC	See HapX	–	na	Hortschansky et al. (2007)
HapX§	AFUA_5G03920	bZip-TF	Repression of iron consumption, activation of iron uptake	−Fe	±	Schrettl et al. (2010a)
MpkA	AFUA_4G13720	MAP kinase A	Repression of SB	na	na	Jain et al. (2011)
PacC*	AFUA_3G11970	$(Cys_2His_2)_3$ TF	Activation of TAFC biosynthesis in alkaline pH	na	na	Eisendle et al. (2004)
SrbA	AFUA_2G01260	bHLH-LZ TF	Activation of iron uptake including SB and RIA	−Fe	–	Blatzer et al. (2011a)
SreA§	AFUA_5G11260	GATA TF	Repression of iron uptake including SB and RIA	+Fe	+	Schrettl et al. (2008)

[1] *Unmarked, function analyzed only in A. fumigatus; *, function analyzed only in A. nidulans; §, function is conserved in A. nidulans and A. fumigatus; °, iron regulatory function found only in A. fumigatus but not in A. nidulans.*

[2] *−Fe, transcriptional upregulation during iron starvation; +Fe, transcriptional upregulation during iron sufficiency; –, constitutively expressed.*

[3] *+, Virulent; ±, partially attenuated virulent; −, avirulent.*

[2,3,4] *na, not analyzed.*

Both *Aspergillus* species employ low-affinity ferrous (Fe^{2+}) iron acquisition as well as siderophore-assisted iron uptake, a high-affinity ferric iron uptake system (Eisendle et al., 2003; Schrettl et al., 2004a). In contrast to *A. nidulans*, *A. fumigatus* possesses a second high-affinity iron uptake system, termed reductive iron assimilation (RIA). Schemes of the mechanisms for iron uptake and storage employed by *Aspergillus* spp. are found in recently published reviews (Haas et al., 2008; Schrettl and Haas, 2011).

LOW-AFFINITY IRON UPTAKE

At the molecular level, low-affinity iron uptake has been characterized exclusively in *Saccharomyces cerevisiae*. The relevant permeases are not specific for ferrous iron but additionally transport

other metals, such as copper and zinc (Kaplan and Kaplan, 2009).

REDUCTIVE IRON ASSIMILATION

Reductive iron assimilation starts with reduction of ferric iron sources to the more soluble ferrous iron by plasma membrane-localized metalloreductases (Kosman, 2010). *A. fumigatus* encodes 15 putative metalloreductases indicating possible redundancy of this enzyme system. The metalloreductase FreB has recently been shown to be involved in RIA (Blatzer et al., 2011b). The reduced ferrous iron is re-oxidized and imported by a protein complex consisting of the ferroxidase FetC and the iron permease FtrA (Schrettl et al., 2004a).

SIDEROPHORE-MEDIATED IRON UPTAKE

The siderophores produced by *A. nidulans* and *A. fumigatus* are shown in **Figure 1**. Both fungal species excrete two siderophores, fusarinine C (FsC) and its derivative triacetylfusarinine C (TAFC), to mobilize extracellular iron. The ferri-forms of FsC and TAFC are taken up by siderophore-iron transporters (SIT), which constitute a subfamily of the major facilitator protein superfamily. SIT act most likely as proton symporters energized by the plasma membrane potential (Haas et al., 2003; Philpott and Protchenko, 2008). SIT-mediated iron uptake appears to be universally conserved in the fungal kingdom, even in species not producing siderophores such as *S. cerevisiae*, *Candida* spp., and *Cryptococcus neoformans* (Schrettl et al., 2004b; Haas et al., 2008; Jung and Kronstad, 2008; Philpott and Protchenko, 2008; Nevitt and Thiele, 2011). A likely reason is the dramatically increased solubility and therefore bioavailability of iron chelated by siderophores. Moreover, siderophores might play a role in microbial warfare as chelation of environmental iron by siderophore-types that are not recognized by competitors might be used to starve competitors of iron. This is counteracted by evolving transporters that recognize xenosiderophores, i.e., siderophores that are not produced by the organism, which enables "stealing" of siderophores. This scenario is supported by the fact that most siderophore-producing bacteria and fungi possess xenosiderophore-specific SITs (Haas et al., 2008). *A. fumigatus* and *A. nidulans* encode 10 and 7 putative SITs, respectively (Haas et al., 2008). Heterologous expression in a *S. cerevisiae* mutant lacking high-affinity iron uptake indicated that the *A. nidulans* SITs MirA and MirB transport the bacterial siderophore enterobactin and TAFC, respectively (Haas et al., 2003).

After uptake, the intracellular release of iron from TAFC and FsC involves hydrolysis of the siderophore backbones by the esterase EstB (Kragl et al., 2007).

IRON STORAGE

Extracellular siderophores are employed by most bacterial and some plant species. In contrast, intracellular siderophores are only found in siderophore-producing fungi. *A. fumigatus* produces two different intracellular siderophores (**Figure 1**), ferricrocin (FC) for hyphal iron storage and distribution and its derivative hydroxyferricrocin (HFC) for conidial iron storage (Schrettl et al., 2007; Wallner et al., 2009). *A. nidulans* lacks HFC and employs FC for both hyphal and conidial iron storage (Eisendle et al., 2003). Additionally, the iron-inducible expression of CccA, the ortholog of the vacuolar iron importer Ccc1p from *S. cerevisiae* (Kaplan and Kaplan, 2009), indicates vacuolar iron storage in *A. fumigatus* (Schrettl et al., 2008). In contrast to bacteria, plants, and animals, fungi lack ferritin-mediated iron storage and detoxification.

SIDEROPHORE BIOSYNTHESIS

Fusarinine C consists of three N^5-anhydromevalonyl-N^5-hydroxyornithine residues cyclically linked by ester bonds. TAFC is the N^2-acetylated FsC. FC is a cyclic hexapeptide with the structure Gly-Ser-Gly-(N^5-acetyl-N^5-hydroxyornithine)$_3$ and HFC is the hydroxylated FC (Haas et al., 2008). The siderophore biosynthesis (SB) pathway was characterized by reverse genetics and is shown in **Figure 2**. The first committed step in the biosynthesis of all four siderophores is the hydroxylation of ornithine catalyzed by the ornithine monooxygenase SidA (Eisendle et al., 2003; Schrettl et al., 2004a). Subsequently, the pathways for biosynthesis of extra and intracellular siderophores split. For extracellular SB the transacylase SidF transfers anhydromevalonyl to hydroxyornithine (Schrettl et al., 2007). The required anhydromevalonyl-CoA moiety is derived from mevalonate by CoA-ligation and

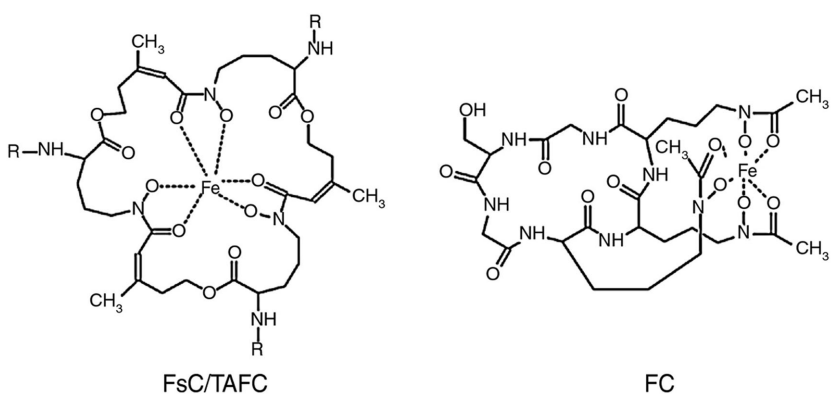


FIGURE 1 | Siderophores produced by *A. fumigatus* and *A. nidulans*: R = H in fusarinine C (FsC) and R = acetyl in triacetylfusarinine C (TAFC); the hydroxylation site in hydroxyferricrocin (HFC) is unknown. The siderophores are shown in the ferric (Fe^{3+}) form.

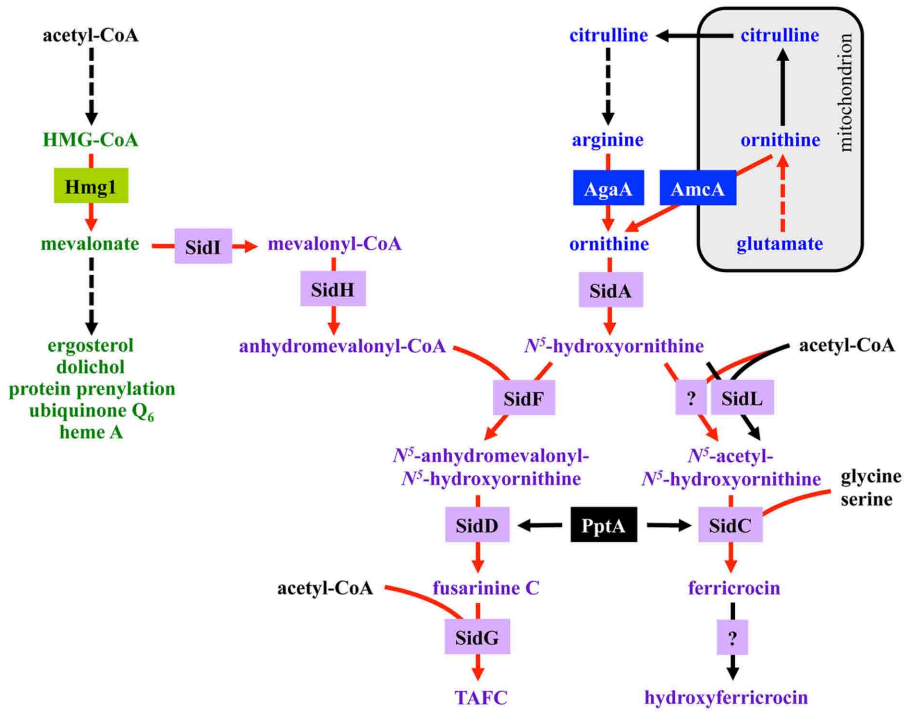


FIGURE 2 | Siderophore biosynthesis (in purple) and its links to the isoprenoid biosynthesis (in green) as well as ornithine/arginine metabolism (in blue). The enzymes, boxed in respective colors, are described in the text and **Table 1**. Enzymatic steps transcriptionally upregulated during iron starvation are marked by red arrows. Broken arrows denote reactions involving more than one enzyme.

dehydration catalyzed by SidI and SidH, respectively (Yasmin et al., 2011). The acetylation of hydroxyornithine for intracellular SB involves two transacetylases, the constitutively expressed SidL and an unidentified enzyme, the activity of which is upregulated by iron starvation (Blatzer et al., 2011c). Assembly of FsC and FC is catalyzed by two different non-ribosomal peptide synthetases (NRPS), SidD and SidC, respectively. TAFC and HFC are formed by SidG-mediated N^2-acetylation of FsC and hydroxylation of FC, respectively (Eisendle et al., 2003; Schrettl et al., 2007). NRPS, polyketide synthases, and the lysine-biosynthetic α-aminoadipate reductase depend on activation by the 4′-phosphopantetheinyl transferase. Consistently, the 4′-phosphopantetheinyl transferases NpgA and PptA are essential for SB in *A. nidulans* and *A. fumigatus*, respectively (Oberegger et al., 2003; Allen et al., 2011). Consequently, it is not surprising that this enzyme was found to be indispensable for phytopathogenic fungi (Horbach et al., 2009).

GENOMIC ORGANIZATION OF GENES INVOLVED IN HIGH-AFFINITY IRON ACQUISITION

Most of the described SB genes above are organized in three gene clusters containing additional genes encoding putative siderophore transporters, the TAFC esterase EstB and genes with uncharacterized functions (Schrettl et al., 2008, 2010a). Exceptions are the genes encoding SidA, SidL, and NpgA/PptA. Notably, SidL and NpgA/PptA are special for another reason as their expression is, in contrast to the other identified structural components of siderophore metabolism, not regulated by iron availability (Oberegger et al., 2003; Blatzer et al., 2011c). Moreover, NpgA/PptA is not exclusively involved in SB. In contrast to *A. fumigatus* and *A. nidulans*, SidA ortholog-encoding genes are clustered with siderophore NRPS in various other fungi, e.g., *Ustilago maydis* and *Neurospora crassa* (Haas et al., 2008). The genes encoding FtrA and FetC form a gene cluster with a common promoter region.

DEFECTS CAUSED BY SIDEROPHORE-DEFICIENCY

Genetic elimination of extracellular siderophores (Δ*sidF* and Δ*sidD* mutants) decreases growth, conidiation, and oxidative stress resistance during iron limitation but not during iron sufficiency, which enables compensation by other iron acquisition systems (Schrettl et al., 2007). Elimination of intracellular siderophores (Δ*sidC* mutant) reduces conidiation and blocks sexual development (as shown in *A. nidulans*) due to the role of FC in intracellular iron transport from substrate-contacting hyphae into aerial hyphae (Eisendle et al., 2006; Schrettl et al., 2007; Wallner et al., 2009). FC-deficiency decreases the conidial iron content by about 50%, which impairs iron-dependent enzymes such as aconitase and catalase A, and thereby decreases conidial size and conidial resistance to oxidative stress (Schrettl et al., 2007; Wallner et al., 2009). Moreover, the lack of FC-mediated iron storage (Δ*sidC* mutants) delays germination during iron starvation (Schrettl et al., 2007). Inactivation of the entire siderophore system (Δ*sidA* mutant) combines the defects caused by inactivation of either extra or intracellular SB and renders *A. fumigatus* extremely sensitive to iron starvation (Schrettl et al., 2004a, 2007).

Both extra and intracellular siderophores are crucial for virulence as elimination of the entire SB (Δ*sidA* mutant) results in absolute avirulence of *A. fumigatus* in a murine model of invasive pulmonary aspergillosis (Schrettl et al., 2004a; Hissen et al., 2005), while deficiency in either extracellular (Δ*sidI*, Δ*sidH*, Δ*sidF*, or Δ*sidD* mutants) or intracellular siderophores (Δ*sidC* mutants) causes partial attenuation of virulence (Schrettl et al., 2007; Yasmin et al., 2011). Conidial FC appears to play a particularly crucial role during initiation of infection because restoration of the conidial HFC content by supplementation with FC during conidiation partially cures the virulence defect of Δ*sidA* conidia (Schrettl et al., 2007). SidG-deficiency, which eliminates TAFC production with concomitant increase of FsC production, affects neither growth nor virulence, indicating that the structural differences between these two siderophores do not play a role in these settings (Schrettl et al., 2007). Consistent with a role in iron acquisition during infection, *A. fumigatus*' siderophores are able to remove iron from host sources, such as transferrin (Hissen et al., 2004; Hissen and Moore, 2005).

Blocking RIA (Δ*ftrA* mutant) does not affect virulence of *A. fumigatus* (Schrettl et al., 2004a). Nevertheless, a putative role of RIA in virulence is indicated by several lines of evidence: (i) elimination of extracellular siderophores causes only partial attenuation of virulence, (ii) mutants lacking both RIA and the siderophore system (Δ*ftrA*Δ*sidA* double mutant) are unable to grow unless supplemented with siderophores or extremely high iron concentrations fueling low-affinity iron uptake (Schrettl et al., 2004a), and (iii) genome-wide expression profiling demonstrated induction of both the siderophore system and RIA during murine infection (McDonagh et al., 2008). Consistently, RIA has been shown to be crucial for virulence of the siderophore-lacking species *C. albicans* and *C. neoformans* (Ramanan and Wang, 2000; Jung et al., 2008).

The siderophore system is important not only for extra, but also for intracellular growth as defects in the siderophore system decrease intracellular growth and survival of *A. fumigatus* after phagocytosis by murine alveolar macrophages, which represent the first line of defense in the lung during pulmonary aspergillosis (Schrettl et al., 2010b). Furthermore, impairment of SB changes the immune response of macrophages after phagocytosis of *A. fumigatus* (Seifert et al., 2008). In agreement, the siderophore system is also critical for virulence of *Histoplasma capsulatum*, a dimorphic fungal pathogen replicating in the yeast form within macrophages (Hwang et al., 2008). The evolutionary conserved role of siderophores in virulence has been confirmed in various other aspergillosis infection models, i.e., a murine cutaneous model, *Drosophila melanogaster*, and *Galleria mellonella* (Ben-Ami et al., 2010; Chamilos et al., 2010; Slater et al., 2011). Moreover, SB is indispensable for the virulence of various phytopathogenic ascomycetes (Oide et al., 2006; Greenshields et al., 2007). In contrast, SB is dispensable and RIA is essential for phytopathogenicity of *U. maydis* (Mei et al., 1993; Eichhorn et al., 2006).

INTEGRATION OF THE SIDEROPHORE-BIOSYNTHETIC PATHWAY IN THE GENERAL METABOLISM

During iron starvation, siderophore production reaches up to 10% of the biomass. In addition, iron starvation dramatically remodels the free amino acid pool of *A. fumigatus* with eight amino acids increasing and three amino acids decreasing more than 1.5-fold (Schrettl et al., 2010a). Among these changes, the approximate sevenfold increase of the siderophore precursor ornithine during iron starvation compared to sufficiency indicates that the enormous ornithine demand for SB is matched by active upregulation of biosynthesis and not by de-repression *via* its consumption. Consistently, blocking siderophore-mediated ornithine consumption by inactivation of SidA (Δ*sidA* mutant) causes a further 2.9-fold increase of the ornithine pool during iron starvation (Schrettl et al., 2010a).

Ornithine is produced in the mitochondrion, and its biosynthetic pathway as well as export via the transporter AmcA to the cytosol is transcriptionally upregulated under iron deprivation (Schrettl et al., 2010a; **Figure 2**). Moreover, ornithine is a precursor of arginine and the conversion of arginine to ornithine in the cytosol by the arginase AgaA is likewise upregulated in response to iron starvation. The metabolic commitment required for siderophore production must however be balanced against the need to maintain other cellular functions as arginine itself is vital for protein biosynthesis. Recently, arginine was found to allosterically activate SidA enzyme activity (Frederick et al., 2011), which appears to connect siderophore production to cellular amino acid homeostasis: SB is stimulated only if the cytosolic arginine pool is sufficient for all of the cell's needs.

Consistent with mevalonate being a precursor for TAFC biosynthesis (**Figure 2**), overexpression of the mevalonate-producing HMG-CoA reductase Hmg1 increases TAFC production, while lovastatin-mediated Hmg1 inhibition blocks TAFC biosynthesis in *A. fumigatus* (Yasmin et al., 2011). Mevalonate is an intermediate of the isoprenoid biosynthetic pathway with ergosterol as the major product, which is one major target of antifungal treatment, i.e., amphotericin B and azoles. In contrast to siderophore production, iron starvation decreases the cellular ergosterol level due to the iron-requirement of ergosterol biosynthesis (Yasmin et al., 2011). Concordant with bilateral demand for mevalonate, blocking mevalonate consumption for TAFC biosynthesis (Δ*sidI* mutant) alters the sterol composition and increases lovastatin resistance during iron starvation (Yasmin et al., 2011). These data demonstrate that statins such as lovastatin, which are widely used as cholesterol-decreasing drugs, have an additional target in siderophore-producing fungi. As SB is crucial for virulence statins might be useful to combat infections with siderophore-producing fungi.

REGULATION OF IRON METABOLISM AND ITS ROLE IN VIRULENCE

Iron starvation has been shown to cause extensive transcriptional remodeling in *A. fumigatus* with about 13% of the genes responding to iron availability (Schrettl et al., 2008, 2010a). As shown in **Figure 3**, in both *A. fumigatus* and *A. nidulans* maintenance of iron homeostasis involves the two central transcription factors (TF) SreA and HapX (Haas et al., 1999; Hortschansky et al., 2007; Schrettl et al., 2008, 2010a). During iron sufficiency, the GATA-factor SreA, postulated by bioinformatic analyses to recognize the consensus sequence ATCWGATAA, represses high-affinity iron uptake, including RIA and the siderophore system, to avoid

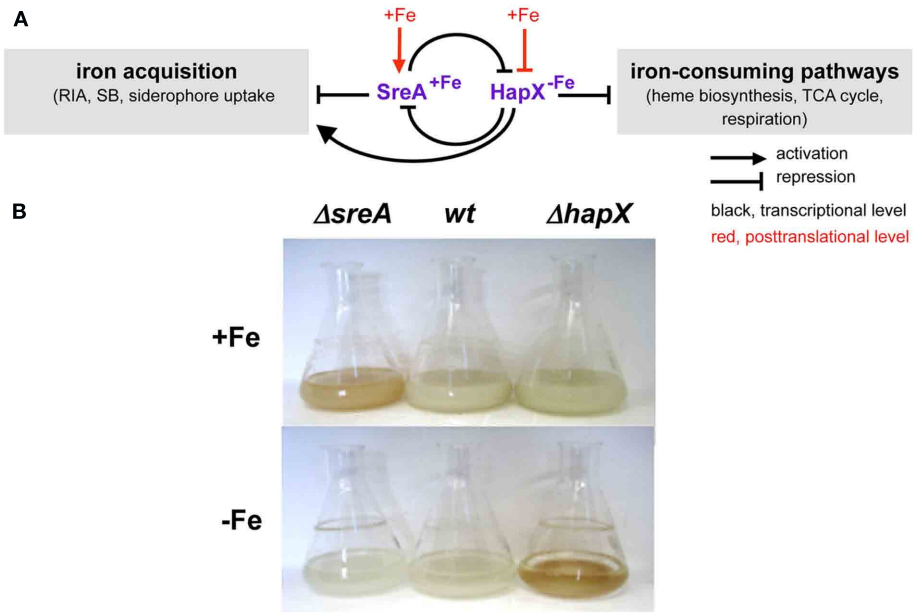


FIGURE 3 | Iron regulation in *Aspergillus* spp. (A) Scheme of SreA- and HapX-mediated iron regulation. (B) Phenotypes of *A. fumigatus* SreA-(Δ*sreA*) and HapX-(Δ*hapX*) deficient mutant strains in 24 h/37°C flask cultures. In contrast to the wild type (*wt*), Δ*hapX* mycelia are reddish colored during iron starvation due to accumulation of protoporphyrin IX, while Δ*sreA* mycelia are reddish colored during iron sufficiency due to accumulation of iron, heme, and FC (Schrettl et al., 2008, 2010a).

iron toxicity (Schrettl et al., 2008). During iron starvation, the bZip-TF HapX represses iron-consuming pathways such as heme biosynthesis, respiration, and ribosome biogenesis to spare iron (Oberegger et al., 2001, 2002; Schrettl et al., 2008, 2010a). Furthermore, HapX activates synthesis of the ribotoxin AspF1 and siderophores in *A. fumigatus*, the latter partly by coordinating SB with supply of its precursor ornithine (Schrettl et al., 2010a). The metabolic deregulation caused by deficiency in either SreA or HapX causes reddish hyphal pigmentation (**Figure 3**). SreA and HapX are interconnected in a negative feed-back loop: SreA represses expression of *hapX* during iron sufficiency, while HapX represses *sreA* during iron starvation. Additionally, both SreA and HapX appear to be regulated posttranslationally by iron blocking HapX function and activating SreA function (Haas et al., 1999; Hortschansky et al., 2007). In *S. pombe*, posttranslational iron sensing by the HapX and SreA orthologs involves the monothiol glutaredoxin Grx4 (Mercier and Labbe, 2009; Jbel et al., 2011; Kim et al., 2011). In *A. nidulans* and *A. fumigatus*, inactivation of both HapX and SreA is synthetically lethal underlining the critical role of iron homeostasis in cellular survival (Hortschansky et al., 2007; Schrettl et al., 2008, 2010a). In agreement with their expression pattern and mode of action, detrimental effects of inactivation of SreA or HapX are confined to growth during iron sufficiency or starvation, respectively (**Figure 3**). Deficiency in HapX, but not SreA, attenuates virulence of *A. fumigatus* in murine models of aspergillosis (Schrettl et al., 2008, 2010a), which emphasizes the crucial role of adaptation to iron limitation in virulence. Most fungal species possess orthologs to SreA and HapX and the important role of HapX orthologs in virulence has been demonstrated in *C. albicans* and *C. neoformans* (Labbe et al., 2007; Haas et al., 2008; Jung et al., 2010; Hsu et al., 2011). Similar to *A. fumigatus*, the SreA ortholog Sfu1 is dispensable for systemic virulence of *C. albicans* (Chen et al., 2011). However, Sfu1 is crucial for persistence of this commensal in the iron-rich gut, which impressively illustrates the importance of adaptation to opposing conditions of iron availability for survival. Remarkably, the fungal prototype *S. cerevisiae* lacks orthologs of SreA, HapX, and SrbA and employs entirely different regulators, Aft1/2 and Cth1/2, which are conserved only in closely related *Saccharomycotina* species (Haas et al., 2008; Kaplan and Kaplan, 2009).

INTERCONNECTION OF IRON METABOLISM WITH OTHER REGULATORY CIRCUITS: OXYGEN, REDOX PH, GLUCONEOGENESIS, MAP KINASE SIGNALING, ZINC

Due to the central metabolic role of iron, a variety of regulatory circuits affect cellular iron handling. As obligate aerobic organisms, *Aspergilli* rely on respiration, which is highly iron-dependent. Therefore, oxygen supply largely influences iron metabolism and *vice versa*. Recently, proteomic analysis of *A. fumigatus* revealed that hypoxia increases the production of proteins involved in glycolysis, the TCA-cycle, and respiration, which is paralleled by increased cellular iron, heme, copper, and zinc contents (Vodisch et al., 2009). The increase in iron/heme is attributable to the iron/heme-dependence of the TCA-cycle and respiration. In *A. fumigatus*, hypoxic adaptation involves SrbA, a member of the "sterol regulatory element binding protein (SREBP)" TF family, which is conserved in most eukaryotes (Willger et al., 2008). These TFs are activated by cellular sterol-depletion to maintain sterol homeostasis (Bien and Espenshade, 2010). In *A. fumigatus*, SrbA-deficiency decreases the cellular ergosterol content

and resistance against azole drugs and blocks hypoxic growth as well as virulence (Willger et al., 2008). Recently, SrbA was found to activate siderophore-mediated iron uptake in response to hypoxia and iron starvation in part by transcriptional activation of HapX (Blatzer et al., 2011a). In agreement with SrbA being involved in iron homeostasis, defects in hypoxic growth and azole resistance caused by SrbA-deficiency are at least partially cured by increased iron supplementation and in particular by de-repression of iron uptake *via* simultaneous inactivation of SreA (Blatzer et al., 2011a). SrbA is transcriptionally upregulated during hypoxia and iron starvation, in both cases likely in response to sterol-depletion and autoregulation as sterol biosynthesis depends on both oxygen and iron. During iron starvation, consumption of the sterol intermediate mevalonate by SB might play an additional role (see above). In agreement, the transcriptional activation of SrbA during iron starvation is independent of SreA and HapX. These data indicate that *A. fumigatus* senses iron not only via HapX and SreA but also via ergosterol biosynthesis and SrbA. Taken together, SrbA coordinates ergosterol biosynthesis and iron metabolism to mediate hypoxia responses and azole resistance. Growth defects during iron starvation of respective gene deletion mutants indicate similar functions in iron homeostasis of the SrbA orthologs in *S. pombe* and *C. neoformans* (Hughes et al., 2005; Chang et al., 2007). Moreover, the SrbA ortholog is also crucial for virulence in *C. neoformans* (Chun et al., 2007). Consequently, the virulence defect of SrbA mutants is possibly attributable not only to the defects in adaptation to hypoxia, but also iron starvation. *S. cerevisiae* lacks an SrbA ortholog and the TFs mediating hypoxic adaptation (Hap1, Mot1, Rox1) are not conserved in *Aspergilli* (Willger et al., 2008), which might be explained by *S. cerevisiae* being an facultative anaerobe.

As shown in *A. nidulans*, HapX functions via physical interaction with the DNA-binding CCAAT-binding complex (CBC; Hortschansky et al., 2007). The CBC is a heterotrimeric DNA-binding complex, which is conserved in all eukaryotes. In *A. nidulans*, inactivation of either one of its subunits, HapB, HapC, or HapE phenocopies HapX inactivation with respect to the defects in adaptation to iron starvation. However, the CBC has HapX-independent functions and is speculated to affect expression of about 30% of all genes. Consistently, CBC-deficiency results in decreased growth and sporulation during both iron sufficiency and starvation (Hortschansky et al., 2007). The mode of discrimination between HapX targets and the remaining CBC targets remains to be elucidated. As mentioned in the introduction, iron metabolism and oxidative stress are intimately intertwined. Therefore it is particularly interesting that, as shown in *A. nidulans*, the CBC senses the redox state of the cell via oxidation and thioredoxin-mediated reduction of evolutionary conserved thiol groups within the HapC histone fold motif (Thon et al., 2010). Oxidation blocks CBC formation and nuclear localization. In line with a role in redox regulation, CBC-deficiency impairs the oxidative stress response. The impact of iron availability on the oxidative stress detoxification system is indicated by iron starvation-mediated transcriptional downregulation of the heme-dependent hyphal catalase B (CatB) and upregulation by the Cu/Zn-superoxide dismutase (SodA; Oberegger et al., 2000, 2001). Excessive iron uptake caused by SreA-deficiency transcriptionally upregulates both enzymes (Oberegger et al., 2001).

The ambient pH impacts iron availability as alkaline conditions decrease iron solubility. In line, neutral compared to acidic conditions upregulate SB and uptake in *A. nidulans* mediated by the pH-reactive TF PacC (Eisendle et al., 2004). Similarly, PacC orthologs mediate upregulation of high-affinity iron uptake – in these species RIA – during alkaline conditions in *S. cerevisiae*, *C. albicans*, and *C. neoformans* (Lamb et al., 2001; Baek et al., 2008). Virulence of *C. albicans* but not *C. neoformans* depends on its PacC ortholog, which is most likely attributable to the occupation of different host niches by these two pathogens (Nobile et al., 2008; O'Meara et al., 2010).

Interestingly, deficiency in the TF AcuM, which is required for gluconeogenesis, decreases siderophore production in *A. fumigatus*, but not in *A. nidulans*, and attenuates virulence in *A. fumigatus* (Liu et al., 2010). Though it is unclear if its effects are direct, AcuM appears to transcriptionally repress SreA.

Recently, iron starvation was found to trigger phosphorylation and nuclear localization of the *A. fumigatus* mitogen-activated protein kinase (MAPK) MpkA, which is involved in maintaining cell wall integrity, protection against ROS, and secondary metabolism (Jain et al., 2011). Moreover, MpkA-deficiency increases siderophore production. The TF targeted by MpkA signaling remains to be identified. Remarkably, despite its dramatic effect on *in vitro* growth rate, MpkA-deficiency does not affect virulence of *A. fumigatus* (Valiante et al., 2008).

Similar to iron, zinc plays a critical role in a diverse array of biochemical processes, but excess of zinc is deleterious. Consequently, regulation of zinc homeostasis by the TF ZafA is essential for virulence of *A. fumigatus*. Iron starvation causes zinc hypersensitive and therefore iron depletion changes cellular zinc handling by downregulating zinc uptake and upregulation of vacuolar zinc detoxification (Yasmin et al., 2009). HapX appears to play a critical role in coordination of zinc and iron homeostasis as its deficiency causes zinc hypersensitivity during iron starvation. These data demonstrate the importance of cellular metal balancing.

CONCLUSION AND PERSPECTIVES

The understanding of the role of iron in fungal pathogenicity has advanced enormously in recent years. Together with the transcriptional upregulation of the high-affinity iron acquisition systems during initiation of murine infection (McDonagh et al., 2008), the attenuated virulence caused by defects in SB or HapX confirms that *A. fumigatus* faces iron limitation during mammalian infection. Thus, human protection against *A. fumigatus* includes growth inhibition by polymorphonuclear leukocytes *via* lactoferrin-mediated iron depletion and possibly siderocalin-mediated scavenging of siderophores (Fluckinger et al., 2004; Zarember et al., 2007). On the other hand, increased bone marrow iron stores represent an independent risk factor for invasive aspergillosis (Kontoyiannis et al., 2007).

The current difficulties in diagnosis and treatment of aspergillosis are reflected by the high mortality rate of this infectious disease (Tekaia and Latge, 2005). The essentiality of iron and the differences in iron handling between mammals and fungi

like *Aspergilli* might help to improve therapy and diagnosis of fungal infections. Specifically, the unique fungal siderophore system represents a promising target for selective therapeutic intervention. Noteworthy, SIT constitute one of few protein families that are unique to fungi (Hsiang and Baillie, 2005). Their absence in prokaryotes and other eukaryotes might enable specific inhibition or drug delivery during infection by a "Trojan horse" approach (Miller et al., 2009), whereby antifungal agents are covalently attached to siderophores and selectively imported by fungi.

Moreover, the enzymatic and regulatory links between siderophore and ergosterol biosynthetic pathways might be crucial for optimization of treatment of infections caused by siderophore-producing fungi. The potential of iron chelation therapy is indicated by the synergistic effect of iron chelators and antifungal drugs demonstrated *in vitro* and in a murine aspergillosis model (Zarember et al., 2009; Ibrahim et al., 2010). Moreover, the recently demonstrated imaging of invasive pulmonary aspergillosis in a rat model, based on positron emission tomography (PET)-visualized fungal accumulation of TAFC-chelated ^{68}Gallium, emphasizes the potential of siderophores in diagnosis of fungal infections (Petrik et al., 2010a,b).

ACKNOWLEDGMENTS

Work in the authors' laboratory is supported by Austrian Science Foundation Grants FWF P-21643-B11 and I-282-B09. Moreover, support by the European Science Foundation via its Research Networking Programme FUMINOMICS is acknowledged.

REFERENCES

Allen, G., Bromley, M., Kaye, S. J., Keszenman-Pereyra, D., Zucchi, T. D., Price, J., Birch, M., Oliver, J. D., and Turner, G. (2011). Functional analysis of a mitochondrial phosphopantetheinyl transferase (PPTase) gene pptB in *Aspergillus fumigatus*. *Fungal Genet Biol.* 48, 456–464.

Almeida, R. S., Wilson, D., and Hube, B. (2009). *Candida albicans* iron acquisition within the host. *FEMS Yeast Res.* 9, 1000–1012.

Baek, Y. U., Li, M., and Davis, D. A. (2008). *Candida albicans* ferric reductases are differentially regulated in response to distinct forms of iron limitation by the Rim101 and CBF transcription factors. *Eukaryotic Cell* 7, 1168–1179.

Ben-Ami, R., Lewis, R. E., Leventakos, K., Latge, J. P., and Kontoyiannis, D. P. (2010). Cutaneous model of invasive aspergillosis. *Antimicrob. Agents Chemother.* 54, 1848–1854.

Bien, C. M., and Espenshade, P. J. (2010). Sterol regulatory element binding proteins in fungi: hypoxic transcription factors linked to pathogenesis. *Eukaryotic Cell* 9, 352–359.

Blatzer, M., Barker, B. M., Willger, S. D., Beckmann, N., Blosser, S. J., Cornish, E. J., Mazurie, A., Grahl, N., Haas, H., and Cramer, R. A. (2011a). SREBP coordinates iron and ergosterol homeostasis to mediate triazole drug and hypoxia responses in the human fungal pathogen *Aspergillus fumigatus*. *PLoS Genet.* 7, e1002374. doi:10.1371/journal.pgen.1002374

Blatzer, M., Binder, U., and Haas, H. (2011b). The metalloreductase FreB is involved in adaptation of *Aspergillus fumigatus* to iron starvation. *Fungal Genet. Biol.* 48, 1027–1033.

Blatzer, M., Schrettl, M., Sarg, B., Lindner, H. H., Pfaller, K., and Haas, H. (2011c). SidL, an *Aspergillus fumigatus* transacetylase involved in biosynthesis of the siderophores ferricrocin and hydroxyferricrocin. *Appl. Environ. Microbiol.* 77, 4959–4966.

Chamilos, G., Bignell, E. M., Schrettl, M., Lewis, R. E., Leventakos, K., May, G. S., Haas, H., and Kontoyiannis, D. P. (2010). Exploring the concordance of *Aspergillus fumigatus* pathogenicity in mice and Toll-deficient flies. *Med. Mycol.* 48, 506–510.

Chang, Y. C., Bien, C. M., Lee, H., Espenshade, P. J., and Kwon-Chung, K. J. (2007). Sre1p, a regulator of oxygen sensing and sterol homeostasis, is required for virulence in *Cryptococcus neoformans*. *Mol. Microbiol.* 64, 614–629.

Chen, C., Pande, K., French, S. D., Tuch, B. B., and Noble, S. M. (2011). An iron homeostasis regulatory circuit with reciprocal roles in *Candida albicans* commensalism and pathogenesis. *Cell Host Microbe* 10, 118–135.

Chun, C. D., Liu, O. W., and Madhani, H. D. (2007). A link between virulence and homeostatic responses to hypoxia during infection by the human fungal pathogen *Cryptococcus neoformans*. *PLoS Pathog.* 3, e22. doi:10.1371/journal.ppat.0030022

Eichhorn, H., Lessing, F., Winterberg, B., Schirawski, J., Kamper, J., Muller, P., and Kahmann, R. (2006). A ferroxidation/permeation iron uptake system is required for virulence in *Ustilago maydis*. *Plant Cell* 18, 3332–3345.

Eisendle, M., Oberegger, H., Buttinger, R., Illmer, P., and Haas, H. (2004). Biosynthesis and uptake of siderophores is controlled by the PacC-mediated ambient-pH Regulatory system in *Aspergillus nidulans*. *Eukaryotic Cell* 3, 561–563.

Eisendle, M., Oberegger, H., Zadra, I., and Haas, H. (2003). The siderophore system is essential for viability of *Aspergillus nidulans*: functional analysis of two genes encoding l-ornithine N 5-monooxygenase (sidA) and a non-ribosomal peptide synthetase (sidC). *Mol. Microbiol.* 49, 359–375.

Eisendle, M., Schrettl, M., Kragl, C., Muller, D., Illmer, P., and Haas, H. (2006). The intracellular siderophore ferricrocin is involved in iron storage, oxidative-stress resistance, germination, and sexual development in *Aspergillus nidulans*. *Eukaryotic Cell* 5, 1596–1603.

Fluckinger, M., Haas, H., Merschak, P., Glasgow, B. J., and Redl, B. (2004). Human tear lipocalin exhibits antimicrobial activity by scavenging microbial siderophores. *Antimicrob. Agents Chemother.* 48, 3367–3372.

Frederick, R. E., Mayfield, J. A., and Dubois, J. L. (2011). Regulated O2 activation in flavin-dependent monooxygenases. *J. Am. Chem. Soc.* 133, 12338–12341.

Ganz, T. (2009). Iron in innate immunity: starve the invaders. *Curr. Opin. Immunol.* 21, 63–67.

Greenshields, D. L., Liu, G., Feng, J., Selvaraj, G., and Wei, Y. (2007). The siderophore biosynthetic gene SID1, but not the ferroxidase gene FET3, is required for full *Fusarium graminearum* virulence. *Mol. Plant Pathol.* 8, 411–421.

Haas, H., Eisendle, M., and Turgeon, B. G. (2008). Siderophores in fungal physiology and virulence. *Annu. Rev. Phytopathol.* 46, 149–187.

Haas, H., Schoeser, M., Lesuisse, E., Ernst, J. F., Parson, W., Abt, B., Winkelmann, G., and Oberegger, H. (2003). Characterization of the *Aspergillus nidulans* transporters for the siderophores enterobactin and triacetylfusarinine C. *Biochem. J.* 371, 505–513.

Haas, H., Zadra, I., Stoffler, G., and Angermayr, K. (1999). The *Aspergillus nidulans* GATA factor SREA is involved in regulation of siderophore biosynthesis and control of iron uptake. *J. Biol. Chem.* 274, 4613–4619.

Halliwell, B., and Gutteridge, J. M. (1984). Oxygen toxicity, oxygen radicals, transition metals and disease. *Biochem. J.* 219, 1–14.

Hissen, A. H., Chow, J. M., Pinto, L. J., and Moore, M. M. (2004). Survival of *Aspergillus fumigatus* in serum involves removal of iron from transferrin: the role of siderophores. *Infect. Immun.* 72, 1402–1408.

Hissen, A. H., and Moore, M. M. (2005). Site-specific rate constants for iron acquisition from transferrin by the *Aspergillus fumigatus* siderophores N′,N″,N‴-triacetylfusarinine C and ferricrocin. *J. Biol. Inorg. Chem.* 10, 211–220.

Hissen, A. H., Wan, A. N., Warwas, M. L., Pinto, L. J., and Moore, M. M. (2005). The *Aspergillus fumigatus* siderophore biosynthetic gene sidA, encoding L-ornithine N5-oxygenase, is required for virulence. *Infect. Immun.* 73, 5493–5503.

Horbach, R., Graf, A., Weihmann, F., Antelo, L., Mathea, S., Liermann, J. C., Opatz, T., Thines, E., Aguirre, J., and Deising, H. B. (2009). Sfp-type 4′-phosphopantetheinyl transferase is indispensable for fungal pathogenicity. *Plant Cell* 21, 3379–3396.

Hortschansky, P., Eisendle, M., Al-Abdallah, Q., Schmidt, A. D., Bergmann, S., Thon, M., Kniemeyer, O., Abt, B., Seeber, B., Werner, E. R., Kato, M., Brakhage, A. A., and Haas, H. (2007). Interaction of HapX with the CCAAT-binding complex – a novel mechanism of gene regulation by iron. *EMBO J.* 26, 3157–3168.

Hsiang, T., and Baillie, D. L. (2005). Comparison of the yeast proteome to other fungal genomes to find core fungal genes. *J. Mol. Evol.* 60, 475–483.

Hsu, P. C., Yang, C. Y., and Lan, C. Y. (2011). *Candida albicans* Hap43 is a repressor induced under low-iron conditions and essential for iron-responsive transcriptional regulation and virulence. *Eukaryotic Cell* 10, 207–225.

Hughes, A. L., Todd, B. L., and Espenshade, P. J. (2005). SREBP pathway responds to sterols and functions as an oxygen sensor in fission yeast. *Cell* 120, 831–842.

Hwang, L. H., Mayfield, J. A., Rine, J., and Sil, A. (2008). Histoplasma requires SID1, a member of an iron-regulated siderophore gene cluster, for host colonization. *PLoS Pathog.* 4, e1000044. doi:10.1371/journal.ppat.1000044

Ibrahim, A. S., Gebremariam, T., French, S. W., Edwards, J. E. Jr., and Spellberg, B. (2010). The iron chelator deferasirox enhances liposomal amphotericin B efficacy in treating murine invasive pulmonary aspergillosis. *J. Antimicrob. Chemother.* 65, 289–292.

Jain, R., Valiante, V., Remme, N., Docimo, T., Heinekamp, T., Hertweck, C., Gershenzon, J., Haas, H., and Brakhage, A. A. (2011). The MAP kinase MpkA controls cell wall integrity, oxidative stress response, gliotoxin production and iron adaptation in *Aspergillus fumigatus*. *Mol. Microbiol.* 82, 39–53.

Jbel, M., Mercier, A., and Labbe, S. (2011). Grx4 monothiol glutaredoxin is required for iron limitation-dependent inhibition of Fep1. *Eukaryotic Cell* 10, 629–645.

Jung, W. H., and Kronstad, J. W. (2008). Iron and fungal pathogenesis: a case study with *Cryptococcus neoformans*. *Cell. Microbiol.* 10, 277–284.

Jung, W. H., Saikia, S., Hu, G., Wang, J., Fung, C. K., D'souza, C., White, R., and Kronstad, J. W. (2010). HapX positively and negatively regulates the transcriptional response to iron deprivation in *Cryptococcus neoformans*. *PLoS Pathog.* 6, e1001209. doi: 10.1371/journal.ppat.1001209

Jung, W. H., Sham, A., Lian, T., Singh, A., Kosman, D. J., and Kronstad, J. W. (2008). Iron source preference and regulation of iron uptake in *Cryptococcus neoformans*. *PLoS Pathog.* 4, e45. doi:10.1371/journal.ppat.0040045

Kaplan, C. D., and Kaplan, J. (2009). Iron acquisition and transcriptional regulation. *Chem. Rev.* 109, 4536–4552.

Kim, K. D., Kim, H. J., Lee, K. C., and Roe, J. H. (2011). Multi-domain CGFS-type glutaredoxin Grx4 regulates iron homeostasis via direct interaction with a repressor Fep1 in fission yeast. *Biochem. Biophys. Res. Commun.* 408, 609–614.

Kontoyiannis, D. P., Chamilos, G., Lewis, R. E., Giralt, S., Cortes, J., Raad, I. I., Manning, J. T., and Han, X. (2007). Increased bone marrow iron stores is an independent risk factor for invasive aspergillosis in patients with high-risk hematologic malignancies and recipients of allogeneic hematopoietic stem cell transplantation. *Cancer* 110, 1303–1306.

Kosman, D. J. (2010). Redox cycling in iron uptake, efflux, and trafficking. *J. Biol. Chem.* 285, 26729–26735.

Kragl, C., Schrettl, M., Abt, B., Sarg, B., Lindner, H. H., and Haas, H. (2007). EstB-mediated hydrolysis of the siderophore triacetylfusarinine C optimizes iron uptake of *Aspergillus fumigatus*. *Eukaryotic Cell* 6, 1278–1285.

Labbe, S., Pelletier, B., and Mercier, A. (2007). Iron homeostasis in the fission yeast *Schizosaccharomyces pombe*. *Biometals* 20, 523–537.

Lamb, T. M., Xu, W., Diamond, A., and Mitchell, A. P. (2001). Alkaline response genes of *Saccharomyces cerevisiae* and their relationship to the RIM101 pathway. *J. Biol. Chem.* 276, 1850–1856.

Liu, H., Gravelat, F. N., Chiang, L. Y., Chen, D., Vanier, G., Ejzykowicz, D. E., Ibrahim, A. S., Nierman, W. C., Sheppard, D. C., and Filler, S. G. (2010). *Aspergillus fumigatus* AcuM regulates both iron acquisition and gluconeogenesis. *Mol. Microbiol.* 78, 1038–1054.

McDonagh, A., Fedorova, N. D., Crabtree, J., Yu, Y., Kim, S., Chen, D., Loss, O., Cairns, T., Goldman, G., Armstrong-James, D., Haynes, K., Haas, H., Schrettl, M., May, G., Nierman, W. C., and Bignell, E. (2008). Sub-telomere directed gene expression during initiation of invasive aspergillosis. *PLoS Pathog.* 4, e1000154. doi:10.1371/journal.ppat.1000154

Mei, B., Budde, A. D., and Leong, S. A. (1993). sid1, a gene initiating siderophore biosynthesis in *Ustilago maydis*: molecular characterization, regulation by iron, and role in phytopathogenicity. *Proc. Natl. Acad. Sci. U.S.A.* 90, 903–907.

Mercier, A., and Labbe, S. (2009). Both Php4 function and subcellular localization are regulated by iron via a multistep mechanism involving the glutaredoxin Grx4 and the exportin Crm1. *J. Biol. Chem.* 284, 20249–20262.

Miller, M. J., Zhu, H., Xu, Y., Wu, C., Walz, A. J., Vergne, A., Roosenberg, J. M., Moraski, G., Minnick, A. A., Mckee-Dolence, J., Hu, J., Fennell, K., Kurt Dolence, E., Dong, L., Franzblau, S., Malouin, F., and Mollmann, U. (2009). Utilization of microbial iron assimilation processes for the development of new antibiotics and inspiration for the design of new anticancer agents. *Biometals* 22, 61–75.

Nevitt, T., and Thiele, D. J. (2011). Host iron withholding demands siderophore utilization for *Candida glabrata* to survive macrophage killing. *PLoS Pathog.* 7, e1001322. doi:10.1371/journal.ppat.1001322

Nobile, C. J., Solis, N., Myers, C. L., Fay, A. J., Deneault, J. S., Nantel, A., Mitchell, A. P., and Filler, S. G. (2008). *Candida albicans* transcription factor Rim101 mediates pathogenic interactions through cell wall functions. *Cell. Microbiol.* 10, 2180–2196.

Oberegger, H., Eisendle, M., Schrettl, M., Graessle, S., and Haas, H. (2003). 4′-phosphopantetheinyl transferase-encoding npgA is essential for siderophore biosynthesis in *Aspergillus nidulans*. *Curr. Genet.* 44, 211–215.

Oberegger, H., Schoeser, M., Zadra, I., Abt, B., and Haas, H. (2001). SREA is involved in regulation of siderophore biosynthesis, utilization and uptake in *Aspergillus nidulans*. *Mol. Microbiol.* 41, 1077–1089.

Oberegger, H., Zadra, I., Schoeser, M., Abt, B., Parson, W., and Haas, H. (2002). Identification of members of the *Aspergillus nidulans* SREA regulon: genes involved in siderophore biosynthesis and utilization. *Biochem. Soc. Trans.* 30, 781–783.

Oberegger, H., Zadra, I., Schoeser, M., and Haas, H. (2000). Iron starvation leads to increased expression of Cu/Zn-superoxide dismutase in *Aspergillus*. *FEBS Lett.* 485, 113–116.

Oide, S., Moeder, W., Haas, H., Krasnoff, S., Gibson, D., Yoshioka, K., and Turgeon, B. G. (2006). NPS6, encoding a non-ribosomal peptide synthetase involved in siderophore-mediated iron metabolism, is a conserved virulence determinant of plant pathogenic ascomycetes. *Plant Cell* 18, 2836–2853.

O'Meara, T. R., Norton, D., Price, M. S., Hay, C., Clements, M. F., Nichols, C. B., and Alspaugh, J. A. (2010). Interaction of *Cryptococcus neoformans* Rim101 and protein kinase A regulates capsule. *PLoS Pathog.* 6, e1000776. doi:10.1371/journal.ppat.1000776

Petrik, M., Haas, H., Dobrozemsky, G., Lass-Florl, C., Helbok, A., Blatzer, M., Dietrich, H., and Decristoforo, C. (2010a). 68Ga-siderophores for PET imaging of invasive pulmonary aspergillosis: proof of principle. *J. Nucl. Med.* 51, 639–645.

Petrik, M., Haas, H., Schrettl, M., Helbok, A., Blatzer, M., and Decristoforo, C. (2010b). In vitro and in vivo evaluation of selected (68)Ga-siderophores for infection imaging. *Nucl. Med. Biol.* doi: 10.1016/j.nucmedbio.2011.09.012. [Epub ahead of print].

Philpott, C. C., and Protchenko, O. (2008). Response to iron deprivation in *Saccharomyces cerevisiae*. *Eukaryotic Cell* 7, 20–27.

Ramanan, N., and Wang, Y. (2000). A high-affinity iron permease essential for *Candida albicans* virulence. *Science* 288, 1062–1064.

Ratledge, C., and Dover, L. G. (2000). Iron metabolism in pathogenic bacteria. *Annu. Rev. Microbiol.* 54, 881–941.

Schrettl, M., Beckmann, N., Varga, J., Heinekamp, T., Jacobsen, I. D., Jochl, C., Moussa, T. A., Wang, S., Gsaller, F., Blatzer, M., Werner, E. R., Niermann, W. C., Brakhage, A. A., and Haas, H. (2010a). HapX-mediated adaptation to iron starvation is crucial for virulence of *Aspergillus fumigatus*. *PLoS Pathog.* 6, e1001124. doi:10.1371/journal.ppat.1001124

Schrettl, M., Ibrahim-Granet, O., Droin, S., Huerre, M., Latge, J. P., and Haas, H. (2010b). The crucial role of the *Aspergillus fumigatus* siderophore system in interaction with alveolar macrophages. *Microbes Infect.* 12, 1035–1041.

Schrettl, M., Bignell, E., Kragl, C., Joechl, C., Rogers, T., Arst, H. N. Jr., Haynes, K., and Haas, H. (2004a). Siderophore biosynthesis but not reductive iron assimilation is essential for *Aspergillus fumigatus* virulence. *J. Exp. Med.* 200, 1213–1219.

Schrettl, M., Winkelmann, G., and Haas, H. (2004b). Ferrichrome in *Schizosaccharomyces pombe* – an iron transport and iron storage compound. *Biometals* 17, 647–654.

Schrettl, M., Bignell, E., Kragl, C., Sabiha, Y., Loss, O., Eisendle, M., Wallner, A., Arst, H. N., Haynes, K., and Haas, H. (2007). Distinct roles for intra and extracellular siderophores during *Aspergillus fumigatus* infection. *PLoS Pathog.* 3, e128. doi:10.1371/journal.ppat.0030128

Schrettl, M., and Haas, H. (2011). Iron homeostasis – Achilles' heel of

Aspergillus fumigatus? *Curr. Opin. Microbiol.* 14, 400–405.

Schrettl, M., Kim, H. S., Eisendle, M., Kragl, C., Nierman, W. C., Heinekamp, T., Werner, E. R., Jacobsen, I., Illmer, P., Yi, H., Brakhage, A. A., and Haas, H. (2008). SreA-mediated iron regulation in *Aspergillus fumigatus*. *Mol. Microbiol.* 70, 27–43.

Seifert, M., Nairz, M., Schroll, A., Schrettl, M., Haas, H., and Weiss, G. (2008). Effects of the *Aspergillus fumigatus* siderophore systems on the regulation of macrophage immune effector pathways and iron homeostasis. *Immunobiology* 213, 767–778.

Slater, J. L., Gregson, L., Denning, D. W., and Warn, P. A. (2011). Pathogenicity of *Aspergillus fumigatus* mutants assessed in *Galleria mellonella* matches that in mice. *Med. Mycol.* 49(Suppl. 1), S107–S113.

Tekaia, F., and Latge, J. P. (2005). *Aspergillus fumigatus*: saprophyte or pathogen? *Curr. Opin. Microbiol.* 8, 385–392.

Thon, M., Al Abdallah, Q., Hortschansky, P., Scharf, D. H., Eisendle, M., Haas, H., and Brakhage, A. A. (2010). The CCAAT-binding complex coordinates the oxidative stress response in eukaryotes. *Nucleic Acids Res.* 38, 1098–1113.

Valiante, V., Heinekamp, T., Jain, R., Hartl, A., and Brakhage, A. A. (2008). The mitogen-activated protein kinase MpkA of *Aspergillus fumigatus* regulates cell wall signaling and oxidative stress response. *Fungal Genet. Biol.* 45, 618–627.

Vodisch, M., Albrecht, D., Lessing, F., Schmidt, A. D., Winkler, R., Guthke, R., Brakhage, A. A., and Kniemeyer, O. (2009). Two-dimensional proteome reference maps for the human pathogenic filamentous fungus *Aspergillus fumigatus*. *Proteomics* 9, 1407–1415.

Wallner, A., Blatzer, M., Schrettl, M., Sarg, B., Lindner, H., and Haas, H. (2009). Ferricrocin, a siderophore involved in intra and transcellular iron distribution in *Aspergillus fumigatus*. *Appl. Environ. Microbiol.* 75, 4194–4196.

Weinberg, E. D. (2009). Iron availability and infection. *Biochim. Biophys. Acta* 1790, 600–605.

Willger, S. D., Puttikamonkul, S., Kim, K. H., Burritt, J. B., Grahl, N., Metzler, L. J., Barbuch, R., Bard, M., Lawrence, C. B., and Cramer, R. A. Jr. (2008). A sterol-regulatory element binding protein is required for cell polarity, hypoxia adaptation, azole drug resistance, and virulence in *Aspergillus fumigatus*. *PLoS Pathog.* 4, e1000200. doi:10.1371/journal.ppat.1000200

Yasmin, S., Abt, B., Schrettl, M., Moussa, T. A., Werner, E. R., and Haas, H. (2009). The interplay between iron and zinc metabolism in *Aspergillus fumigatus*. *Fungal Genet. Biol.* 46, 707–713.

Yasmin, S., Alcazar-Fuoli, L., Grundlinger, M., Puempel, T., Cairns, T., Blatzer, M., Lopez, J. F., Grimalt, J. O., Bignell, E., and Haas, H. (2011). Mevalonate governs interdependency of ergosterol and siderophore biosyntheses in the fungal pathogen *Aspergillus fumigatus*. *Proc. Natl. Acad. Sci. U.S.A.* doi: 10.1073/pnas.1106399108. [Epub ahead of print].

Zarember, K. A., Cruz, A. R., Huang, C. Y., and Gallin, J. I. (2009). Antifungal activities of natural and synthetic iron chelators alone and in combination with azole and polyene antibiotics against *Aspergillus fumigatus*. *Antimicrob. Agents Chemother.* 53, 2654–2656.

Zarember, K. A., Sugui, J. A., Chang, Y. C., Kwon-Chung, K. J., and Gallin, J. I. (2007). Human polymorphonuclear leukocytes inhibit *Aspergillus fumigatus* conidial growth by lactoferrin-mediated iron depletion. *J. Immunol.* 178, 6367–6373.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 December 2011; paper pending published: 03 January 2012; accepted: 16 January 2012; published online: 06 February 2012.

Citation: Haas H (2012) Iron – a key nexus in the virulence of Aspergillus fumigatus. Front. Microbio. ***3****:28. doi: 10.3389/fmicb.2012.00028*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

REVIEW ARTICLE
published: 03 May 2012
doi: 10.3389/fmicb.2012.00164

Salmonella enterica: a surprisingly well-adapted intracellular lifestyle

Thomas Dandekar[1]*, Astrid Fieselmann[1], Jasmin Popp[2] and Michael Hensel[2]

[1] Department of Bioinformatics, Biocenter, University of Würzburg, Würzburg, Germany

[2] Division of Microbiology, University of Osnabrück, Osnabrück, Germany

Edited by:
Reinhard Guthke, Leibniz Institute for Natural Product Research and Infection Biology - Hans-Knoell-Institute, Germany

Reviewed by:
Michael Lalk, University of Greifswald, Germany
Karsten Tedin, Freie Universität Berlin, Germany

***Correspondence:**
Thomas Dandekar, Department of Bioinformatics, Biocenter, University of Würzburg, Am Hubland, 97074 Würzburg, Germany. e-mail: dandekar@biozentrum.uni-wuerzburg.de

The infectious intracellular lifestyle of *Salmonella enterica* relies on the adaptation to nutritional conditions within the *Salmonella*-containing vacuole (SCV) in host cells. We summarize latest results on metabolic requirements for *Salmonella* during infection. This includes intracellular phenotypes of mutant strains based on metabolic modeling and experimental tests, isotopolog profiling using ^{13}C-compounds in intracellular *Salmonella*, and complementation of metabolic defects for attenuated mutant strains towards a comprehensive understanding of the metabolic requirements of the intracellular lifestyle of *Salmonella*. Helpful for this are also genomic comparisons. We outline further recent studies and which analyses of intracellular phenotypes and improved metabolic simulations were done and comment on technical required steps as well as progress involved in the iterative refinement of metabolic flux models, analyses of mutant phenotypes, and isotopolog analyses. *Salmonella* lifestyle is well-adapted to the SCV and its specific metabolic requirements. *Salmonella* metabolism adapts rapidly to SCV conditions, the metabolic generalist *Salmonella* is quite successful in host infection.

Keywords: metabolism, *Salmonella*-containing vacuole, regulation, virulence

INTRODUCTION

Salmonella enterica serovar Typhimurium (*S.* Typhimurium) is an important human gastrointestinal pathogen with an invasive and facultative intracellular lifestyle (Neidhardt, 1996; Eisenreich et al., 2010). Among the various habitats that can be colonized by *Salmonella*, the adaptation to life inside the host cell is of specific interest, since this ability is considered as crucial for systemic infections with fatal outcome. The World Health Organization estimated 1.4 million cases of non-typhoidal *Salmonella* infections. Furthermore, these cause 580 deaths annually even in the United States (World Health Organization, 2005). Infections are often associated with selected subgroups as elderly or patients suffering from HIV and connective tissue disorders (Cummings et al., 2010).

Throughout the intracellular life, *Salmonella* remains in a membrane-bound compartment, which is termed *Salmonella*-containing vacuole or SCV. The SCV is probably a unique compartment that is formed by the combined action of a large number of bacterial virulence factors (**Figure 1**). Virulent *Salmonellae* are able to modify this vacuole in order to escape killing in the endocytic pathway, and to proliferate within host cells (Haraga et al., 2008). The ability to survive and replicate within host cells is closely related to the systemic pathogenesis of *Salmonella* in a murine model of typhoid fever. Mutant strains defective in intracellular replication due to auxotrophies are also attenuated in virulence in an animal murine model of typhoid fever (Fields et al., 1986). *Salmonella* is able to rapidly multiply in various eukaryotic cell lines, but the proliferation appears to be far less rapid within cells in tissues of infected hosts, indicating a more restrictive situation *in vivo* (Mastroeni et al., 2009). The SCV is commonly considered as a nutritional deprived environment, and this notion is based on the phenotypes of auxotrophic strains, analyses of bacterial reporter strains, and microarray analyses. However, the fact that *Salmonella* replicates within the SCV indicates the successful adaptation to this intracellular environment.

Despite the remarkable increase in understanding of the cellular microbiology of *Salmonella* infections and the molecular functions of virulence factors required for intracellular life, the nutritional basis of life of *Salmonella* within the SCV is still not completely understood.

Understanding how *Salmonella* survives and thrives within this compartment and how nutrients are acquired is not only essential for the understanding of the intracellular lifestyle, but might as well open new avenues to therapeutic interference with *Salmonella* infections.

METHODS AND APPROACHES TO GET INSIGHTS INTO INTRACELLULAR NUTRITION

To target metabolism of intracellular *Salmonella*, several approaches have been established covering *in vitro* approaches from analyzing simple growth behavior in full or minimal medium and changes in morphology (Paterson et al., 2009), intracellular replication ability (Bowden et al., 2009, 2010) to more complex transcriptome analysis (Eriksson et al., 2003) or ^{13}C-isotopolog profiling analysis (^{13}C-IPA; Götz and Goebel, 2010) in a cell culture model. Experiments to test if a gene of interest contributes to virulence are mainly done in macrophage cell lines as strains unable to replicate within macrophages proved to be avirulent (Fields et al., 1986). In addition, epithelial cell infection models,

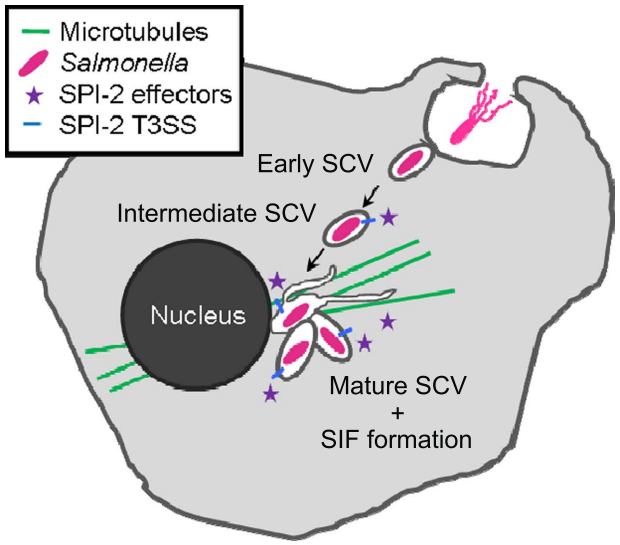


FIGURE 1 | Intracellular lifestyle of *Salmonella*. *Salmonella* is taken up by host cells either by *Salmonella*-induced invasion (T3SS-1 triggered macropinocytosis) or via phagocytosis. By translocating effector proteins via the T3SS-2 into the host cell, the SCV undergoes an altered maturation process characterized by specific endosomal markers. The onset of *Salmonella*-induced filaments (SIFs) formation coincides with the start of *Salmonella* replication between 4 and 6 h after infection. SIFs develop on a microtubule scaffold.

such as HeLa or CaCo-2 cells (Götz and Goebel, 2010) which are targeted first in *Salmonella* infection are under investigation.

The importance of certain metabolic functions and their corresponding pathways can be investigated by intracellular replication assays performed in murine macrophage cell lines with *Salmonella* wild-type (WT) and mutant strains with metabolic defects (Bowden et al., 2009; Lim et al., 2010). Mutant strains can be easily generated by the λ-Red mediated mutagenesis approach (Datsenko and Wanner, 2000). The intracellular replication ability of these mutant strains compared to the WT strain indicates the importance of the respective metabolic pathway for intracellular nutrition and, in turn, intracellular survival and replication. Other useful ways to "scan" for the complete setup of essential genes and metabolic enzymes are transcriptome and proteome analyses of *Salmonella* strains isolated from infected murine macrophage cell lines (Eriksson et al., 2003; Shi et al., 2006) giving direct evidence of genes and proteins expressed under intracellular conditions. The application of ^{13}C-IPA to follow metabolic fluxes in the host and the bacteria cells *in vitro* is on the rise (**Figure 2**). It has been widely used for analyzing metabolic fluxes in different bacterial species like *Escherichia coli* (Fischer and Sauer, 2003) or intracellular pathogens like *Salmonella*, enteroinvasive *E. coli* (EIEC; Götz and Goebel, 2010), *Listeria* (Eylert et al., 2008), and *Legionella* (Eylert et al., 2010).

In vitro approaches provide essential clues about the nutritional status of intracellular *Salmonella*. However, the gained *in vitro* data are often supported by, or compared to, data received from animal models, e.g., mouse (Bowden et al., 2009). A large number of different mutant strains with metabolic defects has been tested for virulence in mouse models (Tchawa Yimga et al., 2006; Bowden et al., 2009; Paterson et al., 2009) and a proteome analysis with *Salmonella* isolated from infected mice (recovered from cecum and spleen) has been performed (Becker et al., 2006) providing information about the essential metabolic enzymes and pathways. Data on metabolite levels would provide important complementary information, but are often difficult to obtain for intracellular bacteria and require complex experimental setups. Another complementary approach is proteome analysis, a modern technique to directly determine enzyme type and amount as well as modifications (e.g., regulatory phosphorylation of metabolic enzymes). The fact that there may be differences between results obtained by *in vitro* and *in vivo* approaches does not render data from *in vitro* experiments questionable. The data may very well be reliable (i.e., reproducible) and of importance for understanding of a limiting number of specific factors. The differences in results from *in vivo* analyses can best be explained by the presence of a large number of additional factors (immune responses, different concentration gradients of nutrients in different tissues, cytokines, etc.) that are often interrelated and affect pathogen survival and replication in host tissue. For example, comparing experimental evidence including proteomics, metabolomics, and survival data on mutants with *in vivo* and *in vitro* conditions, conflicting results can be observed for mutant strains of *Salmonella* with defects in enzymes of the tricarboxylic acid (TCA) cycle. The mutant strains showed even increased replication in a murine RAW macrophage cell line, but were highly reduced in virulence in an animal model (Bowden et al., 2010).

As they are easier to study and analyze, cell culture models for *Salmonella* infections will remain the essential basis for the understanding of the cellular and molecular changes and mechanisms of the intracellular bacterial nutrition. Cell culture experiments are less complex and laborious than animal experiments and offer the possibility to study such important aspects of intracellular nutrition as access of *Salmonella* to host cell nutrients or to nutrients in extracellular medium. Attenuation of certain mutant strains can be directly linked to the lack of a certain metabolite in the cell culture media. Supplementation experiments with labeled carbon sources (C-sources) allow to analyze if and how this nutrients reach *Salmonella* inside its SCV. Data gained from *in vitro* experiments help to establish and improve mathematical models of *Salmonella* during infection. By contrast, the large number of additional and partially unknown factors affecting the pathogen in *in vivo* infection studies would lead to models of extreme complexity.

SALMONELLA INTRACELLULAR LIFESTYLE IN COMPARISON TO LIFESTYLES OF OTHER INTRACELLULAR PATHOGENS

Salmonella is well-adapted to the SCV compartment as seen if compared to cytoplasmic lifestyle (e.g., *Listeria*) or *Legionella* as a further pathogen in a membrane-bound compartment. In *Salmonella*, glucose represents a major C-source. However, gluconeogenesis rather than glycolysis is observed for cytoplasmic listerial intracellular metabolism. This is required to synthesize glucose from available substrates. Glucose is then predominantly degraded in the pentose phosphate pathway (PPP). A complete TCA cycle is observed for *Salmonella*. However, in cytoplasmic

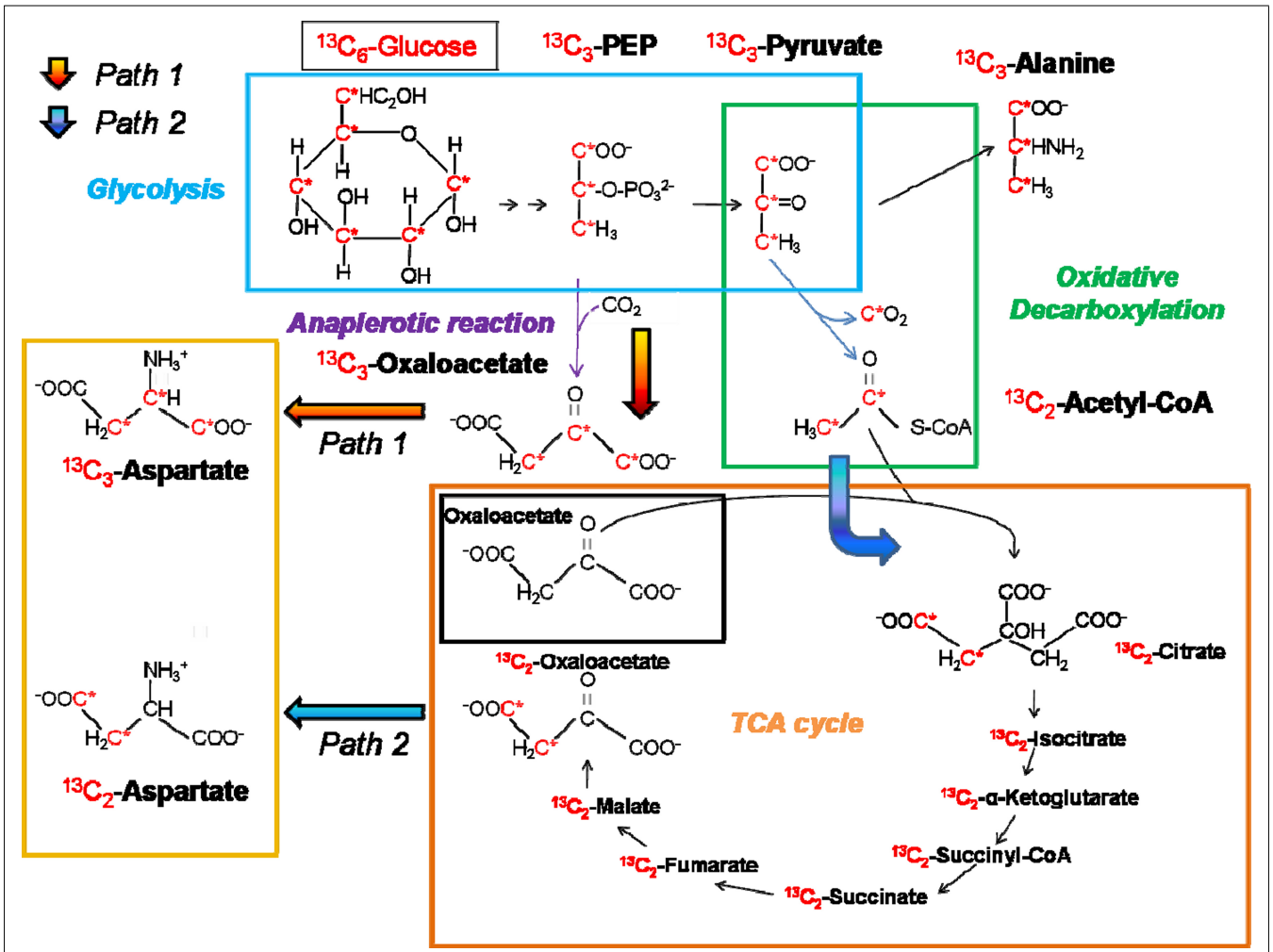


FIGURE 2 | Principles of isotopolog profiling with labeled [U-$^{13}C_6$] glucose. The fate of labeled glucose via different metabolic routes can be followed by isotopolog profiling in analyzing resulting metabolites or products (in this case amino acids). An example is given for the biosynthetic routes of the two aspartate isotopologs $^{13}C_2$-aspartate and $^{13}C_3$-aspartate (represented by orange and blue arrows, respectively) and the origin of a $^{13}C_3$-alanine. Isotopolog studies are described in Eisenreich et al. (2006), Eylert et al. (2008), and Eylert et al. (2010). ^{13}C-atoms are marked in red and indicated by an asterisk.

Listeria the enzyme 2-oxoglutarate dehydrogenase is missing (Glaser et al., 2001; Eisenreich et al., 2006), oxaloacetate is formed by carboxylation of C_3 compounds like pyruvate and the anaplerotic enzyme pyruvate carboxylase becomes crucial (Schar et al., 2010), major C-source is glycerol. *De novo* synthesized listerial amino acids are Ala, Asp, Glu, Ser, Thr, Val, and Gly (Eisenreich et al., 2010), indicating that other amino acids are taken up in the host cell cytoplasm (Schauer et al., 2010).

For comparison, *Legionella pneumophila*, a gram-negative intracellular pathogen and causative agent for Legionnaire's disease was long time supposed to feed solely on amino acids while residing in the host alveolar macrophages (Tesh and Miller, 1981). Indeed, the amino acids Cys, Gln, Ser, and Arg are efficiently used as carbon and energy sources *in vivo* (Wieland et al., 2005). However, recent studies highlight that also glucose is metabolized by *Legionella* during infection of eukaryotic cells. In contrast to *Salmonella*, *Legionella* predominantly degrades glucose by the 2-keto-3-deoxy-phosphogluconate pathway (KDPGP) and only in small quantities by glycolysis. The non-oxidative branch of PPP also accounts for small amounts of glucose catabolism (Harada et al., 2010). ^{13}C-IPA further revealed that there is no evidence for a functional glyoxylate bypass (Eylert et al., 2010), confirming earlier models build on genome sequence analysis (Cazalet et al., 2004). Furthermore, ^{13}C-IPA results indicate that a complete and active TCA cycle occurs in *Legionella* and that the inability to synthesize amino acids *de novo* is only valid for Ile, Leu, Val, Phe, Met, Arg, and Tyr.

COMPUTATIONAL AND SYSTEM BIOLOGICAL ANALYSES OF THE METABOLISM OF INTRACELLULAR *SALMONELLA*

A number of recent studies highlight the importance of system biological modeling of *Salmonella* during infection (Bumann, 2009b; McDermott et al., 2011). Besides the genome information, the analysis of the pathogen proteomes during infection provides

an important basis for infection research, as well as for devising novel control strategies including antibiotics and vaccines (Bumann, 2009a). On this, independent models for *Salmonella* metabolism are built (Becker et al., 2006; AbuOun et al., 2009; Raghunathan et al., 2009; Eisenreich et al., 2010). In such models the flow of metabolites is modeled in terms of pathways. Enzyme chains are calculated such that metabolites are balanced, i.e., consumed and produced in equilibrium, this is called flux balance analysis (FBA). If such a chain of metabolic enzymes can not be dissected any further, this is named an elementary mode. Helpful software tools include the COBRA Toolbox (Schellenberger et al., 2011) and YANAsquare (Schwarz et al., 2007) which can compile available biochemical and genome data systematically and calculate flux distributions by FBA or elementary mode analysis (EMA). Specific tools such as the KEGGbrowser (Schwarz et al., 2007) simplify the direct import of available biochemical data into metabolic network reconstructions even on genome-scale basis. Furthermore, metabolic gaps and dubious annotations for enzymes of central metabolism often occur and therefore different genome annotation software and comparisons are strongly recommended to improve the network reconstruction (Gaudermann et al., 2006).

Recently, a community effort towards a knowledge-base and mathematical model of *S.* Typhimurium strain LT2 has been initiated, resulting in the BiGG knowledge-base of *Salmonella* metabolism (Thiele et al., 2011). A consensus metabolic reconstruction was obtained from two independently developed (Ruppin et al., 2010) metabolic reconstructions for *S.* Typhimurium. The joined reconstruction effort included, furthermore, the development and implementation of a community-based workflow for annotation and corrections including incorporation of thermodynamic information (to decide on reversible and irreversible reactions). By this, metabolite transporters and reactions are more accurately identified and considered. Higher reliable consensus models improve, furthermore, the potential of multi-target drug therapy approaches for specific strains though of course the host response is another important factor to consider.

Our metabolic modeling approach calculating elementary flux modes on central carbon and amino acid metabolism in *S.* Typhimurium indicates that the anaplerotic reactions around phosphoenolpyruvate (PEP) to oxaloacetate are pivotal and occur in many flux modes. For modeling growth on glucose as the sole C-source, PEP carboxylase (*ppc*) plays a central role in directing the flux to the TCA cycle. However, disruption of PEP carboxylase can be partly compensated by alternative routes in the network regarding carbohydrate metabolism. A *ppc*-deficient mutant showed no reduced virulence *in vivo* (Tchawa Yimga et al., 2006), indicating the availability of further C-sources such as amino acids which, through transaminase reactions, also can feed into the TCA cycle. This would give support to a model that contains both glucose and amino acids as C-sources. Here, the flux through PEP carboxylase decreased and the flux through PEP carboxykinase (*pckA*) increased, compared to a medium with only glucose. PEP carboxykinase catalyzes a reaction in the opposite direction of the normal flow cycle, i.e., increasing the flow from oxaloacetate to PEP. The direct conversion of oxaloacetate to aspartate cannot be easily compensated by alternative flux modes. Transaminase activity can be compensated, for instance, loss of two transaminases (*aspC* and *tyrB*) is required before aspartate auxotrophy appears. Moreover, aspartate can also be acquired from the host. However, the production and availability of oxaloacetate is critical, requiring increased PEP carboxylase activity or a reversal of the flow from oxaloacetate to aspartate. This and similar other modeling results suggest that amino acid metabolism is easier impaired and more critical in the SCV than in intracytoplasmic survival (Schauer et al., 2010).

Flux balance analysis and EMA calculate metabolic pathways but the integration of experimental "omics" data allows to better determine metabolic flux strengths (Covert et al., 2004). Tools have thus recently been developed that help to integrate experimental data into metabolic models such as YANAvergence (Liang et al., 2011) fit experimental flux measurements, enzyme activities, gene expression data (Cecil et al., 2011), as well as extracellular metabolite ratios to computational predictions.

Due to the intracellular lifestyle, the host provides further nutrients by transporters (**Figure 3**). Overall, there is an intensive exchange of metabolites during growth in the host cell. Game theoretical approaches treat such interdependencies in a novel mathematical way showing advantages as well as limitations for any specific survival strategy (bacterial pathogen) or opposing strategy from the host (Ruppin et al., 2010; Schuster et al., 2010b, 2011). This becomes even more important as the robust *Salmonella* metabolism including its redundant, overlapping pathway organization limits possibilities for new antimicrobials interfering with its metabolic processes (Becker et al., 2006). Moreover, the potential of central carbon metabolism as a target for microbial defense depends on the environmental factors of the specific niche and the genetic and phenotypic traits of infecting bacteria. For instance, in dormant sub-populations, or "persisters," the uptake of glucose, mannitol, or fructose implies a direct influence in preparatory steps of glycolysis. This potentiates the killing by aminoglycosides (Allison et al., 2011). Thus, for an iterative refinement of computational modeling the integration of transporter reactions (Raghunathan et al., 2009) and a strain-specific analysis (Liang et al., 2011) is crucial. In this regard, a global gene expression analysis of *S.* Typhimurium (Harvey et al., 2011) during colonization of the chicken's cecal lumen and cecal mucosa demonstrates very specific *Salmonella* metabolic adaptations to its environment. For comparison, differences in expression of transporters and in the usage of C-sources regarding three specific niches (cecal lumen, mucosal wall, and the SCV in macrophages and epithelial cells) are illustrated in **Figure 3**.

THE ROLE OF CENTRAL CARBON METABOLISM PATHWAYS DURING INTRACELLULAR SURVIVAL

Salmonella is in fact a pathogen with a very broad and versatile metabolism and as already seen by its comparatively large genome, is a generalist among gram-negative bacteria (**Figure 4**). For instance, *Salmonella* can easily metabolize glucose combining various pathways to supply both energy and amino acids and the same applies for most nutrient sources.

Furthermore, the combination of different pathways guarantees a fine-tuned balance of internal metabolites. Due to this redundancy it is not easy to block growth of *Salmonella* by

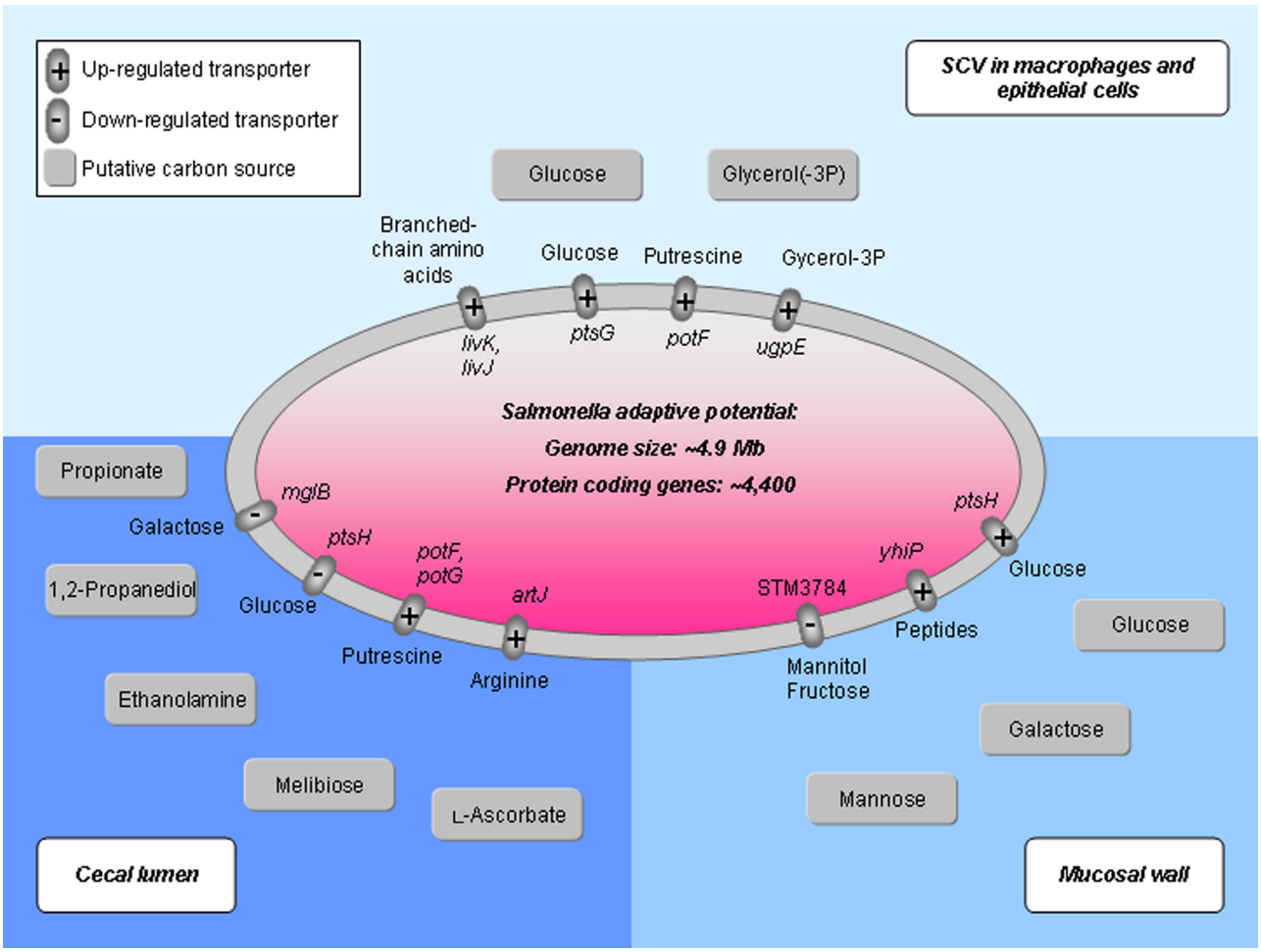


FIGURE 3 | Activity of nutrient transport systems in various host habitats colonized by *Salmonella*. Transporter gene expression and carbon source usage is exemplified according to Hautefort et al. (2008), Eisenreich et al. (2010), and Harvey et al. (2011). During the colonization of the intestine, *Salmonella* first resides in the cecal lumen (bottom left), then attaches to the mucosal layer (bottom right) until it invades macrophages and epithelial cells forming the SCV (upper part). Transporter proteins differentially expressed in the specific niches are represented as small ovals with transported metabolites at the outside of the *Salmonella* cell (central oval) and gene names on the inside. Carbon sources suggested according to literature are illustrated as gray boxes. *Salmonella* genome information is according to McClelland et al. (2001).

antibiotics targeted against key metabolic enzymes of primary metabolism (Becker et al., 2006). In the following we discuss specific *Salmonella* pathways.

THE ROLE OF CATABOLISM OF GLUCOSE

There are three routes for the catabolism of glucose: (i) glycolysis, (ii) the PPP, and (iii) the Entner–Doudoroff pathway also known as the KDPGP. The two latter pathways for glucose utilization seem to be of lesser importance for *Salmonella*. A *Salmonella* mutant strain deficient in *zwf* (encoding glucose-6-phosphate dehydrogenase), catalyzing the first step of both PPP and KDPGP, and a double mutant strain in *gnd* (PPP) and *edd* (KDPGP), or a double mutant in *gnd* (PPP) and *edd* (KDPGP) are not attenuated in proliferation in the murine macrophage cell line RAW 264.7 (own unpublished results). Nevertheless, it was reported that a *zwf* mutant shows reduced virulence in a mouse model of systemic infection (Lundberg et al., 1999). This group refers to the importance of NADPH production in the PPP which is used as electron donor for reductases required for oxidative stress response. However, we think that in our cell culture models the superoxide levels should be less high than in the mouse infection model and this may explain the decreased need for such reductases and for NADPH. Other important PPP products like ribose used for nucleoside synthesis can still be produced by the non-oxidative part of PPP. We suggest that PPP and KDPGP play no important role if glucose is the major substrate. Isotopolog profiling experiments in CaCo-2 cells showed that the internalized glucose is mainly converted by glycolysis and/or KDPGP pathway and excluded PPP as a major route for glucose catabolism (Götz and Goebel, 2010). *Salmonella* studied in epithelial cell lines are significantly less challenged with reactive oxygen intermediates (ROI) and reactive nitrogen intermediates (RNI) responses, hence, the generation of NADPH would be less important than in an animal model. In particular, the activity of the PPP may become

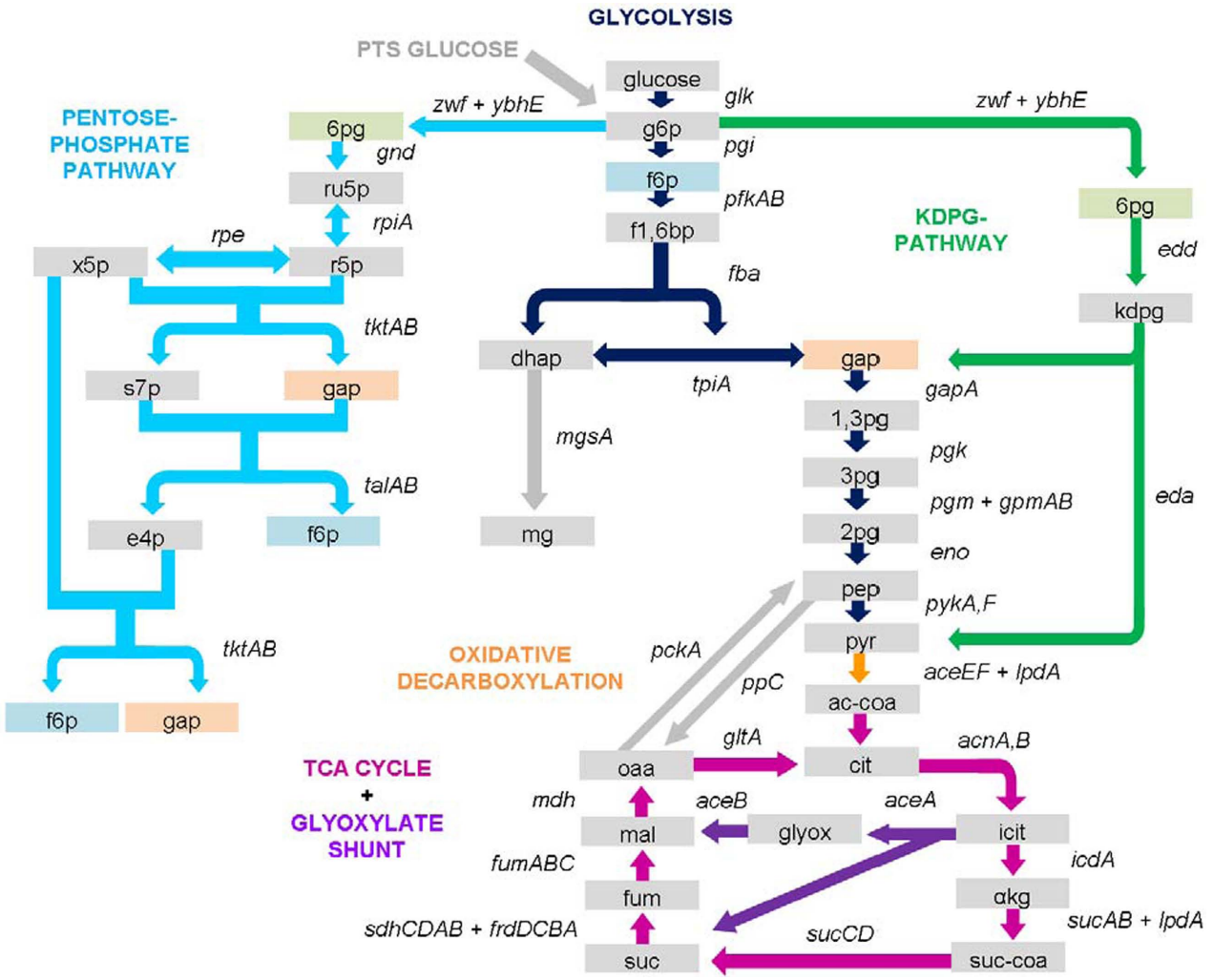


FIGURE 4 | *Salmonella* central carbon metabolism. Shown are genes and enzymes of the central carbon metabolism covering glycolysis, PPP, KDPG/Entner–Doudoroff pathway, oxidative carboxylation, TCA cycle, and anaplerotic reactions. Boxes represent metabolites, arrows indicate enzyme reactions. Metabolites and enzymes are colored according to the color of their pathway name. Arrows are directed according flux direction under glucose metabolism but may vary under different conditions. Detailed information on importance of the illustrated pathways in intracellular infection is outlined in the text including behavior of an array of different enzyme mutations. 1,3pg, 1,3-bisphospho-D-glycerate; 2pg, 2-phospho-D-glycerate; 3pg, 3-phospho-D-glycerate; 6pg, 6-phospho-D-gluconate; αkg, α-ketoglutarate; ac-coa, acetyl-CoA; cit, citrate; dhap, dihydroxyacetone phosphate; e4p, D-erythrose-4-phosphate; f1,6bp, fructose-1,6-bisphosphate; f6p, D-fructose-6-phosphate; fum, fumarate; gap, D-glyceraldehyde-3-phosphate; glyox, glyoxylate; icit, isocitrate; kdpg, 2-dehydro-3-deoxy-D-gluconate-6-phosphate; mal, (S)-malate; mg, methylglyoxal; oaa, oxaloacetate; pep, phosphoenolpyruvate; pyr, pyruvate; r5p, D-ribose-5-phosphate; ru5p, D-ribulose-5-phosphate; s7p, D-sedoheptulose-7-phosphate; suc, succinate; suc-coa, succinyl-CoA; x5p, D-xylulose-5-phosphate.

important if *Salmonella* is challenged with higher degrees of oxidative stress.

The significance of glucose as one of the major C-sources, and glycolysis as the main route for utilization has recently been shown (Bowden et al., 2009). Extending the set of glycolysis mutant strains analyzed by Bowden et al. (2009), work in our groups showed that *eno*, *fba*, *pgk*, *gapA*, or *tpiA* deficient strains are strongly attenuated in intracellular replication and survival in RAW 264.7 cells (unpublished results). In case of a *tpiA* mutant it was demonstrated that reduced growth in rich medium (lysogeny broth, LB) and decreased fitness in mice seems to depend on accumulation of the toxic electrophile methylglyoxal from accumulated dihydroxyacetone phosphate by reaction of the methylglyoxal synthetase (Paterson et al., 2009). Overproduction of methylglyoxal as a result of *tpiA* mutation for *E. coli* has been shown before (Cooper and Anderson, 1970; Cooper, 1984) and although the concentration is lower in the *tpiA* mutant of *Salmonella*, it still leads to reduced growth in medium and lower fitness in mice (Paterson et al., 2009). Mutant strains deficient in *eno*, *fba*, *gapA*, or *pgk* show similar growth characteristics in medium containing just traces of glucose or other C-sources ending up in glycolytic intermediates, e.g., in LB or minimal medium containing ribose. None of these mutant strains was able to grow in these media, but growth was observed on minimal medium with glycerol or PEP

as C-source. The growth defect in the presence of sugars is due to accumulation of phosphorylated sugars, i.e., fructose-6-phosphate (f6p), glucose-6-phosphate (g6p), and fructose-1,6-bisphosphate (f1,6bp), that appear to influence RNA stability as observed in *E. coli* mutant strains (Bock and Neidhardt, 1966; Singer et al., 1991; Morita et al., 2003). Glucose is the main C-source in common cell culture media for eukaryotic host cells along with amino acids. The inability of these mutants to survive intracellularly suggests that the metabolism of glucose through glycolysis and the accumulation of phosphorylated sugars is harmful, caused by the knockout of the aforementioned glycolytic genes. Nevertheless, further analytical approaches such as quantification of metabolites by GC–MS are required to determine the presence and accumulation of these glycolytic intermediates.

OXIDATIVE DECARBOXYLATION OF PYRUVATE – THE CONNECTION TO THE TCA CYCLE

Pyruvate as final product of glycolysis or KDPGP can be further metabolized to acetyl-CoA by reaction of pyruvate dehydrogenase. This production of acetyl-CoA and the connection to the TCA cycle is of great importance as indicated by the reduced growth of a pyruvate dehydrogenase subunit I deficient strain (Δ*aceE*) in rich media (LB) and strong attenuation of intracellular replication. An *aceE* mutant of *S. enterica* serovar Enteritidis tested in a chicken infection model also showed reduced growth compared to a WT strain, a lower invasion rate in HeLa cells and was less resistant to ROI. Chicken macrophages (HD-11) were able to eliminate this strain completely within 24 h after infection (Pang et al., 2011). In non-activated RAW 264.7 cells that produce low amounts of ROI, reduced replication was observed (unpublished observations). The loss of *aceE* may also have influence on the expression of virulence genes important for the defense against ROI. The attenuation of the Δ*aceE* strain may be due to the metabolic burden of virulence gene expression and the lower energy production due to this knockout. Energy is probably mainly produced by glycolysis resulting in less ATP than via respiratory chain of the reduction agents produced in the TCA cycle.

THE ROLE OF THE TCA CYCLE AND ANAPLEROTIC REACTIONS

The TCA cycle plays an important role as a source for precursors for anabolic pathways, e.g., amino acids, and reducing agents used as electron donors in the respiratory chain or for biosynthesis. Due to this central role, it is obvious that in the last years research is focused on this part of the central carbon metabolism. Some groups already showed loss of virulence in a murine model for mutants with defects in the TCA cycle (Tchawa Yimga et al., 2006; Bowden et al., 2010), concluding that the ability to run the full TCA cycle is critical for virulence. The latter group could show that for some of these TCA cycle mutant strains showing reduced virulence in the murine model (*mdh*, *sdhCDAB*, and *sucCD*) an increased replication in resting and activated macrophages (RAW 264.7) could be observed that is probably due to differences in the *in vitro* and *in vivo* environment (available nutrients, etc.). An *sdhA* mutant strain of *E. coli* had an expanded life span due to a lower production of superoxide in contrast to the WT (Gonidakis et al., 2011). The phenomenon may explain the observed increased intracellular replication for the *mdh*, *sdhCDAB*, and *sucCD* strains of *Salmonella in vitro*. A TCA cycle mutant strain with an *icdA* deletion was not tested in this experimental setup but we could observe a strong attenuation in inactivated murine RAW macrophages. The strong attenuation of the *icdA* mutant could be due to the accumulation of an inhibitory product like citrate or isocitrate as was shown for an *E. coli icd* mutant (Lakshmi and Helling, 1976).

In view of the significance of the anaplerotic reaction for survival in macrophages (**Figure 4**), it seems quite convincing that the glyoxylate shunt, which is used during growth on acetate and fatty acids is important for chronic infections but not for acute infections (Fang et al., 2005; Tchawa Yimga et al., 2006). This supports the suggestion that acetate and fatty acids are of no importance as C-sources in acute infection either due to their absence or the presence of more favorable C-sources like glucose. This is confirmed by transcriptome data observing no up-regulation in expression of the isocitrate lyase gene *aceA* (Eriksson et al., 2003) and by observations that *Salmonella* strains with a defect in β-oxidation of fatty acids (*fadD*) also showed the same virulence as the WT strain in a murine model (Tchawa Yimga et al., 2006).

The direct conversion of PEP to oxaloacetate, an important precursor for the amino acids aspartate and asparagine, is another anaplerotic reaction preventing the TCA cycle from idling and is catalyzed by the enzyme PEP carboxylase (Sabe et al., 1984). Loss of this enzyme does not lead to reduced virulence in a murine model (Tchawa Yimga et al., 2006) indicating the presence of C-sources other than glucose, e.g., amino acids, in amounts required for intracellular replication of *Salmonella*. This would make redundant the replenishment of the TCA cycle by this route.

TOWARDS A COMPREHENSIVE UNDERSTANDING OF THE METABOLIC REQUIREMENTS OF THE INTRACELLULAR LIFESTYLE OF *SALMONELLA*

The environmental and nutritional conditions encountered by *Salmonella* within the SCV are still a matter of debate. In the process of SCV formation an acidification of the intravacuolar environment takes place (Rathman et al., 1996). *Salmonella* metabolism during infection adapts to this intracellular life style in the host cell. It is regulated by different transcription factors (Eisenreich et al., 2010; Götz and Goebel, 2010). However, to what extent can the SCV be reached from outside? Are there macrophage-specific strategies of adaptation? A number of recent efforts were stimulated by these questions (Eriksson et al., 2003; Becker et al., 2006; Bowden et al., 2009; Götz and Goebel, 2010) and gene knockout strategies as well as metabolite measurements and modeling were applied to better understand *Salmonella* life inside the cell.

Limitation of magnesium, phosphate, and iron occur as genes for their uptake are up-regulated in *Salmonella* during infection. By gene expression analysis it could also be shown that glycolysis, KDPGP, and TCA cycle are highly expressed in SCV of macrophages and epithelial cells (Hautefort et al., 2008). In contrast, gluconeogenesis is not required for a full virulent phenotype in mice (Tchawa Yimga et al., 2006).

In a number of reports, the nutrient-poor, underfed conditions of the SCV were stressed (Eriksson et al., 2003; Ibarra et al., 2009). This applies even more for the early phases of the infection. Furthermore, detailed studies suggest that intracellular *Salmonella* deploys specific mechanisms to redirect vacuolar transport to make use of host-derived nutrients (Kuhle et al., 2006; Drecktrah et al., 2007; Rajashekar et al., 2008).

Direct comparison of intracellular lifestyles of bacterial pathogens such as *Salmonella* and *Legionella* that reside within a membrane-bound compartment, to those with survival and replication within the cytoplasm of host cells is of interest. For the latter, survival depends on the availability of cytoplasmic host cell metabolites and tight regulation is necessary such as by the pathogenicity factor PrfA in *Listeria* (Eylert et al., 2008) to survive surprisingly poor nutritional conditions (Eriksson et al., 2003; Ibarra et al., 2009). For the *Salmonella* lifestyle in the SCV, also strong metabolic adaptations are necessary, in particular regarding low abundant nutrients and ions, however, carbohydrates are not that limiting. The survival under these different conditions of intracellular life in the host cell indicates successful adaptation to different metabolic limitations.

The expression of virulence factors by *Salmonella* does not come without a price. For example, in a recent study retarded growth of *S.* Typhimurium cells expressing the type III secretion system 1 (T3SS-1) compared to the minus phenotype was observed and the effects on growth kinetics were modeled (Sturm et al., 2011). Growth retardation was at least partially attributable to the expression of the T3SS-1 effector and/or translocon proteins. In spite of this growth penalty, the T3SS-1(+) subpopulation increased from <10% to approximately 60% during the late logarithmic growth phase of an LB batch culture. As shown by experimental data and mathematical modeling, this was attributable to an increasing initiation rate of expression of T3SS-1 genes, in response to environmental cues accumulating during this growth phase. The key is here to mathematically describe the whole system, correctly quantify responses and obtain results from the model which agree with observation. Such models and methods can, furthermore, also be transferred to better understand other genetic variations pertaining, for instance, to effector proteins and estimate the cost of virulence regarding growth and metabolism. Such systems biology approaches to *Salmonella* lifestyle and pathogenicity are expected to grow (Helaine et al., 2010).

Furthermore, a number of technical developments contributed to the understanding of bacterial metabolism during infection. The modeling of genome-scale metabolic networks is now feasible, originally pioneered by the Palsson group (Price et al., 2004a). However, even with strong computational power at hand, detailed modeling requires additional methods to handle the combinatorial explosion of different pathways concerned, for instance, by dividing the metabolic network (Schuster et al., 2002) or sampling averages over high numbers of different modes (Price et al., 2004b). Furthermore, a number of recent developments tackle new software solutions for metabolic modeling including modifications of extreme pathway analysis (Kaleta et al., 2009; Schuster et al., 2010a). Moreover, the analysis of isotopolog data profits from advances in software development (Nanchen et al., 2007). Achieving gene knockout combinations experimentally has been advanced technically (Datsenko and Wanner, 2000; Gerlach et al., 2009) and this has been complemented by *in silico* prediction studies on gene knockouts agreeing well with these experiments. This can also be applied to study survival during infection (Raghunathan et al., 2009).

IMPLICATIONS AND CONCLUSIONS

Pathogenicity can arise from a bacterial specialization where intrinsic cellular functions are especially adapted to the host's environment such as in exclusive human pathogens like *Mycobacterium tuberculosis* (Moller and Hoal, 2010), *Neisseria gonorrhoeae* (Criss and Seifert, 2012), *N. meningitides* (Criss and Seifert, 2012), and *Mycoplasma pneumoniae* (Dumke et al., 2011). In contrast, *Salmonella* is a surprisingly successful intracellular pathogen that is less specialized but rather a generalist with an extraordinary metabolic versatility. Furthermore, *E. coli*, *Pseudomonas* spp., *Klebsiella* spp., and further bacteria are common pathogens with a similar metabolic background which are not highly specialized for a specific host environment.

To underline this, a broad spectrum of central carbon metabolism routes constitutes the repertoire of the *Salmonella* lifestyle (**Figure 4**). Glucose, the predominant C-source during SCV colonization can either be degraded in the glycolytic pathway or the KDPG route, the former one preferentially used. The TCA cycle as the major biosynthetic origin of precursor and provider of reductive agents is complete and is supported by important anaplerotic reactions that lead the metabolic flux to the TCA cycle, mainly by the PEP carboxylase. However, this enzyme can be compensated by other flux modes as shown by EMA. In systemic infections, the PPP generates NADPH required for reductases in oxidative stress response. *Salmonella* is well-equipped to rapidly adapt to various environments during the passage through the host's body.

The flexible metabolic abilities of *Salmonella* make it challenging to elucidate targets to inhibit metabolic functions during infection. The pathogen efficiently modifies its unique vacuolar compartment for its benefits. Additionally, the transport of nutrients to the SCV in response of *Salmonella* effector proteins is currently under investigation. Host–pathogen metabolism is intertwined.

Is there an Achilles' heel of *Salmonella* metabolism in infection? To answer this we suggest a combination of a thorough analysis of central metabolic enzymes such as mutant phenotypes of glycolytic enzymes with systemic analysis such of transporter expression profiles, strain-specific analysis, and projecting those on the very specific environmental conditions under study. Thus, the loss of pyruvate dehydrogenase may have an influence on the expression of virulence genes important for ROI defense. Further, mutants with a defect in the TCA cycle showed reduced virulence in a murine infection model. In *Salmonella*, the T3SS-2 is responsible for translocation of over 20 virulence proteins into the host cell cytoplasm (Haraga et al., 2008). SifA is probably the most prominent among these effectors in maintaining the integrity of the SCV and induction of extensive networks of tubular membrane compartments including *Salmonella*-induced

filaments and recently identified further tubular compartment (Schroeder et al., 2011).

Future work will take up the challenge of this surprisingly well-adapted pathogen by looking more closely at the host–pathogen interaction, not only regarding metabolic interactions for further pathways including secondary metabolism, iso-enzymes, and intertwined links between pathways (including supplementation experiments), but also regarding the regulatory and immune response from the host, for instance, potential immune modifiers as novel approaches to boost host response in severe infections by *Salmonella*. Infectious disease burden in developing countries will doubtless profit from improved hygiene and clean water supply but also protective nutrient additives (e.g., vitamins; golden rice, and similar recent developments) and regarding trace elements such as selenocysteine. Challenges in industrialized countries include persistent infection which can again be coped with metabolic approaches relying, for instance, on elicitors of growth which in turn make the previous persister *Salmonella* vulnerable to standard antibiotics. In general, the exploration of the link between metabolism and infection has to be explored further to improve medical options against *Salmonella*.

ACKNOWLEDGMENT

This work was supported by the German Research Foundation (DFG), grants Da 208/13-2 and He 1964/14-1.

REFERENCES

AbuOun, M., Suthers, P. F., Jones, G. I., Carter, B. R., Saunders, M. P., Maranas, C. D., Woodward, M. J., and Anjum, M. F. (2009). Genome scale reconstruction of a *Salmonella* metabolic model: comparison of similarity and differences with a commensal *Escherichia coli* strain. *J. Biol. Chem.* 284, 29480–29488.

Allison, K. R., Brynildsen, M. P., and Collins, J. J. (2011). Metabolite-enabled eradication of bacterial persisters by aminoglycosides. *Nature* 473, 216–220.

Becker, D., Selbach, M., Rollenhagen, C., Ballmaier, M., Meyer, T. F., Mann, M., and Bumann, D. (2006). Robust *Salmonella* metabolism limits possibilities for new antimicrobials. *Nature* 440, 303–307.

Bock, A., and Neidhardt, F. C. (1966). Properties of a mutant of *Escherichia coli* with a temperature-sensitive fructose-1,6-diphosphate aldolase. *J. Bacteriol.* 92, 470–476.

Bowden, S. D., Ramachandran, V. K., Knudsen, G. M., Hinton, J. C., and Thompson, A. (2010). An incomplete TCA cycle increases survival of *Salmonella* Typhimurium during infection of resting and activated murine macrophages. *PLoS ONE* 5, e13871. doi: 10.1371/journal.pone.0013871

Bowden, S. D., Rowley, G., Hinton, J. C., and Thompson, A. (2009). Glucose and glycolysis are required for the successful infection of macrophages and mice by *Salmonella enterica* serovar Typhimurium. *Infect. Immun.* 77, 3117–3126.

Bumann, D. (2009a). Pathogen proteomes during infection: a basis for infection research and novel control strategies. *J. Proteomics* 73, 2267–2276.

Bumann, D. (2009b). System-level analysis of *Salmonella* metabolism during infection. *Curr. Opin. Microbiol.* 12, 559–567.

Cazalet, C., Rusniok, C., Bruggemann, H., Zidane, N., Magnier, A., Ma, L., Tichit, M., Jarraud, S., Bouchier, C., Vandenesch, F., Kunst, F., Etienne, J., Glaser, P., and Buchrieser, C. (2004). Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat. Genet.* 36, 1165–1173.

Cecil, A., Rikanovic, C., Ohlsen, K., Liang, C., Bernhardt, J., Oelschlaeger, T. A., Gulder, T., Bringmann, G., Holzgrabe, U., Unger, M., and Dandekar, T. (2011). Modeling antibiotic and cytotoxic effects of the dimeric isoquinoline IQ-143 on metabolism and its regulation in *Staphylococcus aureus*, *Staphylococcus epidermidis* and human cells. *Genome Biol.* 12, R24.

Cooper, R. A. (1984). Metabolism of methylglyoxal in microorganisms. *Annu. Rev. Microbiol.* 38, 49–68.

Cooper, R. A., and Anderson, A. (1970). The formation and catabolism of methylglyoxal during glycolysis in *Escherichia coli*. *FEBS Lett.* 11, 273–276.

Covert, M. W., Knight, E. M., Reed, J. L., Herrgard, M. J., and Palsson, B. O. (2004). Integrating high-throughput and computational data elucidates bacterial networks. *Nature* 429, 92–96.

Criss, A. K., and Seifert, H. S. (2012). A bacterial siren song: intimate interactions between *Neisseria* and neutrophils. *Nat. Rev. Microbiol.* 10, 178–190.

Cummings, P. L., Sorvillo, F., and Kuo, T. (2010). Salmonellosis-related mortality in the United States, 1990–2006. *Foodborne Pathog. Dis.* 7, 1393–1399.

Datsenko, K. A., and Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6640–6645.

Drecktrah, D., Knodler, L. A., Howe, D., and Steele-Mortimer, O. (2007). *Salmonella* trafficking is defined by continuous dynamic interactions with the endolysosomal system. *Traffic* 8, 212–225.

Dumke, R., Hausner, M., and Jacobs, E. (2011). Role of *Mycoplasma pneumoniae* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in mediating interactions with the human extracellular matrix. *Microbiology* 157, 2328–2338.

Eisenreich, W., Dandekar, T., Heesemann, J., and Goebel, W. (2010). Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nat. Rev. Microbiol.* 8, 401–412.

Eisenreich, W., Slaghuis, J., Laupitz, R., Bussemer, J., Stritzker, J., Schwarz, C., Schwarz, R., Dandekar, T., Goebel, W., and Bacher, A. (2006). ^{13}C isotopologue perturbation studies of *Listeria monocytogenes* carbon metabolism and its modulation by the virulence regulator PrfA. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2040–2045.

Eriksson, S., Lucchini, S., Thompson, A., Rhen, M., and Hinton, J. C. (2003). Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. *Mol. Microbiol.* 47, 103–118.

Eylert, E., Herrmann, V., Jules, M., Gillmaier, N., Lautner, M., Buchrieser, C., Eisenreich, W., and Heuner, K. (2010). Isotopologue profiling of *Legionella pneumophila*: role of serine and glucose as carbon substrates. *J. Biol. Chem.* 285, 22232–22243.

Eylert, E., Schar, J., Mertins, S., Stoll, R., Bacher, A., Goebel, W., and Eisenreich, W. (2008). Carbon metabolism of *Listeria monocytogenes* growing inside macrophages. *Mol. Microbiol.* 69, 1008–1017.

Fang, F. C., Libby, S. J., Castor, M. E., and Fung, A. M. (2005). Isocitrate lyase (AceA) is required for *Salmonella* persistence but not for acute lethal infection in mice. *Infect. Immun.* 73, 2547–2549.

Fields, P. I., Swanson, R. V., Haidaris, C. G., and Heffron, F. (1986). Mutants of *Salmonella* Typhimurium that cannot survive within the macrophage are avirulent. *Proc. Natl. Acad. Sci. U.S.A.* 83, 5189–5193.

Fischer, E., and Sauer, U. (2003). Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC–MS. *Eur. J. Biochem.* 270, 880–891.

Gaudermann, P., Vogl, I., Zientz, E., Silva, F. J., Moya, A., Gross, R., and Dandekar, T. (2006). Analysis of and function predictions for previously conserved hypothetical or putative proteins in *Blochmannia floridanus*. *BMC Microbiol.* 6, 1. doi: 10.1186/1471-2180-6-1

Gerlach, R. G., Jäckel, D., Holzer, S. U., and Hensel, M. (2009). Rapid oligonucleotide-based recombineering of the chromosome of *Salmonella enterica*. *Appl. Environ. Microbiol.* 75, 1575–1580.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., De Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fsihi, H., Garcia-Del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., De Pablos, B., Perez-Diaz, J. C., Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J., and Cossart, P. (2001). Comparative genomics of *Listeria* species. *Science* 294, 849–852.

Gonidakis, S., Finkel, S. E., and Longo, V. D. (2011). Lifespan extension and paraquat resistance in a ubiqui-deficient *Escherichia coli* mutant depend on transcription factors ArcA and TdcA. *Aging (Albany NY)* 3, 291–303.

Götz, A., and Goebel, W. (2010). Glucose and glucose 6-phosphate as carbon sources in extra- and intracellular growth of enteroinvasive *Escherichia coli* and *Salmonella enterica*. *Microbiology* 156, 1176–1187.

Harada, E., Iida, K., Shiota, S., Nakayama, H., and Yoshida, S. (2010). Glucose metabolism in *Legionella pneumophila*: dependence on the Entner–Doudoroff pathway and connection with intracellular bacterial growth. *J. Bacteriol.* 192, 2892–2899.

Haraga, A., Ohlson, M. B., and Miller, S. I. (2008). *Salmonellae* interplay with host cells. *Nat. Rev. Microbiol.* 6, 53–66.

Harvey, P. C., Watson, M., Hulme, S., Jones, M. A., Lovell, M., Berchieri, A. Jr., Young, J., Bumstead, N., and Barrow, P. (2011). *Salmonella enterica* serovar Typhimurium colonizing the lumen of the chicken intestine grows slowly and upregulates a unique set of virulence and metabolism genes. *Infect. Immun.* 79, 4105–4121.

Hautefort, I., Thompson, A., Eriksson-Ygberg, S., Parker, M. L., Lucchini, S., Danino, V., Bongaerts, R. J., Ahmad, N., Rhen, M., and Hinton, J. C. (2008). During infection of epithelial cells *Salmonella enterica* serovar Typhimurium undergoes a time-dependent transcriptional adaptation that results in simultaneous expression of three type 3 secretion systems. *Cell. Microbiol.* 10, 958–984.

Helaine, S., Thompson, J. A., Watson, K. G., Liu, M., Boyle, C., and Holden, D. W. (2010). Dynamics of intracellular bacterial replication at the single cell level. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3746–3751.

Ibarra, J. A., Knodler, L. A., Sturdevant, D. E., Virtaneva, K., Carmody, A. B., Fischer, E. R., Porcella, S. F., and Steele-Mortimer, O. (2009). Induction of *Salmonella* pathogenicity island 1 under different growth conditions can affect *Salmonella*-host cell interactions *in vitro*. *Microbiology* 156, 1120–1133.

Kaleta, C., De Figueiredo, L. F., and Schuster, S. (2009). Can the whole be less than the sum of its parts? Pathway analysis in genome-scale metabolic networks using elementary flux patterns. *Genome Res.* 19, 1872–1883.

Kuhle, V., Abrahams, G. L., and Hensel, M. (2006). Intracellular *Salmonella enterica* redirect exocytic transport processes in a *Salmonella* pathogenicity island 2-dependent manner. *Traffic* 7, 716–730.

Lakshmi, T. M., and Helling, R. B. (1976). Selection for citrate synthase deficiency in *icd* mutants of *Escherichia coli*. *J. Bacteriol.* 127, 76–83.

Liang, C., Liebeke, M., Schwarz, R., Zuhlke, D., Fuchs, S., Menschner, L., Engelmann, S., Wolz, C., Jaglitz, S., Bernhardt, J., Hecker, M., Lalk, M., and Dandekar, T. (2011). *Staphylococcus aureus* physiological growth limitations: insights from flux calculations built on proteomics and external metabolite data. *Proteomics* 11, 1915–1935.

Lim, S., Kim, M., Choi, J., and Ryu, S. (2010). A mutation in *tdcA* attenuates the virulence of *Salmonella enterica* serovar Typhimurium. *Mol. Cells* 29, 509–517.

Lundberg, B. E., Wolf, R. E. Jr., Dinauer, M. C., Xu, Y., and Fang, F. C. (1999). Glucose 6-phosphate dehydrogenase is required for *Salmonella* Typhimurium virulence and resistance to reactive oxygen and nitrogen intermediates. *Infect. Immun.* 67, 436–438.

Mastroeni, P., Grant, A., Restif, O., and Maskell, D. (2009). A dynamic view of the spread and intracellular distribution of *Salmonella enterica*. *Nat. Rev. Microbiol.* 7, 73–80.

McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., and Wilson, R. K. (2001). Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413, 852–856.

McDermott, J. E., Yoon, H., Nakayasu, E. S., Metz, T. O., Hyduke, D. R., Kidwai, A. S., Palsson, B. O., Adkins, J. N., and Heffron, F. (2011). Technologies and approaches to elucidate and model the virulence program of *Salmonella*. *Front. Microbiol.* 2, 121. doi: 10.3389/fmicb.2011.00121

Moller, M., and Hoal, E. G. (2010). Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. *Tuberculosis (Edinb)* 90, 71–83.

Morita, T., El-Kazzaz, W., Tanaka, Y., Inada, T., and Aiba, H. (2003). Accumulation of glucose 6-phosphate or fructose 6-phosphate is responsible for destabilization of glucose transporter mRNA in *Escherichia coli*. *J. Biol. Chem.* 278, 15608–15614.

Nanchen, A., Fuhrer, T., and Sauer, U. (2007). Determination of metabolic flux ratios from ^{13}C-experiments and gas chromatography-mass spectrometry data: protocol and principles. *Methods Mol. Biol.* 358, 177–197.

Neidhardt, F. C. (1996). *Escherichia coli and Salmonella: Cellular and Molecular Biology*. Washington, DC: ASM Press.

Pang, E., Tien-Lin, C., Selvaraj, M., Chang, J., and Kwang, J. (2011). Deletion of the *aceE* gene (encoding a component of pyruvate dehydrogenase) attenuates *Salmonella enterica* serovar Enteritidis. *FEMS Immunol. Med. Microbiol.* 63, 108–118.

Paterson, G. K., Cone, D. B., Northen, H., Peters, S. E., and Maskell, D. J. (2009). Deletion of the gene encoding the glycolytic enzyme triosephosphate isomerase (*tpi*) alters morphology of *Salmonella enterica* serovar Typhimurium and decreases fitness in mice. *FEMS Microbiol. Lett.* 294, 45–51.

Price, N. D., Reed, J. L., and Palsson, B. O. (2004a). Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* 2, 886–897.

Price, N. D., Schellenberger, J., and Palsson, B. O. (2004b). Uniform sampling of steady-state flux spaces: means to design experiments and to interpret enzymopathies. *Biophys. J.* 87, 2172–2186.

Raghunathan, A., Reed, J., Shin, S., Palsson, B., and Daefler, S. (2009). Constraint-based analysis of metabolic capacity of *Salmonella* typhimurium during host-pathogen interaction. *BMC Syst. Biol.* 3, 38. doi: 10.1186/1752-0509-3-38

Rajashekar, R., Liebl, D., Seitz, A., and Hensel, M. (2008). Dynamic remodeling of the endosomal system during formation of *Salmonella*-induced filaments by intracellular *Salmonella enterica*. *Traffic* 9, 2100–2116.

Rathman, M., Sjaastad, M. D., and Falkow, S. (1996). Acidification of phagosomes containing *Salmonella* typhimurium in murine macrophages. *Infect. Immun.* 64, 2765–2773.

Ruppin, E., Papin, J. A., De Figueiredo, L. F., and Schuster, S. (2010). Metabolic reconstruction, constraint-based analysis and game theory to probe genome-scale metabolic networks. *Curr. Opin. Biotechnol.* 21, 502–510.

Sabe, H., Miwa, T., Kodaki, T., Izui, K., Hiraga, S., and Katsuki, H. (1984). Molecular cloning of the phosphoenolpyruvate carboxylase gene, *ppc*, of *Escherichia coli*. *Gene* 31, 279–283.

Schar, J., Stoll, R., Schauer, K., Loeffler, D. I., Eylert, E., Joseph, B., Eisenreich, W., Fuchs, T. M., and Goebel, W. (2010). Pyruvate carboxylase plays a crucial role in carbon metabolism of extra- and intracellularly replicating *Listeria monocytogenes*. *J. Bacteriol.* 192, 1774–1784.

Schauer, K., Geginat, G., Liang, C., Goebel, W., Dandekar, T., and Fuchs, T. M. (2010). Deciphering the intracellular metabolism of *Listeria monocytogenes* by mutant screening and modelling. *BMC Genomics* 11, 573. doi: 10.1186/1471-2164-11-573

Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., Zielinski, D. C., Bordbar, A., Lewis, N. E., Rahmanian, S., Kang, J., Hyduke, D. R., and Palsson, B. O. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* 6, 1290–1307.

Schroeder, N., Mota, L. J., and Meresse, S. (2011). *Salmonella*-induced tubular networks. *Trends Microbiol.* 19, 268–277.

Schuster, S., De Figueiredo, L. F., and Kaleta, C. (2010a). Predicting novel pathways in genome-scale metabolic networks. *Biochem. Soc. Trans.* 38, 1202–1205.

Schuster, S., Kreft, J. U., Brenner, N., Wessely, F., Theissen, G., Ruppin, E., and Schroeter, A. (2010b). Cooperation and cheating in microbial exoenzyme production – theoretical analysis for biotechnological applications. *Biotechnol. J.* 5, 751–758.

Schuster, S., De Figueiredo, L. F., Schroeter, A., and Kaleta, C. (2011). Combining metabolic pathway analysis with Evolutionary Game Theory: explaining the occurrence of low-yield pathways by an analytic optimization approach. *Biosystems* 105, 147–153.

Schuster, S., Pfeiffer, T., Moldenhauer, F., Koch, I., and Dandekar, T. (2002). Exploring the pathway structure of metabolism: decomposition into subnetworks and application to *Mycoplasma pneumoniae*. *Bioinformatics* 18, 351–361.

Schwarz, R., Liang, C., Kaleta, C., Kuhnel, M., Hoffmann, E., Kuznetsov, S., Hecker, M., Griffiths, G., Schuster, S., and Dandekar, T. (2007). Integrated network reconstruction, visualization and analysis using YANAsquare. *BMC Bioinformatics* 8, 313. doi: 10.1186/1471-2105-8-313

Shi, L., Adkins, J. N., Coleman, J. R., Schepmoes, A. A., Dohnkova, A., Mottaz, H. M., Norbeck, A. D.,

Purvine, S. O., Manes, N. P., Smallwood, H. S., Wang, H., Forbes, J., Gros, P., Uzzau, S., Rodland, K. D., Heffron, F., Smith, R. D., and Squier, T. C. (2006). Proteomic analysis of *Salmonella enterica* serovar Typhimurium isolated from RAW 264.7 macrophages: identification of a novel protein that contributes to the replication of serovar Typhimurium inside macrophages. *J. Biol. Chem.* 281, 29131–29140.

Singer, M., Walter, W. A., Cali, B. M., Rouviere, P., Liebke, H. H., Gourse, R. L., and Gross, C. A. (1991). Physiological effects of the fructose-1,6-diphosphate aldolase ts8 mutation on stable RNA synthesis in *Escherichia coli*. *J. Bacteriol.* 173, 6249–6257.

Sturm, A., Heinemann, M., Arnoldini, M., Benecke, A., Ackermann, M., Benz, M., Dormann, J., and Hardt, W. D. (2011). The cost of virulence: retarded growth of *Salmonella* Typhimurium cells expressing type III secretion system 1. *PLoS Pathog.* 7, e1002143. doi: 10.1371/journal.ppat.1002143

Tchawa Yimga, M., Leatham, M. P., Allen, J. H., Laux, D. C., Conway, T., and Cohen, P. S. (2006). Role of gluconeogenesis and the tricarboxylic acid cycle in the virulence of *Salmonella enterica* serovar Typhimurium in BALB/c mice. *Infect. Immun.* 74, 1130–1140.

Tesh, M. J., and Miller, R. D. (1981). Amino acid requirements for *Legionella pneumophila* growth. *J. Clin. Microbiol.* 13, 865–869.

Thiele, I., Hyduke, D. R., Steeb, B., Fankam, G., Allen, D. K., Bazzani, S., Charusanti, P., Chen, F. C., Fleming, R. M., Hsiung, C. A., De Keersmaecker, S. C., Liao, Y. C., Marchal, K., Mo, M. L., Ozdemir, E., Raghunathan, A., Reed, J. L., Shin, S. I., Sigurbjornsdottir, S., Steinmann, J., Sudarsan, S., Swainston, N., Thijs, I. M., Zengler, K., Palsson, B. O., Adkins, J. N., and Bumann, D. (2011). A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella* Typhimurium LT2. *BMC Syst. Biol.* 5, 8. doi: 10.1186/1752-0509-5-8

Wieland, H., Ullrich, S., Lang, F., and Neumeister, B. (2005). Intracellular multiplication of *Legionella pneumophila* depends on host cell amino acid transporter SLC1A5. *Mol. Microbiol.* 55, 1528–1537.

World Health Organization. (2005). *Drug-resistant Salmonella*. WHO Fact Sheet 139. Geneva: WHO.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 December 2011; accepted: 12 April 2012; published online: 03 May 2012.

Citation: Dandekar T, Fieselmann A, Popp J and Hensel M (2012) Salmonella enterica: a surprisingly well-adapted intracellular lifestyle. Front. Microbio. ***3****:164. doi: 10.3389/fmicb.2012.00164*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

REVIEW ARTICLE
published: 03 February 2012
doi: 10.3389/fmicb.2012.00023

Toward a systemic understanding of *Listeria monocytogenes* metabolism during infection

Thilo M. Fuchs[1,2]*, Wolfgang Eisenreich[3], Tanja Kern[1] and Thomas Dandekar[4]

[1] *Abteilung Mikrobiologie, Zentralinstitut für Ernährungs- und Lebensmittelforschung, Technische Universität München, Freising, Germany*
[2] *Lehrstuhl für Mikrobielle Ökologie, Department Biowissenschaften, Wissenschaftszentrum Weihenstephan, Technische Universität München, Freising, Germany*
[3] *Lehrstuhl für Biochemie, Technische Universität München, Garching, Germany*
[4] *Abteilung Bioinformatik, Theodor-Boveri-Institut (Biozentrum), Universität Würzburg, Würzburg, Germany*

Edited by:
Reinhard Guthke, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany

Reviewed by:
Guoku Hu, Creighton University, USA
Conor P. O'Byrne, NUI Galway, Ireland

****Correspondence:***
Thilo M. Fuchs, Abteilung Mikrobiologie, Zentralinstitut für Ernährungs- und Lebensmittelforschung, Technische Universität München, Weihenstephaner Berg 3, 85350 Freising, Germany.
e-mail: thilo.fuchs@wzw.tum.de

Listeria monocytogenes is a foodborne human pathogen that can cause invasive infection in susceptible animals and humans. For proliferation within hosts, this facultative intracellular pathogen uses a reservoir of specific metabolic pathways, transporter, and enzymatic functions whose expression requires the coordinated activity of a complex regulatory network. The highly adapted metabolism of *L. monocytogenes* strongly depends on the nutrient composition of various milieus encountered during infection. Transcriptomic and proteomic studies revealed the spatial–temporal dynamic of gene expression of this pathogen during replication within cultured cells or *in vivo*. Metabolic clues are the utilization of unusual C_2- and C_3-bodies, the metabolism of pyruvate, thiamine availability, the uptake of peptides, the acquisition or biosynthesis of certain amino acids, and the degradation of glucose-phosphate via the pentose phosphate pathway. These examples illustrate the interference of *in vivo* conditions with energy, carbon, and nitrogen metabolism, thus affecting listerial growth. The exploitation, analysis, and modeling of the available data sets served as a first attempt to a systemic understanding of listerial metabolism during infection. *L. monocytogenes* might serve as a model organism for systems biology of a Gram-positive, facultative intracellular bacterium.

Keywords: *Listeria monocytogenes*, infection, metabolism, systems biology, modeling, intracellular

INTRODUCTION

A successful infection by bacterial pathogens requires multiple adaptation processes including adhesion to host tissues, modulation of the immune response, or toxic activity toward the host defense system. Some pathogens enter epithelial cells or are internalized by professional phagocytes, and these steps are often followed by bacterial manipulation of the host cell actin skeleton and the manipulation of the endocytic route. Most of these processes, which often require specific virulence factors that enable the microbes to overcome the various physical and biochemical barriers of the infected host, have been characterized in detail.

In contrast, little attention has been given to the metabolic requirements and the metabolic flexibility of bacteria during infection, in parts due to limitation of analytical tools, and because the bacterial metabolism *in vivo* and *in vitro* has erroneously been assumed to be similar (Muñoz-Elías and McKinney, 2006). Therefore, our knowledge about the substrates used by pathogens during infection, and, equally important, the effect of a bacterial infection on the metabolism of the host cell is still fragmentary (Joseph and Goebel, 2007; Fuchs et al., 2011; Rohmer et al., 2011). Furthermore, the structural conservation of metabolic enzymes was considered to prevent the identification of microbe-specific inhibitors. Key metabolic enzymes, however, that are specifically required during growth within host cells could constitute a promising new set of possible targets for antibacterial compounds urgently needed or be used for the development of food formulas that suppress growth of pathogenic bacteria (Boigegrain et al., 2005; Becker et al., 2006; Liautard et al., 2006). Recent progress has been made in determining the major carbon sources used by intracellularly replicating pathogens such as *Listeria monocytogenes*, *Shigella flexneri*, and pathogenic *Escherichia coli* (Lucchini et al., 2005; Eylert et al., 2008; Götz and Goebel, 2010; Götz et al., 2010). These data suggest that pathogens, in order to efficiently replicate within a host or its cells, have to coordinate their metabolism with the availability of nutrients during their life cycle (for review, see Eisenreich et al., 2010).

Listeria monocytogenes is a Gram-positive pathogen that mainly affects immunocompromised individuals, pregnant women, and newborns. Severe infections are characterized by bacteremia, meningoencephalitis, abortion, or neonatal sepsis. The most common vehicles of transmission of this saprophytic bacterium to humans are dairy products and other foods including eggs, seafood, and vegetables. Three hundred eighty-six documented cases of listeriosis were reported for 2010 in Germany (Robert Koch-Institut, 2011), and about 1600 in the USA (Centers for Disease Control and Prevention, 2011). The high lethality rate of up to 20–30% despite early antibiotic treatment resulted in

Abbreviations: σ^B, sigma factor; BBCAAs, branched-chain amino acids; GAP, glyceraldehyde-3-phosphate; Hly, listeriolysin; IPA, isotopolog profiling analysis; Pdh, pyruvate dehydrogenase; PlcA/PlcB, phospholipase A/B; PPP, pentose phosphate pathway; PrfA, positive regulatory factor A.

increasing efforts to understand listerial pathogenicity and to find tools against this pathogen (Vázquez-Boland et al., 2001).

Upon uptake by contaminated food, *L. monocytogenes* enter non-phagocytic cells such as epithelial cells, hepatocytes, or fibroblasts by the activity of the surface-associated internalins A and B. In contrast to other facultative intracellular pathogens like *Salmonella*, a hallmark of *L. monocytogenes* is that it is capable to escape from the phagocytic vacuole by disrupting the phagosomal membrane via the expression of listeriolysin (Hly) and phospholipase A (PlcA). Listerial cells thus access the host cell cytoplasm where they are not only able to replicate, but also to actively move by actin polymerization mediated by ActA. Cell-to-cell spreading and subsequent disruption of the vacuolar double-membrane by Hly and PlcB has also been observed. Having passed the gut epithelium, *L. monocytogenes* is capable to resist killing by professional phagocytes. It might disseminate via the lymph and the blood to the liver and the spleen and even cross the blood–brain or the blood–placenta barrier. All main virulence factors are under control of the positive regulatory factor A (PrfA; for more details, see reviews such as Vázquez-Boland et al., 2001; Dussurget et al., 2004; Hamon et al., 2006; Cossart and Toledo-Arana, 2008; Camejo et al., 2011).

Here, we will summarize recent omic-studies relevant for the topic of listerial metabolism during infection, and introduce isotopolog profiling analysis (IPA) as a technique that allows novel insights in metabolic fluxes during infection. Then, metabolic adaptations and requirement of *L. monocytogenes* in cultured cells and *in vivo*, as well as the underlying regulatory factors, will be resumed. Modeling approaches as further tools that pave the way toward a systems level understanding of listerial metabolism during infection will be presented, followed by challenges and future perspectives in this research field.

OMIC-APPROACHES TO DELINEATE METABOLIC TRAITS RELEVANT FOR INFECTION

So far, mainly two strategies have been followed to improve our systemic understanding of metabolic adaptations by listeriae within host compartments: firstly, the differences between the transcriptomes and the proteomes of apathogenic and pathogenic strains, and secondly the analysis of differential gene expression upon the transition from one physiologically relevant condition to another.

COMPARATIVE APPROACHES

Comparative genomics is important to identify factors and pathways contributing to virulence properties, both on the genus level with insights into host specificities and on the species level regarding the biodiversity of certain bacterial lineages (Zhao et al., 2011). Detection of differences between the genomes of *L. monocytogenes* and the apathogenic species *L. innocua* gave first insights into the reasons that cause a *Listeria* species to be pathogenic (Buchrieser et al., 2003). In a triple analysis involving the whole genome sequences of *L. monocytogenes*, *L. welshimeri*, and *L. innocua*, it was shown that genome reduction led to the apathogenic *L. welshimeri*, the latter being derived from early evolutionary events. This finding points to an ancestor more compact than *L. monocytogenes* (Hain et al., 2006).

However, the genetic equipment itself does not sufficiently describe differences of more closely related strains with respect to their virulence properties. Several proteomic approaches have been performed in the context of listerial adaption to the host environment, thus contributing to the systemic understanding of metabolism during infection (Cabanes et al., 2011). A comparative *in vitro* proteomic approach investigated the protein expression profiles of *L. monocytogenes* and *L. innocua* with a focus on the secretome of both species (Trost et al., 2005). *In vitro* comparative transcriptome analysis of *L. monocytogenes* strains revealed differences of the two major lineages/serovar 1/2a, and serovars 4b and 1/2b including stress-related sigma factor B (see below) and virulence factors (Severino et al., 2007). Related studies analyzed the secreted proteomes of *L. monocytogenes* strains belonging to serovars 4b, 1/2b, and 1/2a (Dumas et al., 2008, 2009a,b). Due to the identification of factors possibly involved in substrate degradation, those studies might reveal novel insight into the listerial metabolism *in vivo*. Donaldson et al. (2011) focused on proteome differences between an avirulent and two virulent *L. monocytogenes* strains representing the two lineages mentioned above. Their data revealed that most proteins of the intermediary metabolism are stronger expressed in pathogenic strains in comparison to apathogenic strains. Comparison between the two virulent serovars also revealed metabolic differences in their intramacrophagic proteome, possibly reflecting unequal proliferation rates (Donaldson et al., 2009).

TRANSITION STUDIES

Two pilot studies had investigated the transcriptome of *L. monocytogenes* infecting human epithelial (Caco-2) cells and the murine macrophage cell line P388D1. Both studies revealed that up to 19% of the listerial genes are differentially expressed in comparison to their level of transcription in BHI medium (Chatterjee et al., 2006; Joseph et al., 2006). Differences in the results point to cell-specific metabolic adaptations during the intracellular replication of *L. monocytogenes*, whereas common findings support the assumption that several metabolic traits play a central role for listerial replication *in vivo* (see below). The analysis of the proteome of *L. monocytogenes* strain EGDe replicating in macrophages demonstrated the upregulation of specific metabolic pathways (van de Velde et al., 2009). An *in vitro* proteomic study revealed global changes in gene expression when *L. monocytogenes* enters stationary phase, a growth condition also relevant *in vivo* (Weeks et al., 2004). These data clearly indicate that specific metabolic adaptations significantly contribute to the capability of pathogens to replicate within macrophages.

The analysis of the bacterial response to changing conditions *in vitro* also elucidates the dynamic of metabolism *in vivo*. Wen et al. (2011) investigated the reversible transition from the bacilli-like to the cocci-morphology in the long-term-survival (LTS) phase of *L. monocytogenes* serotype 4b strain F2365 at different growth stages in tryptic soy broth with yeast extract (TSBYE). Transcriptome analysis identified 225 differentially expressed genes ($\geq$4-fold; $P < 0.05$) with the upregulation of metabolic genes including those involved in the synthesis of branched-chain amino acids (BCAAs). Combined treatment with potassium lactate and sodium diacetate led to altered metabolism, including a shift toward fermentative production of acetoin (Stasiewicz et al., 2011).

POWER OF ISOTOPOLOG PERTURBATION STUDIES FOR SYSTEMS BIOLOGY OF INFECTION

The study of biosynthetic pathways and fluxes under *in vivo* conditions is crucial to understand the metabolism and physiology of microorganisms on a systems level (Winder et al., 2011). While many important features on metabolism can be deduced on the basis of genome sequences, RNA transcripts, protein and metabolite profiles, and numeric modeling, the direct observation of metabolic pathways and fluxes during listerial growth in eukaryotic host cells has long been hampered by the lack of an adequate technology. One of the recent methods for the quantitative analysis of metabolite fluxes throughout biological systems is based on growing infected host cells in medium containing stable isotope labeled nutrients (Eylert et al., 2008). For heterotrophic organisms including pathogenic bacteria such as *L. monocytogenes* and their eukaryotic host cells, simple carbohydrates, amino acids, and glycerol are among the typical carbon sources and, therefore, serve as suitable tracers in these labeling studies.

Any perturbation of the natural isotope equilibrium by the supply of such an isotope enriched compound will naturally spread in the experimental system via a large number of enzyme catalyzed reactions. In other words, a broad spectrum of biosynthetic pathways (if not all) is covered by this experimental approach. As a consequence, comprehensive information about the nature and the dynamics of the metabolic network can be obtained. Analytical methods to monitor the distribution of the stable isotope (e.g., ^{13}C) in biosynthetic products are NMR spectroscopy and/or mass spectrometry typically coupled with gas chromatography (GC/MS). Both methods have their specific advantages and disadvantages. Mass spectrometry is a well-established tool to determine isotopolog patterns in amino acids at adequate sensitivity (Zamboni et al., 2009). Signals with high signal-to-noise ratio can be obtained even with small sample amounts, i.e., 10^8 bacterial cells or 1 mg of dry cell pellet is sufficient for the measurement of protein bound amino acids (Eisenreich et al., 2010).

After protein hydrolysis, the resulting amino acids are converted into *tert*-butyldimethylsilyl-derivatives (TBDMS amino acids). On the basis of the mass patterns, the relative fractions of isotopomers (i.e., molecular ions or fragments thereof comprising a defined number of ^{13}C-atoms) can be determined for approximately 50 mass fragments of 15 TBDMS amino acids under realistic conditions. However, from these analytes only 29 fragments of 12 different TBDMS amino acids (glycine, alanine, serine, aspartate, threonine, glutamate, valine, leucine, isoleucine, methionine, tyrosine, and phenylalanine) meet the demands for a reliable isotopomer quantification (Antoniewicz et al., 2007). Only 13 fragments comprise all carbon atoms of the original amino acids (i.e., for glycine, alanine, serine, aspartate, threonine, glutamate, proline, methionine, valine, lysine, histidine, phenylalanine, and tyrosine), whereas all other fragments are observed after loss of one or more carbon atoms from the original amino acid. For most TBDMS amino acids, fragments are detected where C-1 (the carboxylic atom) has been lost during the ionization. Provided that the labeling strategy had generated specific profiles in all of these fragments, the biosynthetic pathways leading to the analyzed amino acids can be identified in typical studies. Moreover, the same data can be used to predict the patterns in the precursors for the respective amino acids as a basis to elucidate the fluxes in the central intermediary metabolisms. For accurate data, the overall ^{13}C-enrichment should be at least 0.2%; better results are obtained with enrichments >1%.

Notably, the measurements can be done in a high throughput manner by automated systems, albeit data processing, and interpretation still need considerable efforts by expert users even when supported by the available software to deconvolute isotopolog enrichments from the original data. Considering the isotope fractions in the molecular masses of the original metabolites and one or more fragments thereof, some limited information can be gleaned about the positional distribution of the ^{13}C-label. However, the positional resolution is still low in comparison with isotopolog profiling by NMR spectroscopy that, in turn, is worse in sensitivity (Eisenreich and Bacher, 2007). Nevertheless, the isotopomer patterns in amino acids detected by GC/MS already reveal considerable information about the carbon fluxes in the bacterial cells and their hosts (see below).

When more than 10^{12} bacterial cells are available (>10 mg of dry cell pellet), one- and two-dimensional NMR spectroscopy can be used in addition to GC/MS. Indeed, high-resolution ^{13}C-NMR spectroscopy is capable to assess ^{13}C-enrichment for each non-equivalent carbon atom of a metabolite under study. Isotopomers carrying one ^{13}C-atom at different positions display completely different ^{13}C-NMR spectra with singlet signals at the chemical shifts for the respective labeled carbon atoms. Notably, this is in sharp contrast to mass spectrometry, where each of these isotopomers gives rise to identical signals due to the identical molecular masses. The same holds true for multiply ^{13}C-labeled isotopologs. Whereas the mass spectra only show the sum of all isotopologs comprising two or more ^{13}C-atoms, respectively, these isotopologs can be better distinguished by NMR spectroscopy. Due to scalar ^{13}C^{13}C couplings, the ^{13}C-NMR signals of a multiply ^{13}C-labeled metabolite appear as specific multiplets in the spectra.

On the basis of the ^{13}C-NMR coupling signatures observed for every single (non-equivalent) carbon atom in a ^{13}C-labeled compound, information on the abundance of ^{12}C and ^{13}C at the respective neighbored carbon atoms is obtained (i.e., in a framework comprising not more than three to four bonds around the index atom). As a result, a set of isotopolog groups can be identified for each ^{13}C-NMR signal. On the basis of the signal intensities (i.e., by deconvolution of the complex coupling patterns providing integral values for each component in the ^{13}C-NMR multiplets), these sets are quantified. For most very small non-symmetrical molecules (i.e., comprising not more than three carbon atoms), all isotopomers/isotopologs display specific NMR signal patterns and can be clearly assigned and quantified on this basis (Eisenreich and Bacher, 2007).

Due to the fact that many of the potential long-range ^{13}C-couplings cannot be completely resolved in the ^{13}C-NMR spectrum of a more complex metabolite (comprising more than three carbon atoms), the observables are typically not sufficient for the direct observation of all individual isotopologs. However, in combination with GC/MS analysis, a sufficient number of constraints can be determined for amino acids to clearly assign molecules with single and multiple labels at adjacent carbon positions. As described below, isotopolog profiling studies have

indeed contributed important and novel insights into the listerial metabolism during infection.

THE INTRACELLULAR NICHE

A hallmark of *L. monocytogenes* is its capability to replicate within host cells. Thus, auxotrophic mutants indicate possible nutrient limitations within a host cell. For example, listerial strains defective in the biosynthesis of aromatic compounds as precursors for menaquinone show a strong attenuation in cell culture and mice infection assays, and a more slightly *in vivo* attenuation was observed for mutants requiring the three aromatic amino acids, threonine, and adenine (Alexander et al., 1993; Marquis et al., 1993; Stritzker et al., 2004). These studies indicate that hosts provide sufficient organic and inorganic compounds to overcome selected auxotrophies. In addition, proteins involved in nucleotide synthesis (PyrD, PyrE, PyrF, and PyrAB) are expressed at a lower level during replication of *L. monocytogenes* within human THP-1 monocytes or macrophages, indicating a sufficient supply of nucleotides from the host cell (Klarsfeld et al., 1994; van de Velde et al., 2009). The reduced replication rate of a *pyrE* mutant in macrophages (Schauer et al., 2010) again might point to specificities of host cell lines with respect to their metabolic status.

The cytosol serves as a compartment that protects pathogens from killing by the host's immune system. Its composition with respect to metabolites is largely unknown, as well as the intracytosolic availability of nutrients. Typically, the cytosol exhibits low concentrations of magnesium, sodium, iron, and calcium ions at neutral pH (reviewed in Ray et al., 2009). In total, up to 100 suitable carbon as well as nitrogen, phosphorus, and sulfur sources are present within eukaryotic cells (Brown et al., 2008). The intracellular milieu is a reducing environment that contains 10 mM reduced glutathione and only 0.5 mM oxidized glutathione (Hwang et al., 1992), a fact important for growth of *L. monocytogenes* that depends on the availability of reduced nitrogen and sulfur sources (Joseph and Goebel, 2007).

It was assumed that facultative intracellular bacteria have developed a balanced strategy to exploit just enough nutrients from the host cell to maintain survival and proliferation in the intracellular niche (Joseph and Goebel, 2007). A prerequisite for such a strategy, however, is that the intracellular compartment is rich in nutrients, a hypothesis for which the following pros and cons might be stated:

Pros: Strikingly, non-pathogenic bacteria such as *Bacillus subtilis* or *E. coli* expressing Hly are able to replicate within the cytosol following phagocytic uptake and escape from the vacuole (Bielecki et al., 1990; Monack and Theriot, 2001). The vacuolar passage, for example by a low pH, might prime the bacterium for intracellular replication since microinjected non-pathogenic *L. innocua*, *E. coli*, *Salmonella enterica* serovar Typhimurium, *Yersinia enterocolitica*, or *B. subtilis* failed to proliferate within the cytosol (Goetz et al., 2001; Slaghuis et al., 2004; Hain et al., 2008). Cons: In case cells are a habitat providing easy access to substrates, much more pathogens beside *L. monocytogenes*, *S. flexneri*, *Burkholderia pseudomallei*, and *Francisella tularensis* would exploit this niche (Ray et al., 2009). Furthermore, all intracellular pathogens exhibit a preference for certain cell types that might depend on differences in the cytosolic nutrient composition. For example, the pathogenic strains EGDe and F2365 could be more metabolically active in J774.1 than in P388D1 macrophages (Chatterjee et al., 2006).

LISTERIAL METABOLISM WITHIN CELLS

Occupation of the intracellular niche plays a pivotal role for *L. monocytogenes* virulence, allowing the pathogen to temporarily escape from the immune system of the host and to initiate systemic infection. The host cell cytosol with cell-specific and dynamic nutrient composition is a complex environment that this intracellular pathogen actively adapts to (Eisenreich et al., 2010). Here, we will resume what is known about the listerial mechanisms of substrate acquisition, and the specific metabolic adaptations of *L. monocytogenes* that allow its successful transition from the extra- to the intracellular milieu (**Figure 1**).

CARBON SOURCES

The intracellular utilization of glucose or glucose-6-phosphate, but also of non-carbohydrate nutrients such as glycerol and amino acids has been demonstrated by IPA studies (Eylert et al., 2008). A multiple mutant unable to take up and catabolize glycerol due to a lack of GlpF, glycerol kinase (GlpK), glycerol-3-phosphate dehydrogenase (GlpD), and DhaK is also attenuated in intracellular growth (own unpublished data). This is in line with the upregulation of these genes within the cytoplasm of host cells (Chatterjee et al., 2006; Joseph et al., 2006) and within the intestine of mice (Toledo-Arana et al., 2009). *In vitro*, genes encoding glycolysis enzymes and genes involved in the metabolism and the biosynthesis of BCAA are downregulated, and those required for gluconeogenesis are upregulated in the presence of glycerol. All PrfA-dependent genes show a higher level of transcription under these conditions (Joseph et al., 2008). On the other hand, a strain lacking Hpt, a transporter involved in the exploitation of hexose phosphate from the host cell, is attenuated *in vivo* and within cells, and its gene, *uhpT*, revealed to be highly upregulated during growth in intracellular strains (Chatterjee et al., 2006; Joseph et al., 2006). In addition, the decreased transcription of *hpr* encoding the phosphocarrier protein of a phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS) during intraepithelial replication reflects the lower amount of Hpr in the absence of glucose (Asanuma and Hino, 2003). These data support the assumption that glucose-6-phosphate rather than glucose is used for bacterial metabolism within hosts (Chico-Calero et al., 2002). This was confirmed by the observation of carbon catabolite derepression during intracellular growth (Joseph et al., 2006), and the finding that a mutant unable to uptake glucose does not show replication attenuation in macrophages or epithelial cells (Stoll and Goebel, 2010). Instead of glycolysis, the oxidative pentose phosphate pathway (PPP) is assumed to be the predominant pathway of sugar metabolism within host cells, since the genes required for glycolysis are down- and those for PPP are upregulated (Chatterjee et al., 2006; Joseph et al., 2006). The non-oxidative branch of PPP is also upregulated in epithelial cells and results in the production of xylulose- and ribose-5-phosphate as precursors of nucleotide biosynthesis (Chatterjee et al., 2006). This was not supported by proteomic analysis that revealed a decreased abundance of two enzymes of this pathway, namely the transaldolase lmo2743 and ribose-5-phosphate epimerase during listerial replication in

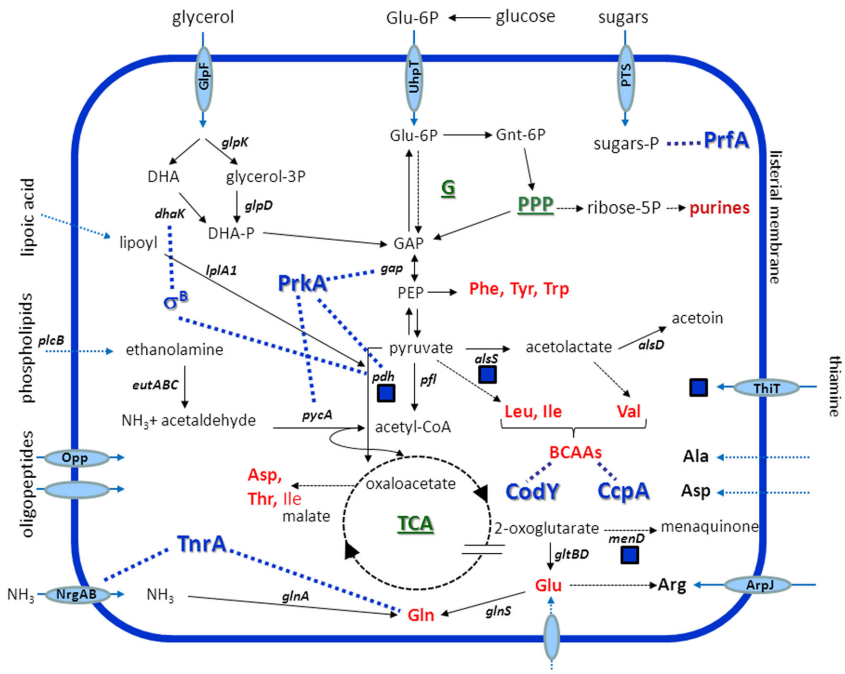


FIGURE 1 | Simplified view of listerial metabolic enzymes, transporters, and pathways relevant during infection. Cofactor thiamine is symbolized by a blue square. Dashed lines indicate putative interactions of master regulators with transporters, enzymes, or metabolites, dashed arrows unknown uptake mechanisms. Blue, shaded genes encode putative interaction partners of PrkA. *De novo* biosynthesized amino acids are shown in red, complex pathways in green. Abbreviations: G, glycolysis; PEP, phosphoenolpyruvate; P, phosphate. Functions of genes are mentioned in the text.

human THP-1 monocytes (van de Velde et al., 2009). This contradiction might be due to cell-type specific metabolite compositions. Another possible explanation was only recently been provided by a proteomic comparison between two pathogenic strains EGDe and F2365 of *L. monocytogenes* with the non-pathogenic strain HCC23. The pattern of intermediary metabolism proteins suggests that an initial period of reduced glycolysis during intracellular replication is followed by a resumption of glycolysis (Donaldson et al., 2011).

NITROGEN METABOLISM

Glutamine, convertible to glutamate, is the optimal nitrogen source for *L. monocytogenes in vitro* (Fisher, 1999). Little is known about the *in vivo* nitrogen metabolism of *L. monocytogenes* that is assumed to use ammonium, arginine, and/or ethanolamine as nitrogen sources during intracellular replication (Buchrieser et al., 2003; Joseph et al., 2006; Joseph and Goebel, 2007). NrgAB that is activated by TnrA at low concentrations of nitrogen sources is responsible for the uptake of ammonium ions. Gene *nrgAB* is strongly downregulated within murine macrophages (Chatterjee et al., 2006), but upregulated in human epithelial cells (Joseph et al., 2006). These data again suggest a cell-type dependent availability of nitrogen sources and/or cell-type specific metabolic adaptations of *L. monocytogenes*.

Ethanolamine lyase EutABC has been demonstrated to be required for wildtype-like intracellular replication of *L. monocytogenes* (Joseph et al., 2006). Ethanolamine might be derived from phoshatidylethanolamine by the activity of listerial phospholipases. Thus, phospholipids ubiquitous in host environments such as the gut might serve as important nitrogen, carbon, and energy source during infection. This hypothesis is in line with the finding that ethanolamine can be utilized under anaerobic conditions only (Price-Carter et al., 2001; Winter et al., 2010; Srikumar and Fuchs, 2011). The role of host peptides in nitrogen metabolism (see below) remains to be investigated in more detail.

AMINO ACID UPTAKE AND *DE NOVO* BIOSYNTHESIS

At least six listerial transporters possibly involved in the uptake of amino acids and oligopeptides were recently identified to contribute to intramacrophagic survival and replication (Joseph et al., 2006; Schauer et al., 2010). Induced expression of the oligopeptide transporter lmo0135–0137 associated with cysteine transport has been demonstrated to contribute to proliferation in Caco-2 cells and *in vivo* (Chatterjee et al., 2006; Schauer et al., 2010). In addition, the oligopeptide permease OppABCDF is not only a prerequisite for intramacrophagic growth, but also for full virulence of *L. monocytogenes* in mice (Borezee et al., 2000; Port and Freitag, 2007). These findings support the assumption that *L. monocytogenes* uses intracellular peptides as a source of amino acids (Marquis et al., 1993).

Transcriptomic data showed that genes involved in the biosynthesis of glutamate, glutamine, and arginine belonging to the glutamate family are upregulated during intraepithelial replication of *L. monocytogenes* (Joseph et al., 2006). Similar findings were made for tryptophan and the BCAAs valine, leucine, and isoleucine. BCAAs are among the most abundant amino acids in proteins, and maintaining their pools is a prerequisite for high-level synthesis of

proteins (Sonenshein, 2007). The downregulation of aminoacyl tRNA synthase genes *glyS*, *serS*, *cysS*, *alaS*, *hisS*, *valS*, *thrS*, *ileS*, *leuS*, *tyrS*, and *trpS* suggests that the respective amino acids are available for *L. monocytogenes* within macrophages (Chatterjee et al., 2006). In a pilot study, this hypothesis was at least partially confirmed by ^{13}C-isotopolog profiling experiments, demonstrating that up to 50% of the amino acids alanine, aspartate, and glutamate are recruited from the host cells during intracellular replication of *L. monocytogenes* (Eylert et al., 2008). Results of the same study indicate a restricted anabolism under these conditions since only 7 of 14 amino acids investigated, namely alanine, aspartate, glutamate, serine, threonine, valine, and glycine ordered by decreased ^{13}C-incorporation, were identified to be *de novo* biosynthesized (Eylert et al., 2008).

CENTRAL METABOLISM

A hallmark of the central metabolism of *L. monocytogenes* is the incomplete tricarboxylic acid (TCA) cycle due to a lack of 2-oxoglutarate dehydrogenase as demonstrated by genome analysis and ^{13}C-labeling studies (Glaser et al., 2001; Eisenreich et al., 2006). 2-Oxoglutarate as a substrate of glutamate synthase links carbon and nitrogen metabolism. Vice versa, amino acids such as arginine, proline, and histidine can be converted to 2-oxoglutarate.

It could be demonstrated that oxaloacetate, the direct or indirect precursor of aspartate and threonine, is generated by an anaplerotic reaction afforded by pyruvate carboxylation (Eylert et al., 2008). This reaction is catalyzed by the ATP-dependent pyruvate carboxylase PycA (Schär et al., 2010). The critical role for PycA for proliferation of *L. monocytogenes in vivo* and within epithelial cells or macrophages makes it an interesting candidate for antilisterial compound search (Schindler and Zähner, 1972). It is worth to note that pyruvate is also the precursor for the synthesis of BCAAs that, as demonstrated by several studies cited here, play a pivotal role during intracellular replication of listeriae. Thus, the biosynthesis of BCAAs that is mainly regulated by catabolite control protein A (CcpA) and CodY in an opposite manner also balances the distribution of pyruvate and 2-oxoglutarate (Sonenshein, 2007).

Oxidative decarboxylation of pyruvate results in acetyl-coenzyme A used for the synthesis of valine and leucine, whereas isoleucine is more indirectly derived via pyruvate carboxylation to oxaloacetate as a precursor for aspartate and threonine synthesis. These data suggest a competition of two metabolic pathways important for intracellular replication, namely the biosynthesis of BCAA and the replenishing of the TCA cycle. This metabolic sink might partially be compensated by the increased and decreased expression of several enzymes involved in amino acid metabolism, among them D-alanyl-D-alanine ligase and D-alanine transaminase, an enzyme essential for virulence (Johnson et al., 2004; van de Velde et al., 2009).

COFACTORS

The cofactors riboflavin, thiamine, biotin, and lipoate are supplements of defined growth media for *L. monocytogenes.* The vitamin thiamine is required by enzymes of the central metabolism including Pdh, and the need of its *de novo* biosynthesis by *L. monocytogenes* suggests its intracellular limitation (Schauer et al., 2009). Its precursors might be derived from glutamine or from 2-oxoglutarate by overexpressed glutamate dehydrogenase (van de Velde et al., 2009). The same authors also argued that the unexpected induced expression of PurQ and PurM results in the formation of 1-(5′-phosphoribosyl)-5-amino-imidazole that might predominantly be used as a thiamine precursor. Indeed, insertional knockout of *purQ* attenuated listerial replication inside macrophages (Schauer et al., 2010). Interestingly, Madeo and coworkers recently identified thiamine to play a yet unknown role in listerial tolerance toward low pH as encountered within host compartments. This was explained by the fact that the thiamine-dependent conversion of pyruvate to acetolactate and further to acetoin comprises two proton-consuming steps (Madeo et al., 2012). The first step is catalyzed by the product of *alsS* induced under acidic conditions (Stasiewicz et al., 2011). The use of host-derived lipoic acid is also relevant for *L. monocytogenes* infection. Lipoyl modification of Pdh subunit E2 by lipoate protein ligase LplA1 was demonstrated to contribute to intracellular replication and virulence of this pathogen (O'Riordan et al., 2003).

METABOLISM OF *L. MONOCYTOGENES IN VIVO*

Fundamental shifts in the expression pattern of genes are involved in the adaptation of listerial metabolism to several host compartments. A comprehensive approach was recently performed using tiling arrays to describe the listerial RNome during transition of *L. monocytogenes* from saprophytism to virulence. To this end, the authors not only studied three physiologically relevant conditions, namely stationary phase, hypoxia, and temperature shifts, but also investigated sigma B (σ^B)- and PrfA-negative mutants (Toledo-Arana et al., 2009). Two PTS required for mannitol and mannose uptake, and ribose-5-phosphate isomerase as part of the non-oxidative PPP are also upregulated in intestine and stationary phase. A remarkable metabolic adaptation to the intestine and the blood is the increased transcription of the gene cluster involved in ethanolamine and 1,2-propanediol utilization including the cofactor cobalamine (see above), of PTS systems for galactitol, fructose, and cellobiose uptake, and of DhaKs involved in glycerol utilization. Interestingly, the gene clusters lmo0315–lmo0318 responsible for thiamine biosynthesis, and lmo1983–1991 involved in the biosynthesis of BCAAs are also upregulated during listerial replication in blood (Toledo-Arana et al., 2009). This resembles previous findings that these metabolic traits are required for intracellular replication (Joseph et al., 2006). In the host intestinal lumen, the authors observed a σ^B-mediated activation of genes, whereas in blood, gene transcription was mainly controlled by PrfA.

Listeria monocytogenes also actively replicates within the spleen of infected mice. Following intravenous application of strain EGDe, 30% of all genes differentially regulated in mice spleen were found to be involved in metabolism (Camejo et al., 2009). Upregulated genes include *uhpT* and, in contrast to epithelial cells and macrophages, genes involved in glycolysis (*gap*, *pgi*, *fbaA*, *pgm*), whereas those involved in the non-oxidative phase of the PPP appeared to be downregulated. Genes involved in the expression of the Pdh complex including its activator, lipoate protein ligase, were induced *in vivo*. The upregulation of genes encoding pyruvate–formate lyase indicates an anoxic degradation of pyruvate to formate and acetyl-CoA in the spleen. The upregulation of

genes responsible for the biosynthesis of aromatic amino acids (*aroA*, *pheA*), of BCAAs and of amino acids of the aspartate and glutamate families is in line with the transcription pattern in epithelial cells and macrophages. An increased expression of mannose-, maltose-, and cellobiose-specific PTS and a decreased expression of PTS responsible for fructose, galactitol, and mannitol uptake was observed. The induction of *glnA* encoding glutamine synthetase might hint to an inactivation of TnrA within the spleen. Remarkably, listerial genes involved in thiamine synthesis are also upregulated during proliferation in mice spleen.

The human intestine is a yet underinvestigated compartment with respect to its impact on bacterial metabolism. Here, *L. monocytogenes* not only encounters predominantly anaerobic conditions as a prerequisite for ethanolamine utilization, but also a higher concentration of carbon dioxide known to act as a signal for bacterial metabolism (Neidhardt et al., 1974; Kröger et al., 2011). It has been hypothesized that the biosynthesis of threonine via oxaloacetate and overexpression of threonine synthase is favored by the high concentration of carbon dioxide (van de Velde et al., 2009). Carbon dioxide is also required for oxidative carboxylation of glucose-6-phosphate, the first reaction in the PPP.

Some of the data on the listerial metabolism in cell culture on the one hand and *in vivo* on the other hand, for example the role of glycolysis, the non-oxidative PPP or the uptake of sugars, are inconsistent. These discrepancies not only indicate to a more active multiplication status in mouse organs compared to cell culture studies using immortal cells (Camejo et al., 2009), but also point to compartment-specific metabolic needs.

REGULATION OF LISTERIAL METABOLISM DURING INFECTION

During invasive infection, *L. monocytogenes* not only proliferates within epithelial cells and macrophages, but encounters a frequent change of the surrounding milieu. Therefore, *in vivo* growth is assumed to require the concerted activity of the following regulators:

CodY, TnrA, AND CcpA

Three global regulators of *L. monocytogenes*, CcpA, TnrA, and CodY sense key metabolites generated in carbon- and nitrogen metabolism (Fisher, 1999; Sonenshein, 2007). CodY, a global transcriptional regulator, seems to play a critical role in the intracellular adaptation of *L. monocytogenes*. CodY is known to lose its repressing activity when intracellular levels of GTP and BCAAs decrease. As shown by Bennett et al. (2007), the CodY regulon comprises genes involved in amino acid metabolism, nitrogen assimilation, and sugar uptake. Derepression of the CodY operon revealed to be essential for listerial replication in various cell culture models and in mice. The authors therefore hypothesized that CodY senses the metabolic state of the host cell cytoplasm and triggers an adaptive listerial response, which is characterized by the induction of genes involved in the *de novo* biosynthesis of amino acids such as glutamate, arginine, histidine, tryptophan, and serine, in the intermediary carbon metabolism, in the uptake of sugars or in peptide, and ammonium transport. The induction and requirement of *argD* encoding *N*-acetylornithine aminotransferase during intracellular replication simultaneously with the induced expression of the arginine transporter ArpJ might reflect a limited availability of this amino acid under intracellular conditions (Klarsfeld et al., 1994; Joseph et al., 2006). Listerial TnrA, probably encoded by lmo1298, is known as the principal global regulator of nitrogen metabolism in *B. subtilis* and represses the synthesis of glutamate synthase (Glt) and glutamine synthetase (GlnA). TnrA expression revealed to be downregulated within epithelial cells (Chatterjee et al., 2006), a finding that links to the inactivation of TnrA in the presence of glutamine (Wray et al., 2001). Less is known about the role of the listerial catabolite CcpA. Repression of sugar utilization pathways by glucose is independent of CcpA (Gopal et al., 2010), and CcpA is not involved in carbon source regulation of virulence genes (Behari and Youngman, 1998).

ALTERNATIVE SIGMA FACTOR B (σ^B)

The transition of *L. monocytogenes* from the saprophytic to the pathogenic lifestyle requires an adequate response to environmental conditions such as low pH, bile stress, carbon starvation, and the presence of reactive oxygen species on the level of transcription and expression (Abram et al., 2008a; Hain et al., 2008; O'Byrne and Karatzas, 2008; Ryan et al., 2008; Soni et al., 2011; Zhang et al., 2011). The regulon controlled by σ^B plays a major role in this adaptation process. Moreover, σ^B has been demonstrated to contribute to invasion of epithelial cells and to gastrointestinal infection (Kim et al., 2004; Garner et al., 2006; McGann et al., 2007; Ollinger et al., 2009). Interestingly, the two most highly represented categories of σ^B-dependent factors were transport and metabolism proteins, among them pyruvate dehydrogenase Pdh, a dihydroxyacetone kinase DhaK, the glycerol transporter GlpF, and a mannose-specific PTS (Kazmierczak et al., 2003). Accordingly, a *L. monocytogenes* σ^B mutant lacking σ^B less efficiently used glycerol as a carbon and energy source (Abram et al., 2008b). An intensive transcriptional reorganization was observed in an *in vivo* study based on tiling microarrays in which a σ^B mutant was used for infection (Toledo-Arana et al., 2009).

POSITIVE REGULATORY FACTOR A

The main virulence regulator of *L. monocytogenes*, PrfA, belongs to the family of cAMP receptor protein (Crp)/fumarate nitrate reductase regulators (de las Heras et al., 2011). PrfA not only activates nine key virulence factors including Hpt, but also controls the expression of further 136 factors including those involved in the metabolic activity of this pathogen during infection (Milohanic et al., 2003). The direct or indirect regulatory PrfA function requires an interplay with σ^B that controls the general stress response of *L. monocytogenes* (Ollinger et al., 2008, 2009). Interestingly, there is a close link between metabolism and virulence in *L. monocytogenes* since the carbon sources utilized by the pathogen affect the expression of virulence genes. Especially, glucose, fructose, mannose, or cellobiose transported by PTS downregulated PrfA-dependent genes (Freitag et al., 2009), a finding that points to a low concentration of these sugars within cells. Vice versa, overproduction of PrfA resulted in growth inhibition in glucose-containing media and in an increased *de novo* biosynthesis of BCAA (Eisenreich et al., 2006; Marr et al., 2006). This sugar-mediated gene repression of PrfA-dependent genes depends rather on inhibition of PrfA than on changes of the PrfA concentration

(Renzoni et al., 1997). Taken together, the intracellular availability of carbon sources might act as a key signal for virulence gene expression in *L. monocytogenes* (de las Heras et al., 2011).

ADDITIONAL REGULATORY MECHANISMS

Another mechanism that regulates the metabolic adaptation of pathogens to their hosts is the phosphorylation status of proteins. In *L. monocytogenes*, a eukaryotic-type serine/threonine-kinase (PrkA, lmo1820) and a serine/threonine-phosphatase (Stp, lmo1821) have been identified that play a role during intracellular replication and *in vivo*, respectively (Archambaud et al., 2005; Zemansky et al., 2009). In a proteomic approach, 62 proteins were identified as interacting partners of PrkA including 19 involved in carbohydrate and amino acid metabolism (Lima et al., 2011). PrkA directly influences important metabolic enzymes such as Pdh, glyceraldehyde-3-phosphate (GAP) dehydrogenase, pyruvate carboxylase PycA, an aminotransferase involved in BCAA synthesis, and the acetate kinase Acs. The role of universal stress proteins (Usps), which support the adaptation to energy deficiencies, in listerial stress resistance and in virulence has only recently been demonstrated (Seifart Gomes et al., 2011). In *E. coli*, the activation of *uspA* is regulated by fructose-6-phosphate (Persson et al., 2007), but a similar mechanism in Gram-positive pathogens remains to be elucidated.

MODELING *L. MONOCYTOGENES* METABOLISM UNDER INFECTION CONDITIONS

Systems biology approaches of listerial infection use a combination of genome-based bioinformatics tools including sequence and domain analysis, function and structure prediction for proteins, and phylogenetic comparisons. The focus of modeling has been on (intracellular) metabolism, pathogenicity factors, host–pathogen interaction, and the immune response of the host (Schauer et al., 2010; Sauer et al., 2011).

To model listerial metabolism *in vitro* and within cells, elementary mode analysis was applied (Eisenreich et al., 2006). This method allows to enumerate all possible enzyme combinations or independent pathways within the bacterial cell. Each pathway balances all internal metabolites involved. Furthermore, the metabolic flux models allows to fill in gaps of knowledge from the retrobiosynthetic analysis of the isotopolog compositions, for instance areas where no labeled metabolites were transported by the metabolic flux.

In a more recent study, quantitative literature data on intracellular replication of defined mutants was combined with the replication rates derived from a genome-scale screening for mutants showing reduced intraepithelial survival of *L. monocytogenes* (Schauer et al., 2010). The application of extreme pathway and elementary mode analysis revealed a critical role of glycerol and purine metabolism, of fucose utilization, and of the synthesis of glutathione, aspartate semialdehyde, serine, and BCAAs during intracellular replication of *L. monocytogenes*. Thus, new targets for antibiotic intervention became visible by this approach. Furthermore, the modeling demonstrated that degradation of glucose indeed occurs to a large extent via the PPP. To achieve those insights in metabolic traits relevant for listerial infection, numerous all-against-all protein sequence comparisons were performed to establish specific and general protein families.

The number of metabolic pathways to consider in such genome-scale models reaches several thousand and more. Two criteria helped here to zoom in on the key pathways involved: Firstly, only those mutants of *L. monocytogenes* were considered that showed wildtype-like replication in rich medium, but impaired growth in macrophages.

Secondly, the pathways critical for survival in the macrophage were analyzed to particularly identify enzymes with a role in at least two of these pathways. The list of key enzymes reoccurring in the affected elementary modes by the different affected mutants could then be calculated and ranked. By this procedure, it was possible to quickly identify the most important listerial pathways required for intracellular survival.

Modeling metabolic pathways often results in combinatorial explosion, i.e., an intractable high number of potential flux modes. However, recent advances have the potential to overcome this disadvantage. For instance, flux distributions were decomposed into elementary flux modes in genome-scale metabolic networks of *E. coli* grown in rich medium containing various carbon sources (Chan and Ji, 2011). Furthermore, software such as EMILiO increases the scope of strain design to include reactions with individually optimized fluxes. Unlike existing approaches experiencing an explosion in complexity to solve this problem, this allowed to generate diverse strain designs regarding production of succinate, L-glutamate, and L-serine (Yang et al., 2011). A third possibility is the application of metabolic flux patterns. These are sets of reactions representing the basic routes through a particular subsystem that are compatible with admissible fluxes in a (possibly) much larger metabolic network (Schuster et al., 2010). The subsystem can be made up by reactions one is interested in, for example production of a certain metabolite. Furthermore, growth, growth boundary, and inactivation models of *L. monocytogenes* have extensively been developed in food research. Validated approaches exploited available experimental data and combined all three aspects using on different food sources as a medium (Coroller et al., 2012). Stochastic-based examples are multiplicative heteroscedastic models taking into account differing growth variance for heterogeneous populations (Cao et al., 2010), or the dependence of successful growth on inoculum and cell counts (Vermeulen et al., 2009).

PERSPECTIVES

Each of the approaches reviewed above has obvious limitations. Interfacing the transcriptome analysis with a mutant library screening underscored that the up- or downregulation of a gene or an operon not necessarily coincides with a phenotype under the same conditions (Joseph et al., 2006), a general drawback of these approaches. On the other hand, a proteomic analysis does not reveal the whole proteome under given conditions. Mutant analysis and IPA, on the other hand, focus on the role of single genes and metabolic traits, thus providing parts of the metabolic puzzle only. Contradictory results derived from different approaches posed the relevance of single metabolic pathways of *L. monocytogenes* for infection in question. Examples are the overexpression

of glutamate dehydrogenase during listerial replication in human THP-1 monocytes in contrast to its decreased expression in strain F2365 (van de Velde et al., 2009; Donaldson et al., 2011), or the induction or repression of the non-oxidative branch of the PPP (see above). These contradictions not only point to host cell-specific or temporary metabolic adaptations of *L. monocytogenes*, but also to the need of more standardized experimental conditions in omic-approaches.

Another prerequisite for systems biology of listerial metabolism is to quantify and qualify the metabolic state of the bacterial and the host cells during different stages of infection (Zamboni and Sauer, 2009). So far, data sets on metabolites have been derived only from *in vitro* approaches under non-physiological conditions. For example, *L. monocytogenes* 10403S cells grown at 37 and 8°C were analyzed (Singh et al., 2011). At low temperature, an increase of solute concentrations such as amino acids, sugars, organic acids, urea cycle intermediates, polyamines, and different compatible solutes was observed in the listerial cytoplasm, thus lowering the freezing point of intracellular water and decreasing ice crystal formation. A statistical analysis (PCA) was instrumental to reveal this system response. Mathematical modeling can consider such effects by adjusting metabolite concentrations and by thermodynamic considerations. Furthermore, compartment models are soon expected to improve such studies (Cheng et al., 2008) including pools and fluxes of the mammalian cell (Lopes et al., 2010). Future research should also consider microcompartments and organelles that both are used to optimize metabolic processes. Compartmentalization is common also in *L. monocytogenes* including carboxysomes responsible for cobalamine-dependent 1,2-propanediol degradation, thus contributing to functional diversity.

To reveal an even more comprehensive picture of listerial metabolism during infection, cells such as primary macrophages and fibroblasts should be considered for future cell culture experiments. It is also of interest to determine whether and to what extent the host metabolism is modulated during infection, for instance by *L. monocytogenes* secreting c-di-AMP (Woodward et al., 2010). In addition to *in vivo* studies using mice and guinea pigs (Cabanes et al., 2008), small animals such as *Caenorhabditis elegans* or *Galleria mellonella* might be established as new model organisms. *C. elegans* is a facile and proven model host for the study of microbial pathogenesis and metabolism (Sifri et al., 2005; Thomsen et al., 2006; Srikumar and Fuchs, 2011). Because both host and pathogen are easily amenable to genetic manipulations, the effect of metabolic perturbations upon infection, changing culture conditions, or gene knockouts can be investigated in this infection model. Notably, the model also reduces costs and ethical constraints in comparison to infection experiments with higher animals. A further advantage of the model is that metabolic fluxes from cells or tissues of an infected host animal to *L. monocytogenes* and vice versa might be followed, an approach that has long been hampered by the lack of technology to identify and quantify metabolism throughout an appropriate biological system (Spanier et al., 2009). *G. mellonella*, the greater wax moth, is another relatively simple, non-mammalian host model system that can be used to assess not only the virulence of listerial strains, but also the contribution of metabolic capacities to larvae killing as exemplified by the attenuated phenotype of an EGDe Δ*uhpT* mutant (Mukherjee et al., 2010). A major disadvantage in contrast to nematodes, however, is the need of subcutaneous instead of oral application of the pathogen.

SUMMARY

Listeria monocytogenes is highly adapted to the cytoplasm of mammalian host cells where it is able to multiply with a generation time comparable to that in rich medium. A comprehensive analysis of its metabolism is a prerequisite for a systemic understanding of *L. monocytogenes* infection. Taking into account all omic-studies and mutant data, certain metabolic enzymes and pathways of *L. monocytogenes* revealed to have a critical role during infection and might serve as new targets for the development of antilisterial compounds (**Figure 1**). Metabolic clues for intracellular and *in vivo* replication are the degradation of phosphorylated glucose via the PPP, the utilization of C_3-bodies as alternative carbon- and energy source, pyruvate as a sink for central metabolism including downstream reactions such as pyruvate decarboxylation, biosynthesis of BCAAs, the availability of thiamine, and the acquisition of host cell derived nitrogen sources including ethanolamine. The regulation of these metabolic capabilities involves the activity of factors such as PrfA and σ^B, resulting in a fine-tuned balance of metabolic flows with the host cell and *in vivo*. For a deeper understanding especially of host–pathogen interdependencies with respect to metabolism, new technologies such as isotopolog profiling, additional animal models including invertebrates, and novel systems biological approaches have to be combined in future research.

ACKNOWLEDGMENTS

Work in the authors' laboratories is supported by grants of the German Research Foundation (DFG), including the priority program SPP1316 (DA 208/13-1, EI-384/6, and FU-375/5).

REFERENCES

Abram, F., Starr, E., Karatzas, K. A., Matlawska-Wasowska, K., Boyd, A., Wiedmann, M., Boor, K. J., Connally, D., and O'Byrne, C. P. (2008a). Identification of components of the sigma B regulon in *Listeria monocytogenes* that contribute to acid and salt tolerance. *Appl. Environ. Microbiol.* 74, 6848–6858.

Abram, F., Su, W.-L., Wiedmann, M., Boor, K. J., Coote, P., Botting, C., Karatzas, K. A. G., and O'Byrne, C. P. (2008b). Proteomic analyses of a *Listeria monocytogenes* mutant lacking sigma B identify new components of the sigma B regulon and highlight a role for sigma B in the utilization of glycerol. *Appl. Environ. Microbiol.* 74, 594–604.

Alexander, J. E., Andrew, P. W., Jones, D., and Roberts, I. S. (1993). Characterization of an aromatic amino acid-dependent *Listeria monocytogenes* mutant: attenuation, persistence, and ability to induce protective immunity in mice. *Infect. Immun.* 61, 2245–2248.

Antoniewicz, M. R., Kelleher, J. K., and Stephanopoulos, G. (2007). Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. *Anal. Chem.* 79, 7554–7559.

Archambaud, C., Gouin, E., Pizarro-Cerda, J., Cossart, P., and Dussurget, O. (2005). Translation elongation factor EF-Tu is a target for Stp, a serine-threonine phosphatase involved in virulence of *Listeria monocytogenes*. *Mol. Microbiol.* 56, 383–396.

Asanuma, N., and Hino, T. (2003). Molecular characterization of HPr and related enzymes, and regulation of HPr phosphorylation in the ruminal bacterium *Streptococcus bovis*. *Arch. Microbiol.* 179, 205–213.

Becker, D., Selbach, M., Rollenhagen, C., Ballmaier, M., Meyer, T. F., Mann, M., and Bumann, D. (2006). Robust *Salmonella* metabolism limits possibilities for new antimicrobials. *Nature* 440, 303–307.

Behari, J., and Youngman, P. (1998). A homolog of CcpA mediates catabolite control in *Listeria monocytogenes* but not carbon source regulation of virulence genes. *J. Bacteriol.* 180, 6316–6324.

Bennett, H. J., Pearce, D. M., Glenn, S., Taylor, C. M., Kuhn, M., Sonenshein, A. L., Andrew, P. W., and Roberts, I. S. (2007). Characterization of *relA* and *codY* mutants of *Listeria monocytogenes*: identification of the CodY regulon and its role in virulence. *Mol. Microbiol.* 63, 1453–1467.

Bielecki, J., Youngman, P., Connelly, P., and Portnoy, D. A. (1990). *Bacillus subtilis* expressing a haemolysin gene from *Listeria monocytogenes* can grow in mammalian cells. *Nature* 345, 175–176.

Boigegrain, R.-A., Liautard, J.-P., and Köhler, S. (2005). Targeting of the virulence factor acetohydroxyacid synthase by sulfonylureas results in inhibition of intramacrophagic multiplication of *Brucella suis*. *Antimicrob. Agents Chemother.* 49, 3922–3925.

Borezee, E., Pellegrini, E., and Berche, P. (2000). OppA of *Listeria monocytogenes*, an oligopeptide-binding protein required for bacterial growth at low temperature and involved in intracellular survival. *Infect. Immun.* 68, 7069–7077.

Brown, S. A., Palmer, K. L., and Whiteley, M. (2008). Revisiting the host as a growth medium. *Nat. Rev. Microbiol.* 6, 657–666.

Buchrieser, C., Rusniok, C., Kunst, F., Cossart, P., and Glaser, P. (2003). Comparison of the genome sequences of *Listeria monocytogenes* and *Listeria innocua*: clues for evolution and pathogenicity. *FEMS Immunol. Med. Microbiol.* 35, 207–213.

Cabanes, D., Lecuit, M., and Cossart, P. (2008). Animal models of *Listeria* infection. *Curr. Protoc. Microbiol.* 9, 9B.1.1–9B.1.17.

Cabanes, D., Sousa, S., and Cossart, P. (2011). "*Listeria* genomics," in *Genomics of Foodborne Bacterial Pathogens*, eds M. Wiedmann and W. Zhang (New York: Springer), 141–70.

Camejo, A., Buchrieser, C., Couvé, E., Carvalho, F., Reis, O., Ferreira, P., Sousa, S., Cossart, P., and Cabanes, D. (2009). In vivo transcriptional profiling of *Listeria monocytogenes* and mutagenesis identify new virulence factors involved in infection. *PLoS Pathog.* 5, e1000449. doi:10.1371/journal.ppat.1000449

Camejo, A., Carvalho, F., Reis, O., Leitao, E., Sousa, S., and Cabanes, D. (2011). The arsenal of virulence factors deployed by *Listeria monocytogenes* to promote its cell infection cycle. *Virulence* 2, 379–394.

Cao, R., Francisco-Fernández, M., and Quinto, E. J. (2010). A random effect multiplicative heteroscedastic model for bacterial growth. *BMC Bioinformatics* 11, 77. doi:10.1186/1471-2105-11-77

Centers for Disease Control and Prevention. (2011). *Statistics of Listeriosis*. Atlanta. Available at: http://www.cdc.gov/listeria/statistics.html

Chan, S. H. J., and Ji, P. (2011). Decomposing flux distributions into elementary flux modes in genome-scale metabolic networks. *Bioinformatics* 27, 2256–2262.

Chatterjee, S. S., Hossain, H., Otten, S., Kuenne, C., Kuchmina, K., Machata, S., Domann, E., Chakraborty, T., and Hain, T. (2006). Intracellular gene expression profile of *Listeria monocytogenes*. *Infect. Immun.* 74, 1323–1338.

Cheng, S., Liu, Y., Crowley, C. S., Yeates, T. O., and Bobik, T. A. (2008). Bacterial microcompartments: their properties and paradoxes. *Bioessays* 30, 1084–1095.

Chico-Calero, I., Suárez, M., González-Zorn, B., Scortti, M., Slaghuis, J., Goebel, W., and Vázquez-Boland, J. A. (2002). Hpt, a bacterial homolog of the microsomal glucose-6-phosphate translocase, mediates rapid intracellular proliferation in *Listeria*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 431–436.

Coroller, L., Kan-King-Yu, D., Leguerinel, I., Mafart, P., and Membré, J.-M. (2012). Modelling of growth, growth/no-growth interface and nonthermal inactivation areas of *Listeria* in foods. *Int. J. Food Microbiol.* 152, 139–152.

Cossart, P., and Toledo-Arana, A. (2008). *Listeria monocytogenes*, a unique model in infection biology: an overview. *Microbes Infect.* 10, 1041–1050.

de las Heras, A., Cain, R. J., Bielecka, M. K., and Vazquez-Boland, J. A. (2011). Regulation of *Listeria* virulence: PrfA master and commander. *Curr. Opin. Microbiol.* 14, 118–127.

Donaldson, J. R., Nanduri, B., Burgess, S. C., and Lawrence, M. L. (2009). Comparative proteomic analysis of *Listeria monocytogenes* strains F2365 and EGD. *Appl. Environ. Microbiol.* 75, 366–373.

Donaldson, J. R., Nanduri, B., Pittman, J. R., Givaruangsawat, S., Burgess, S. C., and Lawrence, M. L. (2011). Proteomic expression profiles of virulent and avirulent strains of *Listeria monocytogenes* isolated from macrophages. *J. Proteomics* 74, 1906–1917.

Dumas, E., Desvaux, M., Chambon, C., and Hébraud, M. (2009a). Insight into the core and variant exoproteomes of *Listeria monocytogenes* species by comparative subproteomic analysis. *Proteomics* 9, 3136–3155.

Dumas, E., Meunier, B., Berdagué, J.-L., Chambon, C., Desvaux, M., and Hébraud, M. (2009b). The origin of *Listeria monocytogenes* 4b isolates is signified by subproteomic profiling. *Biochim. Biophys. Acta* 1794, 1530–1536.

Dumas, E., Meunier, B., Berdagué, J.-L., Chambon, C., Desvaux, M., and Hébraud, M. (2008). Comparative analysis of extracellular and intracellular proteomes of *Listeria monocytogenes* strains reveals a correlation between protein expression and serovar. *Appl. Environ. Microbiol.* 74, 7399–7409.

Dussurget, O., Pizarro-Cerda, J., and Cossart, P. (2004). Molecular determinants of *Listeria monocytogenes* virulence. *Annu. Rev. Microbiol.* 58, 587–610.

Eisenreich, W., and Bacher, A. (2007). Advances of high-resolution NMR techniques in the structural and metabolic analysis of plant biochemistry. *Phytochemistry* 68, 2799–2815.

Eisenreich, W., Dandekar, T., Heesemann, J., and Goebel, W. (2010). Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nat. Rev. Microbiol.* 8, 401–412.

Eisenreich, W., Slaghuis, J., Laupitz, R. ., Bussemer, J., Stritzker, J., Schwarz, C., Schwarz, R., Dandekar, T., Goebel, W., and Bacher, A. (2006). 13C isotopologue perturbation studies of *Listeria monocytogenes* carbon metabolism and its modulation by the virulence regulator PrfA. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2040–2045.

Eylert, E., Schär, J., Mertins, S., Stoll, R., Bacher, A., Goebel, W., and Eisenreich, W. (2008). Carbon metabolism of *Listeria monocytogenes* growing inside macrophages. *Mol. Microbiol.* 69, 1008–1017.

Fisher, S. H. (1999). Regulation of nitrogen metabolism in *Bacillus subtilis*: vive la différence! *Mol. Microbiol.* 32, 223–232.

Freitag, N. E., Port, G. C., and Miner, M. D. (2009). *Listeria monocytogenes* – from saprophyte to intracellular pathogen. *Nat. Rev. Microbiol.* 7, 623–628.

Fuchs, T. M., Eisenreich, W., Heesemann, J., and Goebel, W. (2011). Metabolic adaptation of human pathogenic and related nonpathogenic bacteria to extra- and intracellular habitats. *FEMS Microbiol. Rev.* doi: 10.1111/j.1574-6976.2011.00301.x. [Epub ahead of print].

Garner, M. R., Njaa, B. L., Wiedmann, M., and Boor, K. J. (2006). Sigma B contributes to *Listeria monocytogenes* gastrointestinal infection but not to systemic spread in the guinea pig infection model. *Infect. Immun.* 74, 876–886.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couvé, E., de Daruvar, A., Dehoux, P., Domann, E., Domínguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fsihi, H., García-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gómez-López, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Pérez-Diaz, J. C., Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vázquez-Boland, J. A., Voss, H., Wehland, J., and Cossart, P. (2001). Comparative genomics of *Listeria* species. *Science* 294, 849–852.

Goetz, M., Bubert, A., Wang, G., Chico-Calero, I., Vazquez-Boland, J. A., Beck, M., Slaghuis, J., Szalay, A. A., and Goebel, W. (2001). Microinjection and growth of bacteria in the cytosol of mammalian host cells. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12221–12226.

Gopal, S., Berg, D., Hagen, N., Schriefer, E.-M., Stoll, R., Goebel, W., and Kreft, J. (2010). Maltose and maltodextrin utilization by *Listeria monocytogenes* depend on an inducible ABC transporter which is repressed by glucose. *PLoS ONE* 5, e10349. doi:10.1371/journal.pone.0010349

Götz, A., Eylert, E., Eisenreich, W., and Goebel, W. (2010). Carbon metabolism of enterobacterial human pathogens growing in epithelial colorectal adenocarcinoma

(Caco-2) cells. *PLoS ONE* 5, e10586. doi:10.1371/journal.pone.0010586

Götz, A., and Goebel, W. (2010). Glucose and glucose 6-phosphate as carbon sources in extra- and intracellular growth of enteroinvasive *Escherichia coli* and *Salmonella enterica*. *Microbiology* 156, 1176–1187.

Hain, T., Hossain, H., Chatterjee, S. S., Machata, S., Volk, U., Wagner, S., Brors, B., Haas, S., Kuenne, C. T., Billion, A., Otten, S., Pane-Farre, J., Engelmann, S., and Chakraborty, T. (2008). Temporal transcriptomic analysis of the *Listeria monocytogenes* EGD-e σB regulon. *BMC Microbiol.* 8, 20. doi:10.1186/1471-2180-8-20

Hain, T., Steinweg, C., Kuenne, C. T., Billion, A., Ghai, R., Chatterjee, S. S., Domann, E., Kärst, U., Goesmann, A., Bekel, T., Bartels, D., Kaiser, O., Meyer, F., Pühler, A., Weisshaar, B., Wehland, J., Liang, C., Dandekar, T., Lampidis, R., Kreft, J., Goebel, W., and Chakraborty, T. (2006). Whole-genome sequence of *Listeria welshimeri* reveals common steps in genome reduction with *Listeria innocua* as compared to *Listeria monocytogenes*. *J. Bacteriol.* 188, 7405–7415.

Hamon, M., Bierne, H., and Cossart, P. (2006). *Listeria monocytogenes*: a multifaceted model. *Nat. Rev. Microbiol.* 4, 423–434.

Hwang, C., Sinskey, A. J., and Lodish, H. F. (1992). Oxidized redox state of glutathione in the endoplasmic reticulum. *Science* 257, 1496–1502.

Johnson, J., Jinneman, K., Stelma, G., Smith, B. G., Lye, D., Messer, J., Ulaszek, J., Evsen, L., Gendel, S., Bennett, R. W., Swaminathan, B., Pruckler, J., Steigerwalt, A., Kathariou, S., Yildirim, S., Volokhov, D., Rasooly, A., Chizhikov, V., Wiedmann, M., Fortes, E., Duvall, R. E., and Hitchins, A. D. (2004). Natural atypical *Listeria innocua* strains with *Listeria monocytogenes* pathogenicity island 1 genes. *Appl. Environ. Microbiol.* 70, 4256–4266.

Joseph, B., and Goebel, W. (2007). Life of *Listeria monocytogenes* in the host cells' cytosol. *Microbes Infect.* 9, 1188–1195.

Joseph, B., Mertins, S., Stoll, R., Schar, J., Umesha, K. R., Luo, Q., Muller-Altrock, S., and Goebel, W. (2008). Glycerol metabolism and PrfA activity in *Listeria monocytogenes*. *J. Bacteriol.* 190, 5412–5430.

Joseph, B., Przybilla, K., Stühler, C., Schauer, K., Slaghuis, J., Fuchs, T. M., and Goebel, W. (2006). Identification of *Listeria monocytogenes* genes contributing to intracellular replication by expression profiling and mutant screening. *J. Bacteriol.* 188, 556–568.

Kazmierczak, M. J., Mithoe, S. C., Boor, K. J., and Wiedmann, M. (2003). *Listeria monocytogenes* sigma B regulates stress response and virulence functions. *J. Bacteriol.* 185, 5722–5734.

Kim, H., Boor, K. J., and Marquis, H. (2004). *Listeria monocytogenes* sigma B contributes to invasion of human intestinal epithelial cells. *Infect. Immun.* 72, 7374–7378.

Klarsfeld, A. D., Goossens, P. L., and Cossart, P. (1994). Five *Listeria monocytogenes* genes preferentially expressed in infected mammalian cells: *plcA*, *purH*, *purD*, *pyrE* and an arginine ABC transporter gene, *arpJ*. *Mol. Microbiol.* 13, 585–597.

Kröger, C., Srikumar, S., Ellwart, J., and Fuchs, T. M. (2011). Bistability in myo-inositol utilization by *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* 193, 1427–1435.

Liautard, J.-P., Jubier-Maurin, V., Boigegrain, R.-A., and Köhler, S. (2006). Antimicrobials: targeting virulence genes necessary for intracellular multiplication. *Trends Microbiol.* 14, 109–113.

Lima, A., Duran, R., Schujman, G. E., Marchissio, M. J., Portela, M. M., Obal, G., Pritsch, O., de, M. D., and Cervenansky, C. (2011). Serine/threonine protein kinase PrkA of the human pathogen *Listeria monocytogenes*: biochemical characterization and identification of interacting partners through proteomic approaches. *J. Proteomics* 74, 1720–1734.

Lopes, T. J. S., Luganskaja, T., Vujic Spasic, M., Hentze, M. W., Muckenthaler, M. U., Schümann, K., and Reich, J. G. (2010). Systems analysis of iron metabolism: the network of iron pools and fluxes. *BMC Syst. Biol.* 4, 112. doi:10.1186/1752-0509-4-112

Lucchini, S., Liu, H., Jin, Q., Hinton, J. C. D., and Yu, J. (2005). Transcriptional adaptation of *Shigella flexneri* during infection of macrophages and epithelial cells: insights into the strategies of a cytosolic bacterial pathogen. *Infect. Immun.* 73, 88–102.

Madeo, M., O'Riordan, N., Fuchs, T. M., Utratna, M., Karatzas, K. A. G., and O'Byrne, C. P. (2012). Thiamine plays a critical role in the acid tolerance of *Listeria monocytogenes*. *FEMS Microbiol. Lett.* 326, 137–143.

Marquis, H., Bouwer, H. G., Hinrichs, D. J., and Portnoy, D. A. (1993). Intracytoplasmic growth and virulence of *Listeria monocytogenes* auxotrophic mutants. *Infect. Immun.* 61, 3756–3760.

Marr, A. K., Joseph, B., Mertins, S., Ecke, R., Müller-Altrock, S., and Goebel, W. (2006). Overexpression of PrfA leads to growth inhibition of *Listeria monocytogenes* in glucose-containing culture media by interfering with glucose uptake. *J. Bacteriol.* 188, 3887–3901.

McGann, P., Wiedmann, M., and Boor, K. J. (2007). The alternative sigma factor sigma B and the virulence gene regulator PrfA both regulate transcription of *Listeria monocytogenes* internalins. *Appl. Environ. Microbiol.* 73, 2919–2930.

Milohanic, E., Glaser, P., Coppée, J.-Y., Frangeul, L., Vega, Y., Vázquez-Boland, J. A., Kunst, F., Cossart, P., and Buchrieser, C. (2003). Transcriptome analysis of *Listeria monocytogenes* identifies three groups of genes differently regulated by PrfA. *Mol. Microbiol.* 47, 1613–1625.

Monack, D. M., and Theriot, J. A. (2001). Actin-based motility is sufficient for bacterial membrane protrusion formation and host cell uptake. *Cell. Microbiol.* 3, 633–647.

Mukherjee, K., Altincicek, B., Hain, T., Domann, E., Vilcinskas, A., and Chakraborty, T. (2010). *Galleria mellonella* as a model system for studying *Listeria* pathogenesis. *Appl. Environ. Microbiol.* 76, 310–317.

Muñoz-Elías, E. J., and McKinney, J. D. (2006). Carbon metabolism of intracellular bacteria. *Cell. Microbiol.* 8, 10–22.

Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974). Culture medium for enterobacteria. *J. Bacteriol.* 119, 736–747.

O'Byrne, C. P., and Karatzas, K. A. G. (2008). The role of sigma B (sigma B) in the stress adaptations of *Listeria monocytogenes*: overlaps between stress adaptation and virulence. *Adv. Appl. Microbiol.* 65, 115–140.

Ollinger, J., Bowen, B., Wiedmann, M., Boor, K. J., and Bergholz, T. M. (2009). *Listeria monocytogenes* sigma B modulates PrfA-mediated virulence factor expression. *Infect. Immun.* 77, 2113–2124.

Ollinger, J., Wiedmann, M., and Boor, K. J. (2008). Sigma B- and PrfA-dependent transcription of genes previously classified as putative constituents of the *Listeria monocytogenes* PrfA regulon. *Foodborne Pathog. Dis.* 5, 281–293.

O'Riordan, M., Moors, M. A., and Portnoy, D. A. (2003). *Listeria* intracellular growth and virulence require host-derived lipoic acid. *Science* 302, 462–464.

Persson, O., Valadi, A., Nyström, T., and Farewell, A. (2007). Metabolic control of the *Escherichia coli* universal stress protein response through fructose-6-phosphate. *Mol. Microbiol.* 65, 968–978.

Port, G. C., and Freitag, N. E. (2007). Identification of novel *Listeria monocytogenes* secreted virulence factors following mutational activation of the central virulence regulator, PrfA. *Infect. Immun.* 75, 5886–5897.

Price-Carter, M., Tingey, J., Bobik, T. A., and Roth, J. R. (2001). The alternative electron acceptor tetrathionate supports B12-dependent anaerobic growth of *Salmonella enterica* serovar Typhimurium on ethanolamine or 1,2-propanediol. *J. Bacteriol.* 183, 2463–2475.

Ray, K., Marteyn, B., Sansonetti, P. J., and Tang, C. M. (2009). Life on the inside: the intracellular lifestyle of cytosolic bacteria. *Nat. Rev. Microbiol.* 7, 333–340.

Renzoni, A., Klarsfeld, A., Dramsi, S., and Cossart, P. (1997). Evidence that PrfA, the pleiotropic activator of virulence genes in *Listeria monocytogenes*, can be present but inactive. *Infect. Immun.* 65, 1515–1518.

Robert Koch-Institut. (2011). *Epidemiologisches Bulletin.* Berlin: Robert Koch-Institut.

Rohmer, L., Hocquet, D., and Miller, S. I. (2011). Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. *Trends Microbiol.* 19, 341–348.

Ryan, S., Hill, C., and Gahan, C. G. M. (2008). Acid stress responses in *Listeria monocytogenes*. *Adv. Appl. Microbiol.* 65, 67–91.

Sauer, J.-D., Pereyre, S., Archer, K. A., Burke, T. P., Hanson, B., Lauer, P., and Portnoy, D. A. (2011). *Listeria monocytogenes* engineered to activate the Nlrc4 inflammasome are severely attenuated and are poor inducers of protective immunity. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12419–12424.

Schär, J., Stoll, R., Schauer, K., Loeffler, D. I. M., Eylert, E., Joseph, B., Eisenreich, W., Fuchs, T. M., and Goebel, W. (2010). Pyruvate carboxylase plays a crucial role in carbon metabolism of extra- and intracellularly replicating *Listeria monocytogenes*. *J. Bacteriol.* 192, 1774–1784.

Schauer, K., Geginat, G., Liang, C., Goebel, W., Dandekar, T., and Fuchs, T. M. (2010). Deciphering the intracellular metabolism of *Listeria monocytogenes* by mutant screening

and modelling. *BMC Genomics* 11, 573. doi:10.1186/1471-2164-11-573

Schauer, K., Stolz, J., Scherer, S., and Fuchs, T. M. (2009). Both thiamine uptake and biosynthesis of thiamine precursors are required for intracellular replication of *Listeria monocytogenes*. *J. Bacteriol.* 191, 2218–2227.

Schindler, P. W., and Zähner, H. (1972). Stoffwechselprodukte von Mikroorganismen. *Arch. Mikrobiol.* 82, 66–75.

Schuster, S., de Figueiredo, L. F., and Kaleta, C. (2010). Predicting novel pathways in genome-scale metabolic networks. *Biochem. Soc. Trans.* 38, 1202–1205.

Seifart Gomes, C., Izar, B., Pazan, F., Mohamed, W., Mraheil, M. A., Mukherjee, K., Billion, A., Aharonowitz, Y., Chakraborty, T., and Hain, T. (2011). Universal stress proteins are important for oxidative and acid stress resistance and growth of *Listeria monocytogenes* EGD-e in vitro and in vivo. *PLoS ONE* 6, e24965. doi:10.1371/journal.pone.0024965

Severino, P., Dussurget, O., Vêncio, R. Z. N., Dumas, E., Garrido, P., Padilla, G., Piveteau, P., Lemaître, J.-P., Kunst, F., Glaser, P., and Buchrieser, C. (2007). Comparative transcriptome analysis of *Listeria monocytogenes* strains of the two major lineages reveals differences in virulence, cell wall, and stress response. *Appl. Environ. Microbiol.* 73, 6078–6088.

Sifri, C. D., Begun, J., and Ausubel, F. M. (2005). The worm has turned - microbial virulence modeled in *Caenorhabditis elegans*. *Trends Microbiol.* 13, 119–127.

Singh, A. K., Ulanov, A. V., Li, Z., Jayaswal, R. K., and Wilkinson, B. J. (2011). Metabolomes of the psychrotolerant bacterium *Listeria monocytogenes* 10403S grown at 37˚C and 8˚C. *Int. J. Food Microbiol.* 148, 107–114.

Slaghuis, J., Goetz, M., Engelbrecht, F., and Goebel, W. (2004). Inefficient replication of *Listeria innocua* in the cytosol of mammalian cells. *J. Infect. Dis.* 189, 393–401.

Sonenshein, A. L. (2007). Control of key metabolic intersections in *Bacillus subtilis*. *Nat. Rev. Microbiol.* 5, 917–927.

Soni, K. A., Nannapaneni, R., and Tasara, T. (2011). The contribution of transcriptomic and proteomic analysis in elucidating stress adaptation responses of *Listeria monocytogenes*. *Foodborne Pathog. Dis.* 8, 843–852.

Spanier, B., Lasch, K., Marsch, S., Benner, J., Liao, W., Hu, H., Kienberger, H., Eisenreich, W., and Daniel, H. (2009). How the intestinal peptide transporter PEPT-1 contributes to an obesity phenotype in *Caenorhabditis elegans*. *PLoS ONE* 4, e6279. doi:10.1371/journal.pone.0006279

Srikumar, S., and Fuchs, T. M. (2011). Ethanolamine utilization contributes to proliferation of *Salmonella enterica* serovar Typhimurium in food and in nematodes. *Appl. Environ. Microbiol.* 77, 281–290.

Stasiewicz, M. J., Wiedmann, M., and Bergholz, T. M. (2011). The transcriptional response of *Listeria monocytogenes* during adaptation to growth on lactate and diacetate includes synergistic changes that increase fermentative acetoin production. *Appl. Environ. Microbiol.* 77, 5294–5306.

Stoll, R., and Goebel, W. (2010). The major PEP-phosphotransferase systems (PTSs) for glucose, mannose and cellobiose of *Listeria monocytogenes*, and their significance for extra- and intracellular growth. *Microbiology* 156, 1069–1083.

Stritzker, J., Janda, J., Schoen, C., Taupp, M., Pilgrim, S., Gentschev, I., Schreier, P., Geginat, G., and Goebel, W. (2004). Growth, virulence, and immunogenicity of *Listeria monocytogenes aro* mutants. *Infect. Immun.* 72, 5622–5629.

Thomsen, L. E., Slutz, S. S., Tan, M.-W., and Ingmer, H. (2006). *Caenorhabditis elegans* is a model host for *Listeria monocytogenes*. *Appl. Environ. Microbiol.* 72, 1700–1701.

Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., Loh, E., Gripenland, J., Tiensuu, T., Vaitkevicius, K., Barthelemy, M., Vergassola, M., Nahori, M.-A., Soubigou, G., Régnault, B., Coppée, J.-Y., Lecuit, M., Johansson, J., and Cossart, P. (2009). The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459, 950–956.

Trost, M., Wehmhöner, D., Kärst, U., Dieterich, G., Wehland, J., and Jänsch, L. (2005). Comparative proteome analysis of secretory proteins from pathogenic and nonpathogenic *Listeria* species. *Proteomics* 5, 1544–1557.

van de Velde, S., Delaive, E., Dieu, M., Carryn, S., van Bambeke, F., Devreese, B., Raes, M., and Tulkens, P. M. (2009). Isolation and 2-D-DIGE proteomic analysis of intracellular and extracellular forms of *Listeria monocytogenes*. *Proteomics* 9, 5484–5496.

Vázquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Domínguez-Bernal, G., Goebel, W., González-Zorn, B., Wehland, J., and Kreft, J. (2001). *Listeria* pathogenesis and molecular virulence determinants. *Clin. Microbiol. Rev.* 14, 584–640.

Vermeulen, A., Gysemans, K. P. M., Bernaerts, K., Geeraerd, A. H., Debevere, J., Devlieghere, F., and van Impe, J. F. (2009). Modelling the influence of the inoculation level on the growth/no growth interface of *Listeria monocytogenes* as a function of pH, aw and acetic acid. *Int. J. Food Microbiol.* 135, 83–89.

Weeks, M. E., James, D. C., Robinson, G. K., and Smales, C. M. (2004). Global changes in gene expression observed at the transition from growth to stationary phase in *Listeria monocytogenes* ScottA batch culture. *Proteomics* 4, 123–135.

Wen, J., Deng, X., Li, Z., Dudley, E. G., Anantheswaran, R. C., Knabel, S. J., and Zhang, W. (2011). Transcriptomic response of *Listeria monocytogenes* during the transition to the long-term-survival phase. *Appl. Environ. Microbiol.* 77, 5966–5972.

Winder, C. L., Dunn, W. B., and Goodacre, R. (2011). TARDIS-based microbial metabolomics: time and relative differences in systems. *Trends Microbiol.* 19, 315–322.

Winter, S. E., Thiennimitr, P., Winter, M. G., Butler, B. P., Huseby, D. L., Crawford, R. W., Russell, J. M., Bevins, C. L., Adams, L. G., Tsolis, R. M., Roth, J. R., and Bäumler, A. J. (2010). Gut inflammation provides a respiratory electron acceptor for *Salmonella*. *Nature* 467, 426–429.

Woodward, J. J., Iavarone, A. T., and Portnoy, D. A. (2010). c-di-AMP secreted by intracellular *Listeria monocytogenes* activates a host type I interferon response. *Science* 328, 1703–1705.

Wray, L. V., Zalieckas, J. M., and Fisher, S. H. (2001). *Bacillus subtilis* glutamine synthetase controls gene expression through a protein-protein interaction with transcription factor TnrA. *Cell* 107, 427–435.

Yang, L., Cluett, W. R., and Mahadevan, R. (2011). EMILiO: a fast algorithm for genome-scale strain design. *Metab. Eng.* 13, 272–281.

Zamboni, N., Fendt, S.-M., Rühl, M., and Sauer, U. (2009). (13)C-based metabolic flux analysis. *Nat. Protoc.* 4, 878–892.

Zamboni, N., and Sauer, U. (2009). Novel biological insights through metabolomics and 13C-flux analysis. *Curr. Opin. Microbiol.* 12, 553–558.

Zemansky, J., Kline, B. C., Woodward, J. J., Leber, J. H., Marquis, H., and Portnoy, D. A. (2009). Development of a mariner-based transposon and identification of *Listeria monocytogenes* determinants, including the peptidyl-prolyl isomerase PrsA2, that contribute to its hemolytic phenotype. *J. Bacteriol.* 191, 3950–3964.

Zhang, Q., Feng, Y., Deng, L., Feng, F., Wang, L., Zhou, Q., and Luo, Q. (2011). SigB plays a major role in *Listeria monocytogenes* tolerance to bile stress. *Int. J. Food Microbiol.* 145, 238–243.

Zhao, H., Chen, J., Fang, C., Xia, Y., Cheng, C., Jiang, L., and Fang, W. (2011). Deciphering the biodiversity of *Listeria monocytogenes* lineage III strains by polyphasic approaches. *J. Microbiol.* 49, 759–767.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 December 2011; paper pending published: 30 December 2011; accepted: 13 January 2012; published online: 03 February 2012.

Citation: Fuchs TM, Eisenreich W, Kern T and Dandekar T (2012) Toward a systemic understanding of Listeria monocytogenes metabolism during infection. Front. Microbio. ***3****:23. doi: 10.3389/fmicb.2012.00023*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 12 March 2012
doi: 10.3389/fmicb.2012.00085

An interspecies regulatory network inferred from simultaneous RNA-seq of *Candida albicans* invading innate immune cells

Lanay Tierney[1†], Jörg Linde[2†], Sebastian Müller[2], Sascha Brunke[3,4], Juan Camilo Molina[3], Bernhard Hube[3,5], Ulrike Schöck[6], Reinhard Guthke[2] and Karl Kuchler[1]*

[1] *Christian Doppler Laboratory for Infection Biology, Max F. Perutz Laboratories, Medical University of Vienna, Vienna, Austria*
[2] *Research Group Systems Biology and Bioinformatics, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Jena, Germany*
[3] *Department Microbial Pathogenicity Mechanisms, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Jena, Germany*
[4] *Center for Sepsis Control and Care, Jena University Hospital, Jena, Germany*
[5] *Friedrich Schiller University, Jena, Germany*
[6] *GATC Biotech AG, Konstanz, Germany*

Edited by:
Franziska Mech, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany

Reviewed by:
Thomas Dandekar, University of Wuerzburg, Germany
Oliver Kurzai, Friedrich Schiller University Jena, Germany

***Correspondence:**
Karl Kuchler, Christian Doppler Laboratory for Infection Biology, Max F. Perutz Laboratories, Medical University of Vienna, Campus Vienna Biocenter, Dr. Bohr-Gasse 9/2, A-1030 Vienna, Austria.
e-mail: karl.kuchler@meduniwien.ac.at

†*Lanay Tierney and Jörg Linde have contributed equally to this work.*

The ability to adapt to diverse micro-environmental challenges encountered within a host is of pivotal importance to the opportunistic fungal pathogen *Candida albicans*. We have quantified *C. albicans* and *M. musculus* gene expression dynamics during phagocytosis by dendritic cells in a genome-wide, time-resolved analysis using simultaneous RNA-seq. A robust network inference map was generated from this dataset using Net*Gener*ator, predicting novel interactions between the host and the pathogen. We experimentally verified predicted interdependent sub-networks comprising Hap3 in *C. albicans*, and Ptx3 and Mta2 in *M. musculus*. Remarkably, binding of recombinant Ptx3 to the *C. albicans* cell wall was found to regulate the expression of fungal Hap3 target genes as predicted by the network inference model. Pre-incubation of *C. albicans* with recombinant Ptx3 significantly altered the expression of Mta2 target cytokines such as IL-2 and IL-4 in a Hap3-dependent manner, further suggesting a role for Mta2 in host–pathogen interplay as predicted in the network inference model. We propose an integrated model for the functionality of these sub-networks during fungal invasion of immune cells, according to which binding of Ptx3 to the *C. albicans* cell wall induces remodeling via fungal Hap3 target genes, thereby altering the immune response to the pathogen. We show the applicability of network inference to predict interactions between host–pathogen pairs, demonstrating the usefulness of this systems biology approach to decipher mechanisms of microbial pathogenesis.

Keywords: host–pathogen, RNA-seq, network inference, modeling, reverse engineering, *Candida*, dendritic cells

INTRODUCTION

Both host and pathogenic species have evolved a plethora of strategies to rapidly adapt to the changing environmental dynamics within the infection milieu. However, the extent of this complexity has only recently been investigated through the use of system biology approaches (reviewed in Rizzetto and Cavalieri, 2011). On the molecular level, these adaptations are mediated by complex interaction networks, which sense these environmental changes and transmit the information throughout the cell, leading to a cascade of changes in gene and eventually protein expression. Understanding these underlying interaction networks is important to elucidate how organisms and defense mechanisms interact during microbial infection processes. Genome-wide integrative approaches for modeling have become increasingly popular (Rizzetto and Cavalieri, 2011) due to the availability of high-throughput sequencing technologies, including RNA sequencing (RNA-seq). These technologies now allow for the parallel sequencing of millions of nucleotide sequences simultaneously (Wang et al., 2009; Zhang et al., 2011). One major advantage to using sequencing approach rather than microarrays is that it is a species-independent platform, allowing for an in-depth investigation of non-model organism species, as well as multiple organisms from a single experiment.

In many cases, the underlying interaction networks between the organisms of interest are unknown. Network inference uses reverse engineering techniques (Hecker et al., 2009b; Marbach et al., 2010) to predict unknown interaction networks based on high-throughput gene expression data. A number of approaches have been established to predict inference networks including Bayesian network modeling (Friedman et al., 2000), information theoretical approaches (Butte and Kohane, 2000; Faith et al., 2007), regression based models (D'Haeseleer et al., 1999; Hecker et al., 2009a), and differential equation models (Holter et al., 2001; Guthke et al., 2005, 2007). Biological networks are scale free networks composed of nodes and edges, where nodes represent the objects of interest and edges show the relations between those objects (Le Novere

et al., 2009). Biological interaction networks often use nodes to represent genes or proteins, and edges to show either a direct or indirection interaction, such as protein binding or transcriptional regulation (Barabasi and Oltvai, 2004). Network inference has been successfully applied to a variety of biological scenarios, including the modeling of immune diseases (Guthke et al., 2005; Hecker et al., 2009a), full-genomic models of *Escherichia coli* (Faith et al., 2007), and more recently, small scale networks describing fungal infections (Linde et al., 2010). So far these model have only focused on a single species and have not addressed host–pathogen interactions.

In the present work, we have generated the first interspecies computational model of molecular host–pathogen interactions. We used RNA-seq expression data from an infection time course of *Candida albicans* and bone marrow-derived dendritic cells (BMDCs) from *M. musculus*. *C. albicans* is one of the most prevalent opportunistic human fungal pathogens. Although *C. albicans* normally colonizes the human host, a variety of factors, most notably immune suppression, can lead to dissemination of fungal cells throughout the body. This dissemination can lead to a wide range of diseases, from thrush to multi-organ failure (Gudlaugsson et al., 2003). We focused on dendritic cells as our model host based of their function as antigen-presenting cells, their specialization in pathogen recognition, and their greater role in activating and modulating adaptive immune responses (Netea et al., 2008; Bourgeois et al., 2010). We experimentally verified predicted subnetworks of the interspecies inferred regulatory network, which identifies a role of the transcription factor Hap3 in *C. albicans* during *in vitro* infection. We find that fungal Hap3 is regulated by murine Ptx3, a soluble pattern recognition receptor acting as an opsonin for pathogens (Diniz et al., 2004). We show that Ptx3 binding to *C. albicans* regulates fungal Hap3 target genes, altering the immune response in dendritic cells. Based on the regulation of downstream cytokines and the regulation of *MTA2* mRNA in a Hap3-dependent manner, we provide indirect evidence for a role for Mta2, a member of the nucleosome remodeling and histone deacetylase complex NuRD (Manavathi et al., 2007). We propose a mechanism whereby Ptx3 binding to *C. albicans* leads to cell wall remodeling via fungal Hap3 target genes, thereby changing the ability of the fungi to be recognized by immune cells. The experimental verification of the predicted interspecies interactions is proof-of-principle that network inference can be used to investigate microbial pathogenesis. We suggest that this could be a useful method to identify potential antifungal target genes.

MATERIALS AND METHODS

CANDIDA STRAINS AND GROWTH CONDITIONS

All strains were routinely grown on YPD plates (1% yeast extract, 2% peptone, 2% glucose, 2% agar) and in standard rich media YPD (1% yeast extract, 2% peptone, 2% glucose) for liquid culture at 30˚C. Fungal cells were collected in the logarithmic growth phase by a brief centrifugation, washed in sterile PBS, and diluted for all interaction studies. The following strains were used in this study: *C. albicans* clinical isolate SC5314 (Gillum et al., 1984) and homozygous knock-out of Hap3 (*hap3*Δ/*hap3*Δ) and revertant strain (*hap3*Δ/*hap3*Δ + *CIp10* (*HAP3*, *URA3*), abbreviated in the text as *hap3*Δ/*hap3*Δ + *HAP3*), were generated from the strain BWP17 (*ura3*::*imm434*/*ura3*::*imm434 iro1*/*iro1*::*imm434 his1*::*hisG*/*his1*::*hisG arg4*::*hisG*/*arg4*::*hisG*) by stepwise deletion of both alleles using PCR-amplified *HAP3*::*ARG4* and *HAP3*::*HIS1* cassettes (Gola et al., 2003) and a cIP10 plasmid containing *HAP3* and its promoter and terminator sequences integrated at the RP10 locus (Murad et al., 2000).The homozygous knock-out of *cda2* and revertant were kindly provided by Neil Gow (Aberdeen, UK).

CELL CULTURE OF PRIMARY IMMUNE CELLS FROM MOUSE BONE MARROW

Bone marrow was differentiated to either BMDCs or bone marrow-derived macrophages (BMDMs) from the femurs of 7- to 9-week-old wild type C57BL/6 mice and assessed for homogeneity as previously described using a panel of marker antibodies (Bourgeois et al., 2009).

FUNGAL-MAMMALIAN CELL CO-CULTURE

Fungal-mammalian cell co-cultures were performed as previously described (Bourgeois et al., 2009). Briefly, immune cells were plated at a density of 1.0×10^5 cells/cm^2 in sterile cell culture dishes and incubated with fungal cells at a multiplicity of infection (MOI) of five fungal cells per immune cell. Samples were incubated at 37˚C in 5% CO_2, 95% humidity for up to 24 h.

cDNA PREPARATION FOR RNA-seq

Total RNA was isolated from immune cells and *C. albicans* using the SV total RNA isolation system (Promega, Madison, MI, USA) following manufacturers instructions. To obtain RNA mixtures from both *C. albicans* and BMDCs, cells were first scraped in the provided lysis buffer, followed by homogenization with 200 µl of 0.5 mm acid-washed glass beads (Sigma-Aldrich, St. Louis, MO, USA) in a Fast Prep-24 cooling block at 4˚C (MP Biomedicals Europe, Illkirch, France) for 45 s at 5 m/s. Ribosomal RNA was depleted from 10 µg of pooled total RNA samples using the RiboMinus eukaryote kit for RNA-seq (Invitrogen, Carlsbad, CA, USA) and concentrated using the corresponding RiboMinus Concentration Module (Invitrogen) following manufactures instructions for three independent biological repeats. For each sample, 1 µg of ribosomal-depleted RNA was converted into cDNA using the SMARTer PCR cDNA Synthesis kit and the Advantage 2 polymerase mix (Clontech, Mountain View, CA, USA). PCR amplifications were performed on 1/10 of the first strand synthesis reaction for 18 cycles of 90˚C for 1 min, 95˚C for 15 s, 65˚C for 30 s, and 68˚C for 6 min on a GeneAmp PCR system 9700 (Applied Biosystems, Carlsbad, CA, USA), and purified on ChromaSpin columns (Clontech, Mountain View, CA, USA). The resulting cDNAs were sequenced on the Genome Analyzer IIx at GATC (Konstanz, Germany) using 36 bp, single run, indexed read mode.

SEQUENCE READ MAPPING, PRE-PROCESSING, AND DATA NORMALIZATION

All sequencing reads were mapped using TopHat 1.2.0 (Trapnell et al., 2009) against the SC5314 *C. albicans* assembly 21 (Skrzypek et al., 2010) and the *M. musculus* UCSC version mm9 from the ENSEMBL database (Flicek et al., 2011). Mapping was carried out using the default settings in which only unique hits were kept for

further analysis. The gene expression and normalization analysis was performed as previously described (Mortazavi et al., 2008). Genes were tested for differential expression using the bioconductor package baySeq (Hardcastle and Kelly, 2010) relative to the 0-min infection time point. The analysis was carried out for *C. albicans* and *M. musculus* genes individually.

CLUSTERING AND OVER-REPRESENTED GENE ONTOLOGY TERMS

Fuzzy c-means clustering (Bezdek, 1992) was applied to the two expression matrices of differentially expressed genes from *C. albicans* and *M. musculus*. The optimal number of clusters was estimated as previously described (Guthke et al., 2005; Linde et al., 2010). Functional categorization and significantly over-represented categories were identified using the tool FungiFun (Priebe et al., 2011). All four hierarchical levels of Funcat (Ruepp et al., 2004) and Gene Ontology (Ashburner et al., 2000) categorization were used in this study.

NETWORK INFERENCE PREDICTION AND MEASURING INTERACTION ROBUSTNESS

Network inference was performed as previously described using the Net*Gene*rator tool (Guthke et al., 2005; Linde et al., 2010). Briefly, Net*Gene*rator is based on a set of linear differential equations and models the temporal change of the expression intensity $x_i(t)$ of gene i $(i = 1 \ldots n)$ at time t as the weighted sum of the expression intensities of all other genes and an external stimulus $u(t)$ at time t. The external stimulus $u(t)$ is modeled as a stepwise constant function representing the change from no host–pathogen interaction to the onset of the interaction. The tool aims to identify a network structure, which best fits to the measured RNA-seq data, while it minimizes the number of predicted interactions (Guthke et al., 2005). Thus, a sparse network is inferred.

Net*Gene*rator offers the possibility to integrate prior knowledge (i.e., putative regulatory interactions based on additional data besides the initial time series expression data). Based on the confidence of the prior knowledge source, it is possible to score each proposed interaction. The confidence of the prior knowledge is based on the level of experimentation used to verify a specific interaction and the number of independent experiments showing the same interaction. Since different data sources might be contradictory, prior knowledge is softly integrated, i.e., if a proposed interaction contradicts the measured data too much it can be removed by Net*Gene*rator. Furthermore, the tool may add new interactions not covered by the prior knowledge in order to fit to the measured data. In this study, prior knowledge from public databases was softly integrated (Guthke et al., 2005). Each proposed interaction was scored in an additive manner based on the confidence of the prior knowledge source as follows: direct evidence that a gene is involved in a host–pathogen interaction (confidence score = 0.5), co-expression of two genes (confidence score = 0.25), and the occurrence of the respective transcription factor binding motif in the upstream intergenic regions of genes (confidence score = 0.125). Prior knowledge was obtained from GeneMania[1], IntAct (Aranda et al., 2010), BioGrid (Stark et al., 2011), the *C. albicans* database (Skrzypek et al., 2010), the mouse genome database (Blake et al., 2011), and a number of peer-reviewed publications (Lane et al., 2001; Doedt et al., 2004; Martchenko et al., 2004; Zhao et al., 2004; Fradin et al., 2005; Oberholzer et al., 2006; Wang et al., 2006; Spira et al., 2007; Thewes et al., 2007; Zakikhany et al., 2007; Almeida et al., 2008; Baek et al., 2008; Nobile et al., 2008; Frohner et al., 2009; Griffin et al., 2009; Raman et al., 2009; Sellam et al., 2009; Hinze et al., 2010; Hou et al., 2010; Smith et al., 2010; Wachtler et al., 2011) summarized in **Figure 1C**. Putative regulatory interactions were tested for robustness using two methods. First, Gaussian noise was introduced with a mean of 0 and SD 0.05 to the estimated mRNA concentrations for 1000 iterations. Secondly, predicted interactions were screened for robustness against changes in prior knowledge by iterating the modeling approach 1000 times while randomly skipping 10% of all interactions in the set of prior knowledge for each run. Only edges that were confirmed by more than 50% of the iterations were considered to be robust and used in the resultant model.

REAL-TIME qPCR ANALYSIS

RNA sample preparation, reverse transcription, and real-time PCR were performed previously described (Bourgeois et al., 2009) using the following primers: mouse β-*Actin*, forward 5′-GCGTGACATCAAAGAGAAG-3′ reverse 5′-AGGAGCCAGAGCAGTAATC-3′ (RTPrimerDB)[2] mouse *MTA2*, forward 5′-CACTGCTATAGCCTCACGCC-3′, reverse 5′-GCTAG GAGCTGGAACC TCAC-3′, mouse *PTX3*, forward 5′-CCTGCTTT GTGCTCTCTGGT-3′, reverse 5′-TCTCCAGCATGATGAACAGC-3′ (Diniz et al., 2004), *C. albicans TUP1*, forward 5′-GACTACGCCTCAAACGAAGC-3′ reverse 5′-TGGTGCCACAAT CTGTTGTT-3′, *C. albicans FRE6* forward 5′-CCGGTAAACATCC ATTCCAC-3′, reverse 5′-TTGATCCAAATGCCATT-CAA-3′, *C. albicans SEF1*, forward 5′-GTGGAGGACTCGTTCATGGT-3′, reverse 5′-TGAACCAGCACGATTCAGAG-3′, *C. albicans RIP1*, forward 5′-TGCTGACAGAGTCAAGA-AACC-3′ reverse 5′-GAACCAACCACCGAAATCAC-3′ as determined using the sequence analysis software Vector NTI (Invitrogen, Carlsbad, CA, USA). Results were calculated using the $\Delta\Delta ct$ method and are expressed as the fold of the gene expression of interest versus the expression of a housekeeping gene in *M. musculus* (β-*Actin*) or *C. albicans* (*RIP1*) in treated versus untreated conditions.

CYTOKINE QUANTIFICATION FROM CO-CULTURE SUPERNATANTS

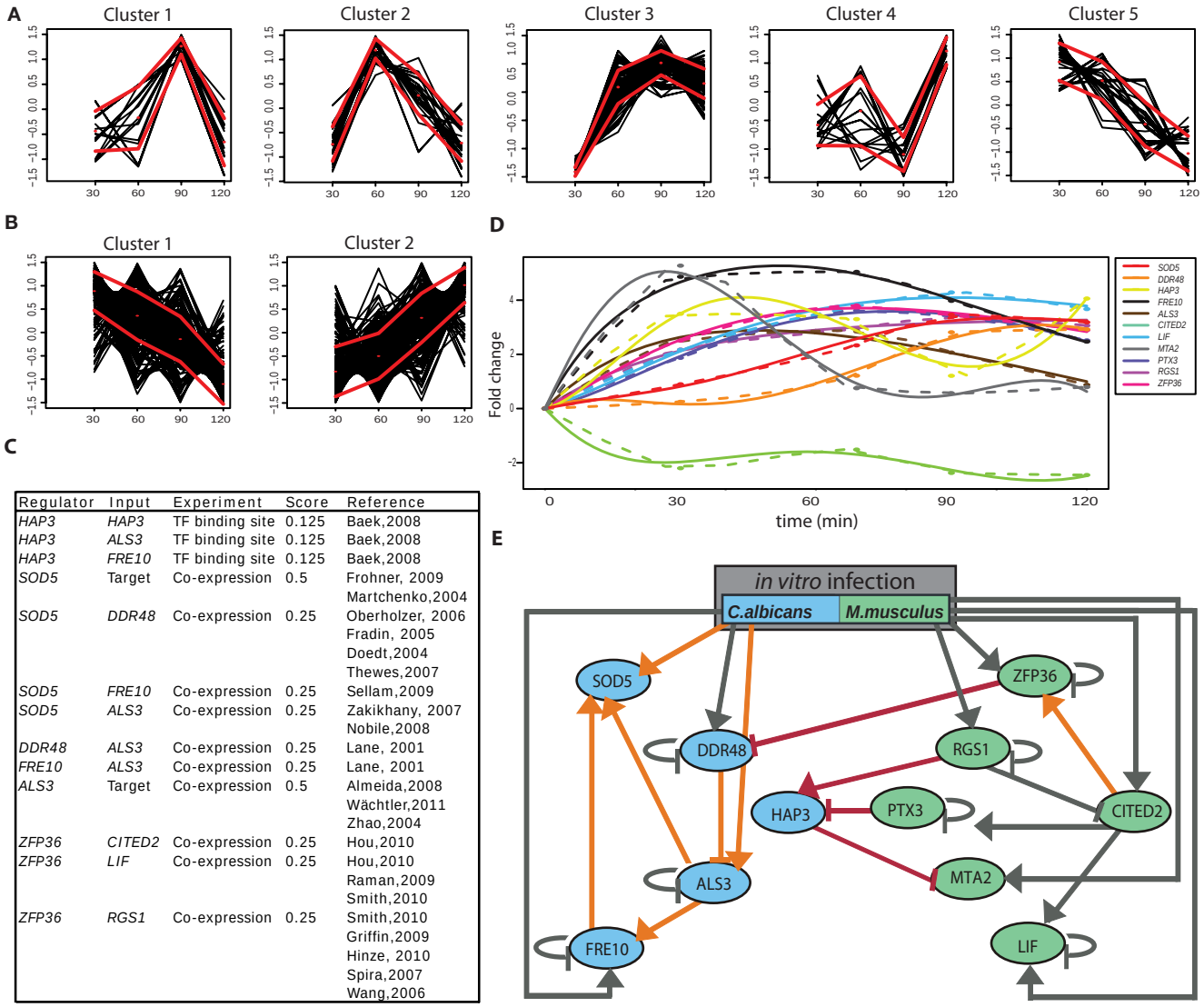
The amount of cytokines released into cell culture supernatants by immune cells during *in vitro* interaction studies with heat killed *C. albicans* were assayed after 24 h of co-culture using the mouse IL-2, IL-4, or TNFα Ready-set-go ELISA kit (R&D Systems, Minneapolis, MN, USA) or the Mouse Cytokine Array Panel A kit (R&D Systems) according to the manufacturers instructions.

BINDING AND LABELING Ptx3 *IN VITRO*

Recombinant mouse Ptx3 (rmPtx3) protein (R&D Systems) was reconstituted in sterile PBS and diluted for all experiments. Some 0.5×10^6 fungal cells were incubated for 1 h at 37°C with 5 μg reconstituted rmPtx3. Ptx3 was labeled with the primary antibody against Ptx3 (Abnova, Taiwan) and secondarily

[1] http://genemania.org/data/

[2] http://medgen.ugent.be/rtprimerdb/index.php



Regulator	Input	Experiment	Score	Reference
HAP3	*HAP3*	TF binding site	0.125	Baek,2008
HAP3	*ALS3*	TF binding site	0.125	Baek,2008
HAP3	*FRE10*	TF binding site	0.125	Baek,2008
SOD5	Target	Co-expression	0.5	Frohner, 2009 Martchenko,2004
SOD5	*DDR48*	Co-expression	0.25	Oberholzer, 2006 Fradin, 2005 Doedt,2004 Thewes,2007
SOD5	*FRE10*	Co-expression	0.25	Sellam,2009
SOD5	*ALS3*	Co-expression	0.25	Zakikhany, 2007 Nobile,2008
DDR48	*ALS3*	Co-expression	0.25	Lane, 2001
FRE10	*ALS3*	Co-expression	0.25	Lane, 2001
ALS3	Target	Co-expression	0.5	Almeida,2008 Wächtler,2011 Zhao,2004
ZFP36	*CITED2*	Co-expression	0.25	Hou,2010
ZFP36	*LIF*	Co-expression	0.25	Hou,2010 Raman,2009 Smith,2010
ZFP36	*RGS1*	Co-expression	0.25	Smith,2010 Griffin,2009 Hinze, 2010 Spira,2007 Wang,2006

FIGURE 1 | RNA-seq predicts regulatory host-fungus interactions. (A) *M. musculus* clustering of differentially expressed genes (scaled and centered $\log_2$ values) over the time course of the infection. Black lines represent the expression of individual genes. The mean (dotted line, red) and the variance (solid, red) are shown. **(B)** *C. albicans* clustering of differentially expressed genes (scaled and centered $\log_2$ values) over the time course of the infection. Black lines represent the expression of individual genes. The mean (dotted line, red) and the variance (solid, red) are shown. **(C)** Prior knowledge incorporated into the inference network for all genes, where the input corresponds to the external perturbation (co-culture) and regulation by target corresponds to induction by the co-culture environment but not associated with a specific gene set. **(D)** Measured (dots), interpolated (dotted line), simulated (solid line) gene expression for all genes used in the inference model from Net*Gener*ator over the time course of the infection. **(E)** Inferred network model between *C. albicans* and *M. musculus*, where all *C. albicans* (blue) and *M. musculus* genes (green) in the study are included. The following interactions are represented on the model: predicted interactions based on the RNA-seq data set from individual species where no prior knowledge exists (gray), predicted interactions between a *C. albicans* and *M. musculus* gene where no prior knowledge exists (red), or where prior knowledge exists and corresponds to expression data set (orange). Here, activation is shown as an arrow and a repression with a bar. The "*C. albicans*/*M. musculus*" rectangle represents the influence from the external perturbation (co-culture during *in vitro* infection) on the gene expression level.

labeled with goat-anti-rabbit 649 Dylight (Thermo Scientific, Rockford, Illinois). Fungal cell wall chitin was labeled using 10 μM of Calcofluor White (Sigma-Aldrich). Intracellular labeling of Ptx3 was performed using the BD Cytofix/Cytoperm Fixation/Permeabilization kit with BD GolgiPlug protein transport inhibitor (BD Biosciences, Heidelberg) after 6 h of *C. albicans* infection following manufactures instructions. Preparations were assessed by flow cytometry or visualized on an Olympus Cell-R live imaging unit (Olympus, Essex, UK) for all experiments.

STATISTICAL ANALYSIS FOR INFERENCE MODEL VERIFICATION

Statistical analysis of data was performed using the GraphPad Prism graphing and analysis software (GraphPad Software, San Diego, USA) for all *in vitro* experiments excluding the RNA-seq

analysis described above. Statistical significance was assessed using with the Student *t*-test and a *p*-value <0.05 was considered significant.

RESULTS

INFERRED REGULATORY NETWORK IDENTIFIES NOVEL INTERSPECIES HOST–PATHOGEN INTERACTIONS

We used massively parallel RNA sequencing of cDNA (RNA-seq) obtained from co-cultures of *C. albicans* and *M. musculus* BMDCs over 2 h to model an infection time course from fungal adhesion to early host cell lysis. In total, we obtained approximately 120 million reads, which were mapped to the *C. albicans* 21 assembly or *M. musculus* mm9 genome and analyzed each for differential gene expression relative to the pre-infection state. We identified 545 differentially expressed genes for *C. albicans* and 240 for *M. musculus* over the complete time course.

The small number of measured data points for each gene over the time course restricts the modeling approach to a limited number of genes. If there was no pre-selection of the genes, or a large number of genes were to be used, it would result in an over-fitting of the measured data that would not produce a robust inference model. For this reason, it is necessary to select a set of relevant genes to be represented by nodes in the network model. To identify candidate genes in *C. albicans* and *M. musculus*, all differentially expressed genes were first clustered (Bezdek, 1992) by their kinetics during the time course (**Figures 1A,B**). From each cluster, one or more representative genes were chosen for use within the model. Several considerations were taken into account for the selection of candidate genes. In *C. albicans*, we preferentially chose genes that have been either annotated as virulence genes (i.e., adhesion, hyphal formation, or response to host) or strongly respond to infection or infection-like conditions (i.e., temperature stress, nutrient limitation, or iron regulation). For *M. musculus*, we prioritized genes with phenotypes relating to the immune defense or response, or susceptibility to pathogens in a systemic mouse model of infection.

A number of recent studies have shown the reverse engineering approach is greatly improved by the integration of different data sources (Werhli and Husmeier, 2007; Gustafsson et al., 2008; Hecker et al., 2009a,b). We therefore collected putative regulatory interactions based on additional data obtained from literature, referred to as prior knowledge, for each gene. Based on the confidence of the prior knowledge source, a score is attributed to each interaction (see Materials and Methods). Since different data sources might be contradictory, prior knowledge was softly integrated so that if a proposed interaction contradicts the measured data to a great extent, it can be removed from the resulting network (see Materials and Methods). Genes with no known or predicted function were therefore excluded from the analysis. Based on these criteria, we narrowed our gene lists to five from *C. albicans* and six from *M. musculus*. Prior knowledge scores (**Figure 1C**) and expression kinetics (**Figure 1D**) for the candidate genes were combined in Net*Gene*rator to generate the final network inference (**Figure 1E**).

To verify the fit of the model to the actual expression kinetics of the candidate genes, we first used Net*Gene*rator to interpolate and simulate gene expression for the measured data points of each gene (**Figure 1D**). The closer expression profiles for the individual genes fit to the measured data points, the better the inference prediction is. We found a close relationship between the simulated and measured data points, showing that the Net*Gene*rator model is representative of the measured data. The final interspecies network was based on these predictions (**Figure 1E**). Only edges that were robust against the addition of Gaussian noise and partial skipping of prior knowledge were used in the construction of the model. The final network predicts 21 putative edges, including 4 interspecies edges.

To specifically test the robustness of the interspecies edges experimentally, we focused on a sub-network composed of a single *C. albicans* transcription factor Hap3 that was predicted in the inference model to contain two interactions with *M. musculus* genes. These interactions include a predicted blunt or repressing edge between fungal Hap3 by murine Ptx3, and a predicted blunt edge of murine Mta2 by fungal Hap3 itself.

THE BINDING OF MURINE Ptx3 REGULATES Hap3 TARGET GENES

Pentraxin 3 (Ptx3) is a soluble pattern recognition receptor that has been previously shown to function as an opsonin to facilitate pathogen uptake by phagocytic cells in a dectin-1 dependent manner (Diniz et al., 2004). To determine if *M. musculus* Ptx3 blocked fungal Hap3 function or the expression of Hap3-regulated genes in *C. albicans* as suggested by the inference model, we first asked whether Ptx3 was induced upon infection with *Candida* cells. Using intracellular staining for Ptx3, we detected a strong fluorescence signal in BMDMs infected with *C. albicans*, whereas no signal was detected in BMDMs alone or *ptx3*$^{-/-}$ macrophages (**Figures 2A,B**). Furthermore, we found an increase in *PTX3* mRNA levels in BMDCs (**Figure 2C**) and BMDMs (data not shown) after 1 h of *C. albicans* infection, verifying that Ptx3 is indeed induced in our experimental setup. Interestingly, the amount of *PTX3* induced significantly decreased in the absence of Hap3 (**Figure 2C**). We detected a similarly significant decrease in *PTX3* induction in the gene containing the Hap3 binding box, Cda2, a predicted chitin deacetylase in *C. albicans*.

Ptx3 has been previously shown to bind to numerous fungi, including *Aspergillus fumigatus* (Jaillon et al., 2007) as well as zymosan-coated particles (Diniz et al., 2004). Therefore, we asked whether recombinant mouse Ptx3 (rmPtx3) could also bind to the *C. albicans* cell wall. We assessed rmPtx3 binding using fluorescence microscopy and flow cytometry. Fungal cells pre-incubated with rmPtx3 for 1 h at 37°C showed surface localization of Ptx3 to the cell wall. No signal was visible on cells treated with the PE-conjugated secondary alone or the untreated control (**Figure 2E**). Notably, the labeling pattern of Ptx3 coincided with areas of expected chitin exposure. Hence, we also stained *C. albicans* cells with Calcofluor White, a fluorescent dye that binds to exposed chitin (Bulawa et al., 1995). Interestingly, the localization of rmPtx3 on the *C. albicans* overlapped with the signal detected for Calcofluor White alone (**Figure 2E**), suggesting that Ptx3 might bind to accessible cell surface chitin. We confirmed and quantified the amount of Ptx3 binding to the cell wall using flow cytometry. Based on the binding observed in the fluorescence microscopy, we analyzed our data using both the complete pool of *C. albicans* cells as well as discriminating yeast and hyphal forms (**Figure 2D**; **Figure A1** in Appendix). We detected a high level of

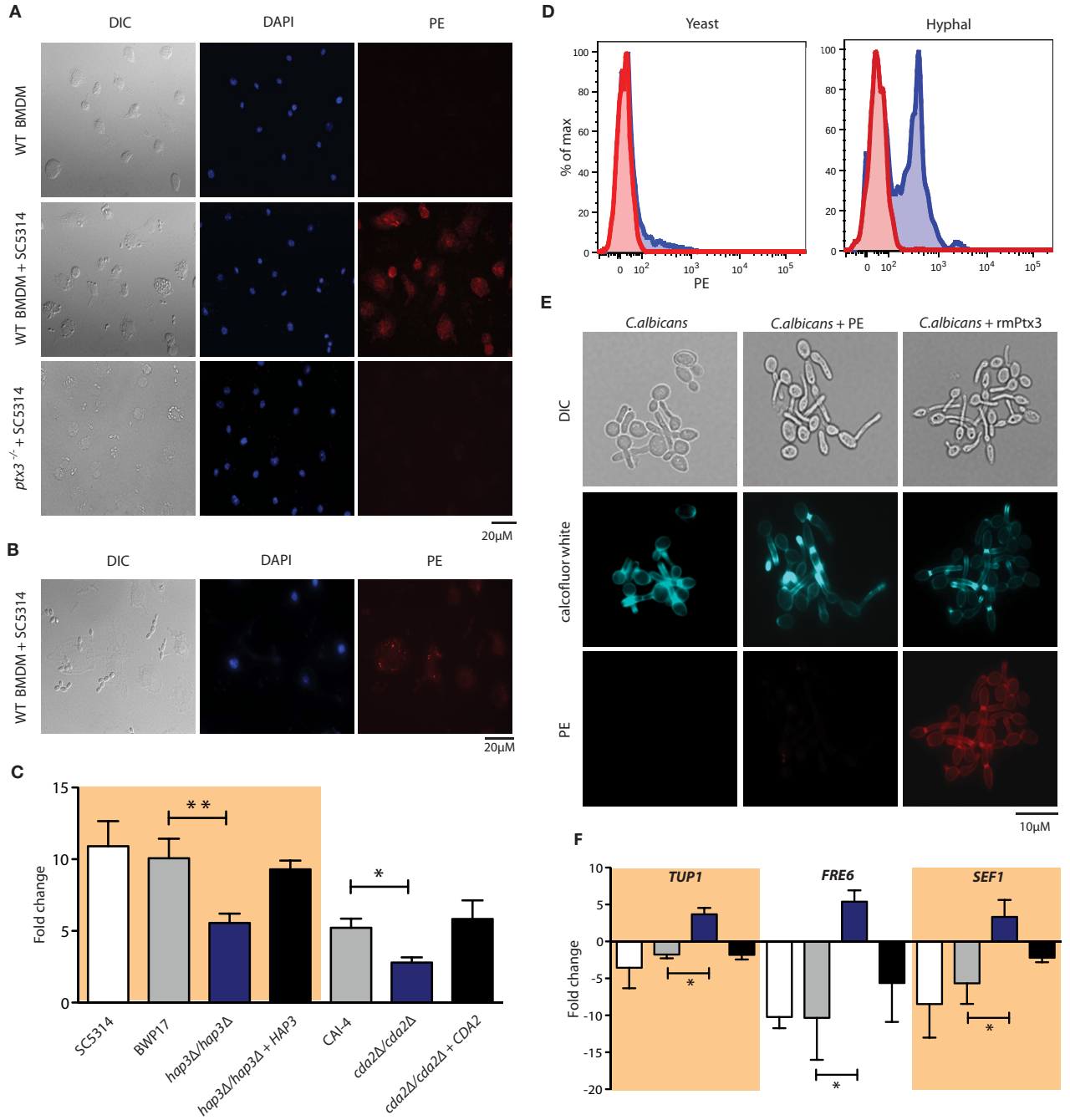


FIGURE 2 | Binding of rmPtx3 to *C. albicans* mediates the expression of Hap3 target genes. (A) Intracellular labeling of endogenous Ptx3 induction after 6 h of *C. albicans* stimulation of macrophages derived from wild type or *ptx3*$^{-/-}$ bone marrow. **(B)** Intracellular labeling of endogenous Ptx3 after 6 h of *C. albicans* stimulation of macrophages derived from wild type bone marrow. Macrophages directly associated with fungal cells and show a strong signal for endogenous Ptx3, while those not associated have only background signal levels. **(C)** qPCR of *PTX3* in BMDCs after 1 h of infection with different *C. albicans* strains. Results represent the mean of 3 pooled experiments ± SD. **(D)** FACS analysis of wild type strain SC5314 after 1-h treatment with rmPtx3, where untreated cells stained with PE only (red) and rmPtx3 and SC5314 (blue) are shown. Cells were gated by size to differentiate yeast and hyphal morphologies. **(E)** Fluorescence microscopy of SC5314 after 1 h pre-treatment with 5 μg rmPtx3 (red) or 10 μM Calcofluor White (blue). **(F)** qPCR of predicted targets genes of Hap3. SC5314 (white), BWP17 (gray), *hap3*Δ/*hap3*Δ (blue), and *hap3*Δ/*hap3*Δ + *HAP3* (black) after 1 h pre-incubation with 5 μg rmPtx3 are shown. Results represent the mean of 3 pooled experiments ± SD.

binding in fungal hyphae compared to yeast form cells. This is consistent with our fluorescence microscopy analysis, where we detected much stronger signals on the hyphal cell walls compared to the bud scars on yeast form cells.

Given that rmPtx3 binds to the fungal cell surface, we assessed if rmPtx3 binding influenced the expression of predicted fungal Hap3 target genes as predicted by the inferred network using qPCR (**Figure 2F**). There are 10 predicted target genes of Hap3 that were recently identified in a network inference study using microarray data from *C. albicans* during *in vitro* epithelial infection, where iron is assumed to be limited (Linde et al., 2010). Out of the 10 putative Hap3 target genes, we found three, *TUP1*, *FRE6*, and *SEF1*, whose expressions were significantly decreased in *C. albicans* after rmPtx3 binding, verifying their functionality as Hap3 target genes. Interestingly, the levels of these genes increased in the Hap3 knock-out. These data strongly suggest that their down-modulation upon binding of rmPtx3 is Hap3-dependent.

CANDIDA ALBICANS BOUND BY Ptx3 ATTENUATES THE IMMUNE RESPONSE IN BMDCs

Recently it was shown that the binding of recombinant human Ptx3 increases *A. fumigatus* conidia phagocytosis and influences cytokine production. Those mice lacking Ptx3 were additionally found to be more susceptible to *A. fumigatus* infection (Moalli et al., 2011). To determine if the binding of Ptx3 to *C. albicans* changed the cytokine production of murine immune cells *in vitro*, we first investigated the gross immune response using a cytokine array after 24 h of co-culture with BMDCs (**Figure A2** in Appendix). RmPtx3-bound *C. albicans* induced multiple cytokines compared to untreated *C. albicans* including IL-2, a cytokine regulated by Mta2, as well as the inflammatory cytokines KC, JE, and TNFα (**Figures A2** in Appendix). Interestingly, when we then compared Hap3 knock-out cells pre-incubated with rmPtx3, we detected a general increase in these cytokines in addition to IL-23, IL-17, IL-16, and IL-10 that were not detected in using wild type *C. albicans* (**Figure A2** in Appendix).

Since the cytokine array is a qualitative assessment of cytokine production with a relatively high detection threshold, we verified the changes in cytokine levels for the cytokines most relevant to the inference model, namely, IL-2 and IL-4, both target cytokines of Mta2, and TNFα, a pleiotropic inflammatory cytokine, by ELISA. Mta2 is a member of the NuRD (nucleosome remodeling and histone deacetylase) complex in *M. musculus* (Manavathi et al., 2007) and predicted in our network model as repressed by Hap3. The cytokines IL-2 and IL-4 produced during the host immune response were both recently identified as targets of the Mta2/NuRD complex (Lu et al., 2008). We found that in BMDCs, *MTA2* increased in the absence of Hap3, suggesting that Hap3 might indirectly regulate expression of Mta2 (**Figure 3A**). We quantified cellular cytokine release using wild type BMDCs with *C. albicans*, wild type BMDCs with rmPtx3-bound *C. albicans*, and $ptx3^{-/-}$ BMDCs. We found that the production of IL-2 and IL-4 significantly increased in the absence of Hap3 (**Figures 3B,C**). This increase was augmented by *C. albicans* cells pre-incubated with rmPtx3, confirming our observation from the cytokine array that there is a general increase in the production of these cytokines when there is an increase in Ptx3. Interestingly, compared to wild type BMDCs, the basal level of cytokine production of IL-2 and IL-4 increased in $ptx3^{-/-}$ BMDCs, corresponding to our inferred network prediction that the loss of its predicted negative regulators, Ptx3 and fungal Hap3, would increase the expression of Mta2 and thereby increase the expression of its target cytokines. In Hap3 knock-out cells, we found both on the cytokine arrays and by ELISA a significant decrease in TNFα (**Figure 3D**). These data show that the binding of Ptx3 to fungal cells alters the cytokine production by immune cells in a Hap3-dependent manner, and the regulation of Mta2 target cytokines indirectly suggests an involvement of Mta2 as predicted by the network inference model.

IDENTIFYING CELL SURFACE Hap3 TARGET GENES

To identify how immune cells could detect the regulation of the transcription factor Hap3 in *C. albicans*, we searched for putative Hap3 target genes that could have more direct contact with immune cells, including: cell wall, plasma membrane or secretory proteins. We focused on *C. albicans* genes of cluster 2, since their expression strongly increased expression over the time course of invasion (**Figure 1B**). Within this cluster, we scanned for genes harboring the binding site of the Hap-complex in their upstream regulatory regions (Baek et al., 2008). We further narrowed down the candidate list by removing genes that did not have a predicted cellular localization or function in the *C. albicans* database (Skrzypek et al., 2010). Following these selection criteria, nine candidate genes were left that we used to infer an additional network in combination with Ptx3, Hap3, and Mta2 to determine if an interaction could be inferred with a protein that could come in direct contact with immune cells (**Figure A3** in Appendix). To increase the reliability of the putative Hap3 interactions within the new interaction network, we included the validated interactions from our experiments within this study (repression of *HAP3* by Ptx3 and *MTA2* by Hap3), as additional prior knowledge. Of all of the candidate genes, only the activation of *CDA2* (a putative chitin deacetylase in *C. albicans*) by Hap3, was robust against Gaussian noise and partial skipping of prior knowledge.

DISCUSSION

In this study, we aimed to infer a network that predicts interactions between host and pathogenic species under infection settings. To our knowledge, this is the first network inference approach predicting host–pathogen interactions. This approach allowed for the prediction, identification, and experimental verification of interdependent sub-networks composed of a single *C. albicans* transcription factor Hap3, and the *M. musculus* genes Ptx3 and Mta2. The experimental validation suggests a putative mechanism to explain how these interactions could be regulated during infection of immune cells by fungal pathogens.

Our modeling approach was fundamentally based on differential equations, which have been previously used to infer regulatory network models (Toepfer et al., 2007; Linde et al., 2010). This approach is generally suitable for time series data. Nevertheless, this approach is inappropriate for large-scale modeling, because a large number of genes incorporated into a differential equation based model leaves open a number of parameters to be identified. This may result in an over-fitting of the data. The modeling approach uses four attempts to prevent over-fitting. First, we restrict the number of genes within the model such that a smaller number of parameters need to be identified. Second, it aims at inferring a sparse network where many parameters are zero. Thirdly, it makes use of re-sampling techniques where the data are

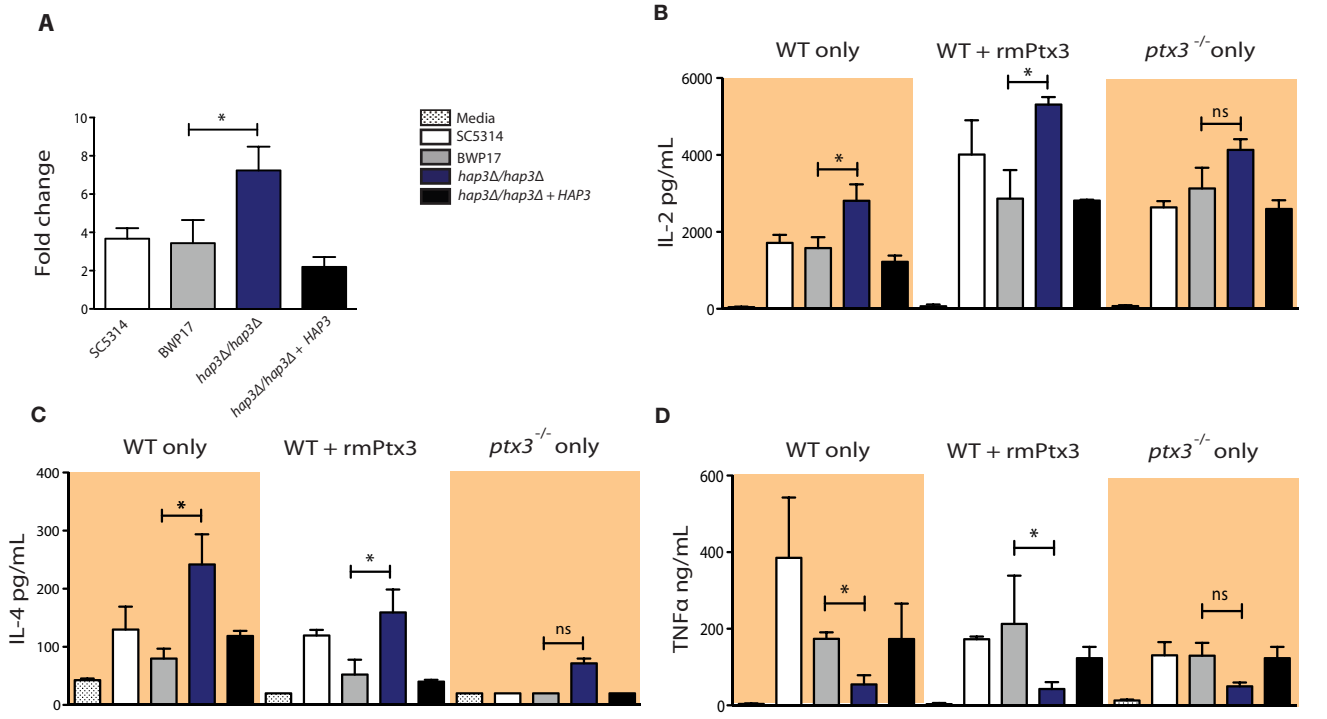


FIGURE 3 | Mouse *MTA2*, IL-2 and IL-4 increase upon stimulation with *C. albicans* in a Hap3-dependent manner. (A) qPCR of *M. musculus MTA2* expression in BMDCs after 1 h incubation with fungal cells at an MOI of five fungal cells to immune cells are shown. Results represent the mean of 3 pooled experiments ± SD. **(B–D)** ELISA measurement of IL-2 **(B)**, IL-4 **(C)**, and TNFα **(D)** from the supernatants of BMDCs after 24 h incubation with fungal cells in WT BMDCs, *C. albicans* cells pre-incubated with rmPtx3 on wild type BMDCs, or fungal cells on *ptx3*$^{-/-}$ BMDCs. Incubations were performed at an MOI of five fungal cells to immune cells. Results represent the mean of 2 pooled experiments ± SD.

perturbed in a random manner. Finally, we make use of prior knowledge guiding the inferred structure to a knowledge-based solution. Thus skipping incorrect network structures.

Gene expression levels, as well as available prior biological knowledge, were used to aid in the narrowing of genes that we chose to incorporate into the model. For this reason, genes where no biological knowledge was available were excluded from further analysis. However, we cannot exclude the possibility that additional genes of unknown function might also play a role in our inference model. This remains a limitation of the modeling approach, in so far as predictions can only be made for genes where a reasonable amount or prior knowledge is available. The genes incorporated into the model represent only one possible scenario of interactions and we do not exclude the possibility that other genes may play a role under other conditions. We have already started to take first step for a full-genomic network modeling for *C. albicans* utilizing a compendium of all available expression data (Altwasser et al., 2012). Moreover, we primarily focused on genes acting as putative network "hubs" in their organisms (Bulawa et al., 1995). Hubs are genes such as transcription factors that regulate many other downstream genes within a network either directly or indirectly. Hubs were chosen because they are less likely to have redundant roles. Therefore, we would expect a stronger phenotype than investigating genes that are sparsely connected. This also means that the interactions we are investigating are more likely to be indirect and should be interpreted with caution.

From our original candidate gene list, we inferred *HAP3* as a putative network hub targeted by innate immune cells. Interestingly, several putative target genes of Hap3 identified in this study are predicted to localize to the plasma membrane, cell wall, or are involved with cell wall reorganization in *C. albicans*. The fungal cell wall is a dynamic structure, which undergoes significant structural and molecular composition remodeling throughout its life cycle, as well as in response to a variety of external stimuli (Chaffin, 2008). As Hap3 in *C. albicans* is a transcription factor up-regulated under iron-limiting conditions (Linde et al., 2010), it is likely that its function during fungal recognition or phagocytosis by immune cells is indirect. Of all of the candidate cell surface Hap3 targets, only Cda2, a putative chitin deacetylase forms a robust interaction with Hap3 within the second network (**Figure A3** in Appendix). Chitin deacetylase enzymes exists in both intracellular and secreted forms in different fungi, where they hydrolyzes the acetamido group in the *N*-acetylglucosamine units of chitin and chitosan, leaving glucosamine units and acetic acid form as byproducts (Zhao et al., 2010). Chitin deacetylases exist in both *Saccharomyces cerevisiae* (Martinou et al., 2002) and in the opportunistic fungal pathogen *Cryptococcus neoformans*, where they have been suggested as an antifungal target due to their severe effect on cell wall integrity (Baker et al., 2007). Notably, chitin deacetylases are secreted during different developmental stages of some other fungi (Zhao et al., 2010). For example, in *Colletotrichum lindemuthianum*, a plant fungal pathogen, chitin deacetylases are

exclusively secreted during hyphal penetration into plant tissue (Tokuyasu et al., 1996). We find that Ptx3 induction is decreased in the *CDA2* knock-out (**Figure 2C**), further suggesting a possible connection to the inferred network model. These data are consistent with the overlap of Ptx3 staining and that of Calcofluor White (**Figure 2E**), which binds to exposed chitin. Therefore, it is tempting to speculate that the recognition of *C. albicans* by immune cells triggers the production of this enzyme to induce cell wall remodeling as an evasion strategy. However, further work beyond the scope of this study is needed to decipher the specific function of Cda2 in *C. albicans* and its connection to Hap3.

We observed that upon binding of rmPtx3 to fungal cells, the *C. albicans* virulence genes *TUP1*, *FRE6*, and *SEF1* mRNA levels significantly decreased in a Hap3-dependent manner (**Figure 2F**). Tup1 has a well-characterized role as a key regulator in *C. albicans* morphogenesis (Braun and Johnson, 1997). We cannot exclude the possibility that Hap3 and Tup1 may have similar or even complementary functions during interaction with immune cells. Interestingly, both Tup1 and Fre6 are either directly or indirectly involved in the *C. albicans* cell wall homeostasis. Tup1 is a multifunctional transcriptional co-repressor of filamentous growth in *C. albicans* whose lack leads to constitutive filamentous growth (Braun and Johnson, 1997; Park and Morschhäuser, 2005). Fre6 is an uncharacterized protein, for which *in silico* predictions suggest it to reside in the plasma membrane with a putative functional similarity to the ferric reductase Fre10, an important protein in iron acquisition (Knight et al., 2005). Therefore, their regulation upon binding or phagocytosis might play an additional role in cell wall remodeling during infection. Since fungal cells experience severe iron-limiting condition within phagosomes of host cells, Hap3 and Fre6 appear as logical candidates involved in this reciprocal interaction. Likewise, Sef1 regulates iron uptake, and has recently been shown to promote virulence in a mouse model of bloodstream infections (Chen et al., 2011). Interestingly, it was shown that knock-out mice lacking Ptx3 are hyper-susceptible to *A. fumigatus* (Moalli et al., 2010). However, no *in vivo* work has been performed to date using *ptx3*$^{-/-}$ mice and *C. albicans*. A recent study has shown that the activation of the complement system via the lectin pathway can be triggered via a complex of Ptx3 and mannose binding lectin (MBL) on *C. albicans* mannan *in vitro* (Ma et al., 2011). They showed the MBL–Ptx3 complex could enhance the deposition of the complement components C3 and C4 and thereby increase phagocytosis of *C. albicans* by polymorphonuclear leukocytes. It has previously been shown that C3 knock-out mice are additionally more susceptible to *C. albicans* infections (Han et al., 2001). Therefore, it is possible that the absence of Ptx3 could result in reduced activation of the complement pathway and reduced fungal killing. *In vivo* studies using *ptx3*$^{-/-}$ mice would be needed to investigate this hypothesis.

We found that the expression of *MTA2* and the regulation of its downstream targets such as cytokines IL-2 and IL-4, are increased during immune cell invasion by *C. albicans* in a Hap3-dependent manner (**Figures 3B,C**). Moreover, we found that an altered-immune response is one consequence of rmPtx3 binding. Mta2 knock-out mice display partial embryonic lethality, while the surviving mice develop lupus-like autoimmune symptoms, including severe developmental phenotypes (Lu et al., 2008). Mta2 is uniquely associated with the NuRD chromatin complex, which has both nucleosome remodeling and histone deacetylase activity (Feng and Zhang, 2003). Although there has been no data

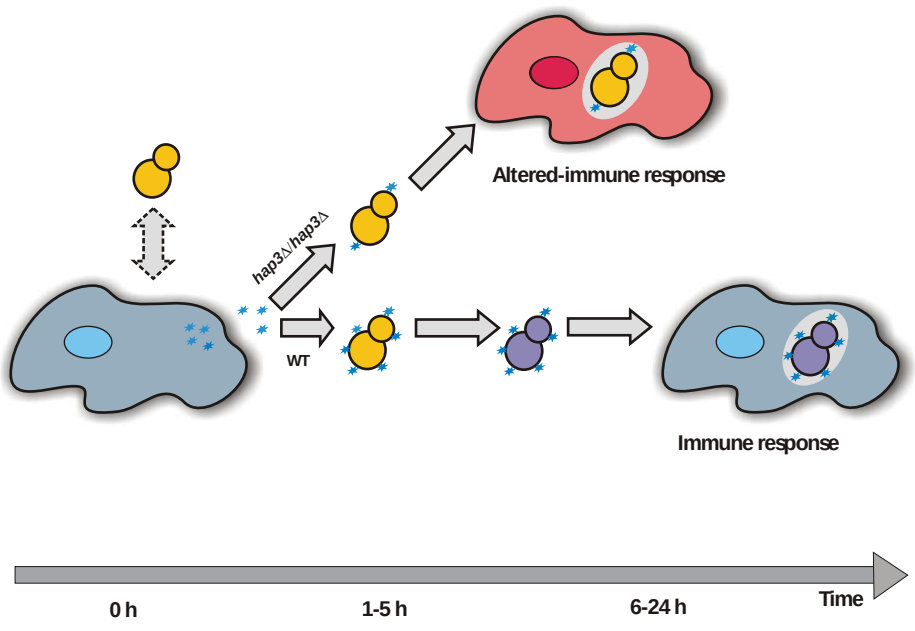


FIGURE 4 | A proposed model for the mode of action of Ptx3 on *C. albicans*. After mutual recognition between *C. albicans* (yellow) and host immune cells (blue), Ptx3 (blue, stars) is released into the surrounding milieu where it can bind to the invading fungal cell wall. The binding of Ptx3 induces a change in the *C. albicans* cell wall (purple) after the activation of Hap3 target genes, influencing its recognition by immune cells and the subsequent immune response. Arrows represent progression of time during an infection.

to date in fungi indicating a role for host chromatin in pathogenesis, recent work in bacteria and viruses (Hamon and Cossart, 2008; Rohde, 2011) shows that chromatin remodeling is induced in host cells during invasion. Consistent with these observations, our data suggests that the regulation of *MTA2* may affect chromatin remodeling in immune cells in the response to fungal pathogens. The resultant altered-immune response may be disadvantageous to the pathogen because it would promote fungal clearance.

We propose that Hap3 constitutes a target hub of *C. albicans*, which actively regulates immune responses through the reorganization of the *C. albicans* cell wall during invasion of innate immune cells (**Figure 4**). Specifically, we propose a model in which the binding of Ptx3 released from immune cells to *C. albicans* cell wall triggering the reorganization of the *C. albicans* cell wall and plasma membrane via the activation of Hap3 target genes. This reorganization in turn changes the recognition of the fungus by immune cells and attenuates the host immune response. This work demonstrates the possibility to experimentally verify predicted host–pathogen relationships based on an interspecies model of network inference, showing that inference modeling can be used in the investigation of microbial pathogenesis. We propose that this method could be useful for the identification of antifungal target genes.

CONTRIBUTION

Lanay Tierney designed research, performed experiments, analyzed data and co-wrote the manuscript. Jörg Linde generated network inference maps, analyzed data, and co-wrote the manuscript. Sebastian Müller performed the sequence alignment and normalization. Sascha Brunke, Juan Camilo Molina, and Bernhard Hube provided *C. albicans* Hap3 deletion strains. Reinhard Guthke and Karl Kuchler designed the research and co-wrote the manuscript.

ACKNOWLEDGMENTS

We would like to thank Alberto Mantovani and Cecilia Garlanda (Milan, Italy) for kindly providing *ptx3*$^{-/-}$ bone marrow. We are also indebted to Neil Gow (Aberdeen, UK) for kindly providing the Cda2 knock-out *C. albicans* strains. This work was supported by a grant from the Christian Doppler Research Society to Karl Kuchler, and in part by the FWF-DACH grant of the Austrian Science Foundation (FWF-Proj.: I-746-B11) to Karl Kuchler and Bernhard Hube. Jörg Linde was supported by the excellence graduate school "Jena School for Microbial Communication (JSMC)." We would like to thank laboratory members for critical reading and helpful comments on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Microbial_Immunology/10.3389/fmicb.2012.00085/abstract

Table S1 | RPKM values for *C. albicans* genes over the infection time course.

Table S2 | RPKM values for *M. musculus* genes over the infection time course.

REFERENCES

Almeida, R. S., Brunke, S., Albrecht, A., Thewes, S., Laue, M., Edwards, J. E., Filler, S. G., and Hube, B. (2008). The hyphal-associated adhesin and invasin Als3 of *Candida albicans* mediates iron acquisition from host ferritin. *PLoS Pathog.* 4, e1000217. doi:10.1371/journal.ppat.1000217

Altwasser, R., Linde, J., Buyko, E., Hahn, U., and Guthke, R. (2012). Genome-wide scale-free network inference for *Candida albicans*. *Front. Microbiol.* 3:51. doi:10.3389/fmicb.2012.00051

Aranda, B., Achuthan, P., Alam-Faruque, Y., Armean, I., Bridge, A., Derow, C., Feuermann, M., Ghanbarian, A. T., Kerrien, S., Khadake, J., Leroy, C., Menden, M., Michaut, M., Montecchi-Palazzi, L., Neuhauser, S. N., Orchard, S., Perreau, V., Roechert, B., van Eijk, K., and Hermjakob, H. (2010). The IntAct molecular interaction database in 2010. *Nucleic Acids Res.* 38, D525–D531.

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., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* 25, 25–29.

Baek, Y. U., Li, M., and Davis, D. A. (2008). *Candida albicans* ferric reductases are differentially regulated in response to distinct forms of iron limitation by the Rim101 and CBF transcription factors. *Eukaryot. Cell* 7, 1168–1179.

Baker, L. G., Specht, C. A., Donlin, M. J., and Lodge, J. K. (2007). Chitosan, the deacetylated form of chitin, is necessary for cell wall integrity in *Cryptococcus neoformans*. *Eukaryot. Cell* 6, 855–867.

Barabasi, A. L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113.

Bezdek, J. (1992). *Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data*. New York: Institute of Electrical and Electronics Engineers (IEEE) Press.

Blake, J. A., Bult, C. J., Kadin, J. A., Richardson, J. E., and Eppig, J. T. (2011). The mouse genome database (MGD): premier model organism resource for mammalian genomics and genetics. *Nucleic Acids Res.* 39, D842–D848.

Bourgeois, C., Majer, O., Frohner, I., and Kuchler, K. (2009). *In vitro* systems for studying the interaction of fungal pathogens with primary cells from the mammalian innate immune system. *Methods Mol. Biol.* 470, 125–139.

Bourgeois, C., Majer, O., Frohner, I. E., Tierney, L., and Kuchler, K. (2010). Fungal attacks on mammalian hosts: pathogen elimination requires sensing and tasting. *Curr. Opin. Microbiol.* 13, 401–408.

Braun, B. R., and Johnson, A. D. (1997). Control of filament formation in *Candida albicans* by the transcriptional repressor *TUP1*. *Science* 277, 105–109.

Bulawa, C. E., Miller, D. W., Henry, L. K., and Becker, J. M. (1995). Attenuated virulence of chitin-deficient mutants of *Candida albicans*. *Proc. Natl. Acad. Sci. U.S.A.* 92, 10570–10574.

Butte, A. J., and Kohane, I. S. (2000). Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. *Pac. Symp. Biocomput.* 5, 418–429.

Chaffin, W. L. (2008). *Candida albicans* cell wall proteins. *Microbiol. Mol. Biol. Rev.* 72, 495–544.

Chen, C., Pande, K., French, S. D., Tuch, B. B., and Noble, S. M. (2011). An iron homeostasis regulatory circuit with reciprocal roles in *Candida albicans* commensalism and pathogenesis. *Cell Host Microbe* 10, 118–135.

D'Haeseleer, P., Wen, X., Fuhrman, S., and Somogyi, R. (1999). Linear modelling of mRNA expression levels during CNS development and injury. *Pac. Symp. Biocomput.* 4, 41–52.

Diniz, S. N., Nomizo, R., Cisalpino, P. S., Teixeira, M. M., Brown, G. D., Mantovani, A., Gordon, S., Reis, L. F., and Dias, A. A. (2004). *PTX3* function as an opsonin for the dectin-1-dependent internalization of zymosan by macrophages. *J. Leukoc. Biol.* 75, 649–656.

Doedt, T., Krishnamurthy, S., Bockmuhl, D. P., Tebarth, B., Stempel, C., Russell, C. L., Brown, A. J., and Ernst, J. F. (2004). APSES proteins regulate morphogenesis and metabolism in *Candida albicans*. *Mol. Biol. Cell* 15, 3167–3180.

Faith, J. J., Hayete, B., Thaden, J. T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J. J., and Gardner, T. S. (2007). Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* 5, e8. doi:10.1371/journal.pbio.0050008

Feng, Q., and Zhang, Y. (2003). The NuRD complex: linking histone modification to nucleosome remodeling. *Curr. Top. Microbiol. Immunol.* 274, 269–290.

Flicek, P., Amode, M. R., Barrell, D., Beal, K., Brent, S., Chen, Y., Clapham, P., Coates, G., Fairley, S., Fitzgerald, S., Gil, L., Gordon, L., Hendrix, M., Hourlier, T., Johnson, N., Kähäri, A. K., Keefe, D., Keenan, S., Kinsella, R., Komorowska, M., Koscielny, G., Kulesha, E., Larsson, P., Longden, I., McLaren, W., Muffato, M., Overduin, B., Pignatelli, M., Pritchard, B., Riat, H. S., Ritchie, G. R., Ruffier, M., Schuster, M., Sobral, D., Tang, Y. A., Taylor, K., Trevanion, S., Vandrovcova, J., White, S., Wilson, M., Wilder, S. P., Aken, B. L., Birney, E., Cunningham, F., Dunham, I., Durbin, R., Fernández-Suarez, X. M., Harrow, J., Herrero, J., Hubbard, T. J., Parker, A., Proctor, G., Spudich, G., Vogel, J., Yates, A., Zadissa, A., and Searle, S. M. (2011). Ensembl 2011. *Nucleic Acids Res.* 39, D800–D806.

Fradin, C., De Groot, P., MacCallum, D., Schaller, M., Klis, F., Odds, F. C., and Hube, B. (2005). Granulocytes govern the transcriptional response, morphology and proliferation of *Candida albicans* in human blood. *Mol. Microbiol.* 56, 397–415.

Friedman, S. R., Kottiri, B. J., Neaigus, A., Curtis, R., Vermund, S. H., and Des Jarlais, D. C. (2000). Network-related mechanisms may help explain long-term HIV-1 seroprevalence levels that remain high but do not approach population-group saturation. *Am. J. Epidemiol.* 152, 913–922.

Frohner, I. E., Bourgeois, C., Yatsyk, K., Majer, O., and Kuchler, K. (2009). *Candida albicans* cell surface superoxide dismutases degrade host-derived reactive oxygen species to escape innate immune surveillance. *Mol. Microbiol.* 71, 240–252.

Gillum, A. M., Tsay, E. Y., and Kirsch, D. R. (1984). Isolation of the *Candida albicans* gene for orotidine-5'-phosphate decarboxylase by complementation of *S. cerevisiae* ura3 and *E. coli* pyrF mutations. *Mol. Gen. Genet.* 198, 179–182.

Gola, S., Martin, R., Walther, A., Dunkler, A., and Wendland, J. (2003). New modules for PCR-based gene targeting in *Candida albicans*: rapid and efficient gene targeting using 100 bp of flanking homology region. *Yeast* 20, 1339–1347.

Griffin, T. A., Barnes, M. G., Ilowite, N. T., Olson, J. C., Sherry, D. D., Gottlieb, B. S., Aronow, B. J., Pavlidis, P., Hinze, C. H., Thornton, S., Thompson, S. D., Grom, A. A., Colbert, R. A., and Glass, D. N. (2009). Gene expression signatures in polyarticular juvenile idiopathic arthritis demonstrate disease heterogeneity and offer a molecular classification of disease subsets. *Arthritis Rheum.* 60, 2113–2123.

Gudlaugsson, O., Gillespie, S., Lee, K., Vande Berg, J., Hu, J., Messer, S., Herwaldt, L., Pfaller, M., and Diekema, D. (2003). Attributable mortality of nosocomial candidemia, revisited. *Clin. Infect. Dis.* 37, 1172–1177.

Gustafsson, M., Björkengren, J., and Tegne, J. (2008). "Soft integration of data for reverse engineering," in *International Conference on Systems Biology*, Göteborg, 127–127.

Guthke, R., Albrecht, D., Brackhage, A. A., and Möller, U. (2007). Discovery of gene regulatory networks in *Aspergillus fumigatus. Lect. Notes Bioinform.* 4366, 22–41.

Guthke, R., Moller, U., Hoffmann, M., Thies, F., and Topfer, S. (2005). Dynamic network reconstruction from gene expression data applied to immune response during bacterial infection. *Bioinformatics* 21, 1626–1634.

Hamon, M. A., and Cossart, P. (2008). Histone modifications and chromatin remodelling during bacterial infections. *Cell Host Microbe* 4, 100–109.

Han, Y., Kozel, T. R., Zhang, M. X., MacGill, R. S., Carroll, M. C., and Cutler, J. E. (2001). Complement is essential for protection by an IgM and an IgG3 monoclonal antibody against experimental, hematogenously disseminated candidiasis. *J. Immunol.* 167, 1550–1557.

Hardcastle, T. J., and Kelly, K. A. (2010). Bayseq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* 11, 422. doi:10.1186/1471-2105-11-422

Hecker, M., Goertsches, R. H., Engelmann, R., Thiesen, H. J., and Guthke, R. (2009a). Integrative modelling of transcriptional regulation in response to antirheumatic therapy. *BMC Bioinformatics* 10, 262. doi:10.1186/1471-2105-10-262

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

Hinze, C. H., Fall, N., Thornton, S., Mo, J. Q., Aronow, B. J., Layh-Schmitt, G., Griffin, T. A., Thompson, S. D., Colbert, R. A., Glass, D. N., Barnes, M. G., and Grom, A. A. (2010). Immature cell populations and an erythropoiesis gene-expression signature in systemic juvenile idiopathic arthritis: implications for pathogenesis. *Arthritis Res. Ther.* 12, R123.

Holter, N. S., Maritan, A., Cieplak, M., Fedoroff, N. V., and Banavar, J. R. (2001). Dynamic modelling of gene expression data. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1693–1698.

Hou, J., Aerts, J., den Hamer, B., van Ijcken, W., den Bakker, M., Riegman, P., van der Leest, C., van der Spek, P., Foekens, J. A., Hoogsteden, H. C., Grosveld, F., and Philipsen, S. (2010). Gene expression-based classification of non-small cell lung carcinomas and survival prediction. *PLoS ONE* 5, e10312. doi:10.1371/journal.pone.0010312

Jaillon, S., Peri, G., Delneste, Y., Fremaux, I., Doni, A., Moalli, F., Garlanda, C., Romani, L., Gascan, H., Bellocchio, S., Bozza, S., Cassatella, M. A., Jeannin, P., and Mantovani, A. (2007). The humoral pattern recognition receptor *PTX3* is stored in neutrophil granules and localizes in extracellular traps. *J. Exp. Med.* 204, 793–804.

Knight, S. A., Vilaire, G., Lesuisse, E., and Dancis, A. (2005). Iron acquisition from transferrin by *Candida albicans* depends on the reductive pathway. *Infect. Immun.* 73, 5482–5492.

Lane, S., Birse, C., Zhou, S., Matson, R., and Liu, H. (2001). DNA array studies demonstrate convergent regulation of virulence factors by Cph1, Cph2, and Efg1 in *Candida albicans*. *J. Biol. Chem.* 276, 48988–48996.

Le Novere, N., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E., Wegner, K., Aladjem, M. I., Wimalaratne, S. M., Bergman, F. T., Gauges, R., Ghazal, P., Kawaji, H., Li, L., Matsuoka, Y., Villéger, A., Boyd, S. E., Calzone, L., Courtot, M., Dogrusoz, U., Freeman, T. C., Funahashi, A., Ghosh, S., Jouraku, A., Kim, S., Kolpakov, F., Luna, A., Sahle, S., Schmidt, E., Watterson, S., Wu, G., Goryanin, I., Kell, D. B., Sander, C., Sauro, H., Snoep, J. L., Kohn, K., and Kitano, H. (2009). The systems biology graphical notation. *Nat. Biotechnol.* 27, 735–741.

Linde, J., Wilson, D., Hube, B., and Guthke, R. (2010). Regulatory network modelling of iron acquisition by a fungal pathogen in contact with epithelial cells. *BMC Syst. Biol.* 4, 148. doi:10.1186/1752-0509-4-148

Lu, X., Kovalev, G. I., Chang, H., Kallin, E., Knudsen, G., Xia, L., Mishra, N., Ruiz, P., Li, E., Su, L., and Zhang, Y. (2008). Inactivation of NuRD component Mta2 causes abnormal T cell activation and lupus-like autoimmune disease in mice. *J. Biol. Chem.* 283, 13825–13833.

Ma, Y. J., Doni, A., Skjoedt, M. O., Honore, C., Arendrup, M., Mantovani, A., and Garred, P. (2011). Heterocomplexes of mannose-binding lectin and the pentraxins *PTX3* or serum amyloid P component trigger cross-activation of the complement system. *J. Biol. Chem.* 286, 3405–3417.

Manavathi, B., Singh, K., and Kumar, R. (2007). MTA family of coregulators in nuclear receptor biology and pathology. *Nucl. Recept. Signal* 5, e010.

Marbach, D., Prill, R. J., Schaffter, T., Mattiussi, C., Floreano, D., and Stolovitzky, G. (2010). Revealing strengths and weaknesses of methods for gene network inference. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6286–6291.

Martchenko, M., Alarco, A. M., Harcus, D., and Whiteway, M. (2004). Superoxide dismutases in *Candida albicans*: transcriptional regulation and functional characterization of the hyphal-induced SOD5 gene. *Mol. Biol. Cell* 15, 456–467.

Martinou, A., Koutsioulis, D., and Bouriotis, V. (2002). Expression, purification, and characterization of a cobalt-activated chitin deacetylase (Cda2p) from *Saccharomyces cerevisiae. Protein Expr. Purif.* 24, 111–116.

Moalli, F., Doni, A., Deban, L., Zelante, T., Zagarella, S., Bottazzi, B., Romani, L., Mantovani, A., and Garlanda, C. (2010). Role of complement and Fcγ receptors in the protective activity of the long pentraxin *PTX3* against *Aspergillus fumigatus. Blood* 116, 5170–5180.

Moalli, F., Paroni, M., Veliz Rodriguez, T., Riva, F., Polentarutti, N., Bottazzi, B., Valentino, S., Mantero, S., Nebuloni, M., Mantovani, A., Bragonzi, A., and Garlanda, C. (2011). The therapeutic potential of the humoral pattern recognition molecule *PTX3* in chronic lung infection caused by *Pseudomonas aeruginosa. J. Immunol.* 186, 5425–5434.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628.

Murad, A. M., Lee, P. R., Broadbent, I. D., Barelle, C. J., and Brown, A. J. (2000). CIp10, an efficient and convenient integrating vector for *Candida albicans. Yeast* 16, 325–327.

Netea, M. G., Brown, G. D., Kullberg, B. J., and Gow, N. A. (2008). An integrated model of the recognition

of *Candida albicans* by the innate immune system. *Nat. Rev. Microbiol.* 6, 67–78.

Nobile, C. J., Solis, N., Myers, C. L., Fay, A. J., Deneault, J. S., Nantel, A., Mitchell, A. P., and Filler, S. G. (2008). *Candida albicans* transcription factor Rim101 mediates pathogenic interactions through cell wall functions. *Cell. Microbiol.* 10, 2180–2196.

Oberholzer, U., Nantel, A., Berman, J., and Whiteway, M. (2006). Transcript profiles of *Candida albicans* cortical actin patch mutants reflect their cellular defects: contribution of the Hog1p and Mkc1p signaling pathways. *Eukaryot. Cell* 5, 1252–1265.

Park, Y. N., and Morschhäuser, J. (2005). *Candida albicans* MTLα tup1Δ mutants can reversibly switch to mating-competent, filamentous growth forms. *Mol. Microbiol.* 58, 1288–1302.

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

Raman, T., O'Connor, T. P., Hackett, N. R., Wang, W., Harvey, B. G., Attiyeh, M. A., Dang, D. T., Teater, M., and Crystal, R. G. (2009). Quality control in microarray assessment of gene expression in human airway epithelium. *BMC Genomics* 10, 493. doi:10.1186/1471-2164-10-493

Rizzetto, L., and Cavalieri, D. (2011). Friend or foe: using systems biology to elucidate interactions between fungi and their hosts. *Trends Microbiol.* 19, 509–515.

Rohde, J. R. (2011). Microbiology. *Listeria* unwinds host DNA. *Science* 331, 1271–1272.

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

Sellam, A., Al-Niemi, T., McInnerney, K., Brumfield, S., Nantel, A., and Suci, P. A. (2009). A *Candida albicans* early stage biofilm detachment event in rich medium. *BMC Microbiol.* 9, 25. doi:10.1186/1471-2180-9-25

Skrzypek, M. S., Arnaud, M. B., Costanzo, M. C., Inglis, D. O., Shah, P., Binkley, G., Miyasato, S. R., and Sherlock, G. (2010). New tools at the *Candida* genome database: biochemical pathways and full-text literature search. *Nucleic Acids Res.* 38, D428–D432.

Smith, J. J., Deane, N. G., Wu, F., Merchant, N. B., Zhang, B., Jiang, A., Lu, P., Johnson, J. C., Schmidt, C., Bailey, C. E., Eschrich, S., Kis, C., Levy, S., Washington, M. K., Heslin, M. J., Coffey, R. J., Yeatman, T. J., Shyr, Y., and Beauchamp, R. D. (2010). Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* 138, 958–968.

Spira, A., Beane, J. E., Shah, V., Steiling, K., Liu, G., Schembri, F., Gilman, S., Dumas, Y. M., Calner, P., Sebastiani, P., Sridhar, S., Beamis, J., Lamb, C., Anderson, T., Gerry, N., Keane, J., Lenburg, M. E., and Brody, J. S. (2007). Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat. Med.* 13, 361–366.

Stark, C., Breitkreutz, B. J., Chatr-Aryamontri, A., Boucher, L., Oughtred, R., Livstone, M. S., Nixon, J., Van Auken, K., Wang, X., Shi, X., Reguly, T., Rust, J. M., Winter, A., Dolinski, K., and Tyers, M. (2011). The BioGRID interaction database: 2011 update. *Nucleic Acids Res.* 39, D698–D704.

Thewes, S., Kretschmar, M., Park, H., Schaller, M., Filler, S. G., and Hube, B. (2007). *In vivo* and *ex vivo* comparative transcriptional profiling of invasive and non-invasive *Candida albicans* isolates identifies genes associated with tissue invasion. *Mol. Microbiol.* 63, 1606–1628.

Toepfer, S., Guthke, R., Driesch, D., Woetzel, D., and Pfaff, M. (2007). The Net*Gene*rator algorithm: reconstruction of gene regulatory networks. *Lect. Notes Bioinform.* 4366, 119–130.

Tokuyasu, K., Ohnishi-Kameyama, M., and Hayashi, K. (1996). Purification and characterization of extracellular chin deacetylase from *Colletotrichum lindemuthianum*. *Biosci. Biotechnol. Biochem.* 60, 1598–1603.

Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111.

Wachtler, B., Wilson, D., Haedicke, K., Dalle, F., and Hube, B. (2011). From attachment to damage: defined genes of *Candida albicans* mediate adhesion, invasion and damage during interaction with oral epithelial cells. *PLoS ONE* 6, e17046. doi:10.1371/journal.pone.0017046

Wang, Q., Diskin, S., Rappaport, E., Attiyeh, E., Mosse, Y., Shue, D., Seiser, E., Jagannathan, J., Shusterman, S., Bansal, M., Khazi, D., Winter, C., Okawa, E., Grant, G., Cnaan, A., Zhao, H., Cheung, N. K., Gerald, W., London, W., Matthay, K. K., Brodeur, G. M., and Maris, J. M. (2006). Integrative genomics identifies distinct molecular classes of neuroblastoma and shows that multiple genes are targeted by regional alterations in DNA copy number. *Cancer Res.* 66, 6050–6062.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63.

Werhli, A. V., and Husmeier, D. (2007). Reconstructing gene regulatory networks with Bayesian networks by combining expression data with multiple sources of prior knowledge. *Stat. Appl. Genet. Mol. Biol.* 6, 15.

Zakikhany, K., Naglik, J. R., Schmidt-Westhausen, A., Holland, G., Schaller, M., and Hube, B. (2007). *In vivo* transcript profiling of *Candida albicans* identifies a gene essential for interepithelial dissemination. *Cell. Microbiol.* 9, 2938–2954.

Zhang, J., Chiodini, R., Badr, A., and Zhang, G. (2011). The impact of next-generation sequencing on genomics. *J. Genet. Genomics* 38, 95–109.

Zhao, X., Oh, S. H., Cheng, G., Green, C. B., Nuessen, J. A., Yeater, K., Leng, R. P., Brown, A. J., and Hoyer, L. L. (2004). ALS3 and ALS8 represent a single locus that encodes a *Candida albicans* adhesin; functional comparisons between Als3p and Als1p. *Microbiology* 150, 2415–2428.

Zhao, Y., Park, R. D., and Muzzarelli, R. A. (2010). Chitin deacetylases: properties and applications. *Mar. Drugs* 8, 24–46.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 December 2011; accepted: 20 February 2012; published online: 12 March 2012.

Citation: Tierney L, Linde J, Müller S, Brunke S, Molina JC, Hube B, Schöck U, Guthke R and Kuchler K (2012) An interspecies regulatory network inferred from simultaneous RNA-seq of Candida albicans invading innate immune cells. Front. Microbio. ***3****:85. doi: 10.3389/fmicb.2012.00085*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

APPENDIX

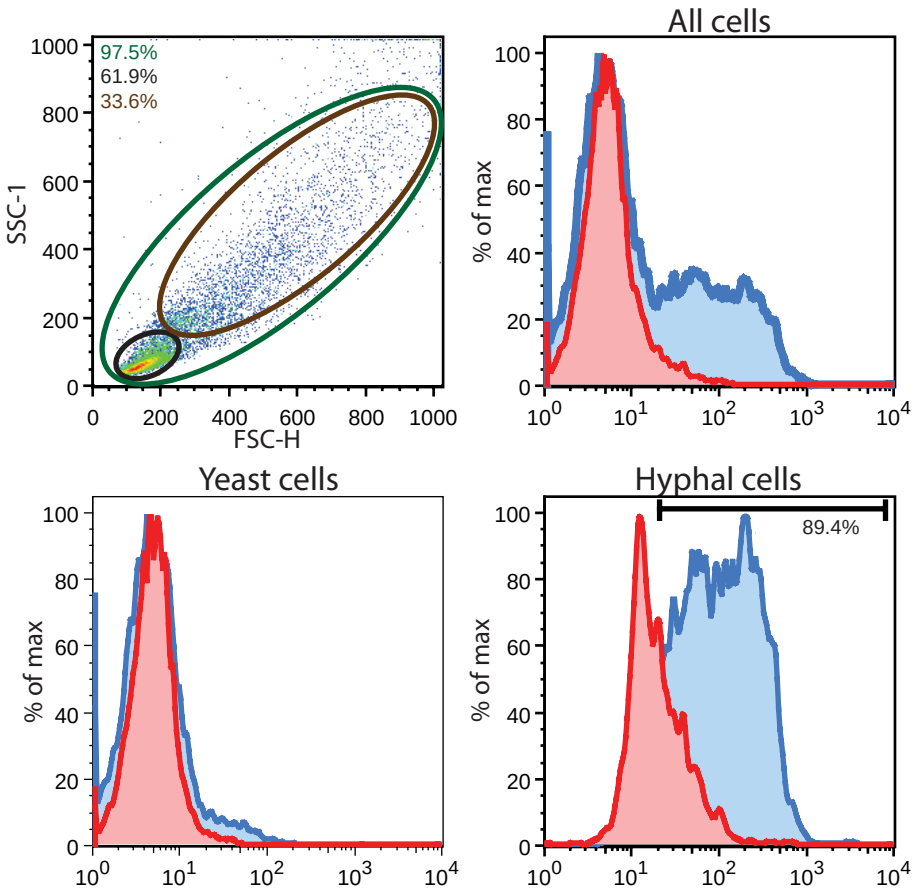


FIGURE A1 | FACS analysis of rmPtx3 binding to *C. albicans* yeast and hyphal cells. SC5314 after 1-h treatment with rmPtx3. Cells were gated according to morphology based on size, all *Candida* cells analyzed (green gate), yeast form only (black gate) and hyphal form only (the brown gate). Histograms for untreated cells (red), and treated rmPtx3 and SC5314 (blue) are shown.

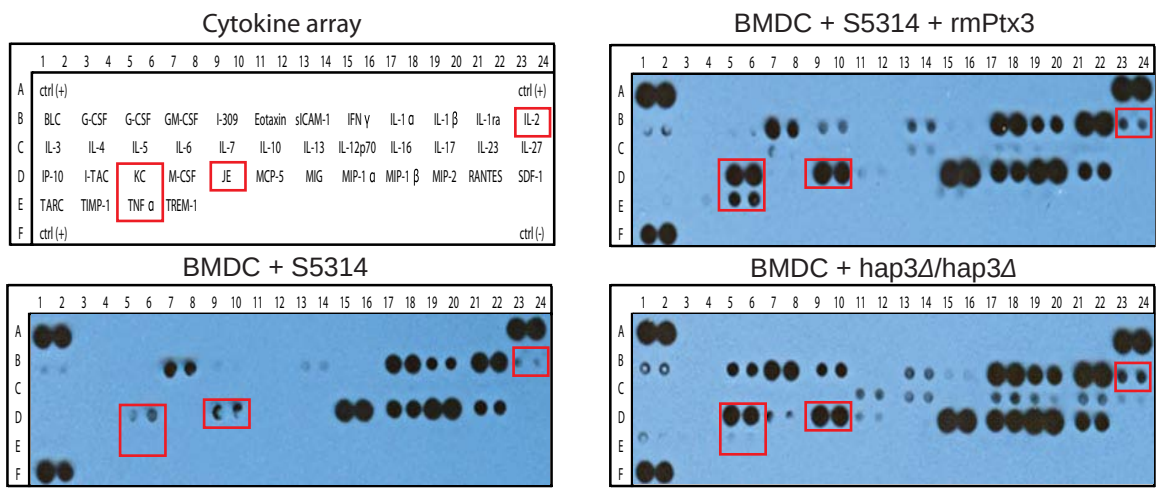


FIGURE A2 | rmPtx3 binds to the *C. albicans* surface and alters the immune response. The cytokine array panel for *M. musculus* (R&D systems) where each spot represents an individual cytokine (in duplicate) for supernatants from BMDCs treated fungal cells pre-treated with rmPtx3 after 24 h.

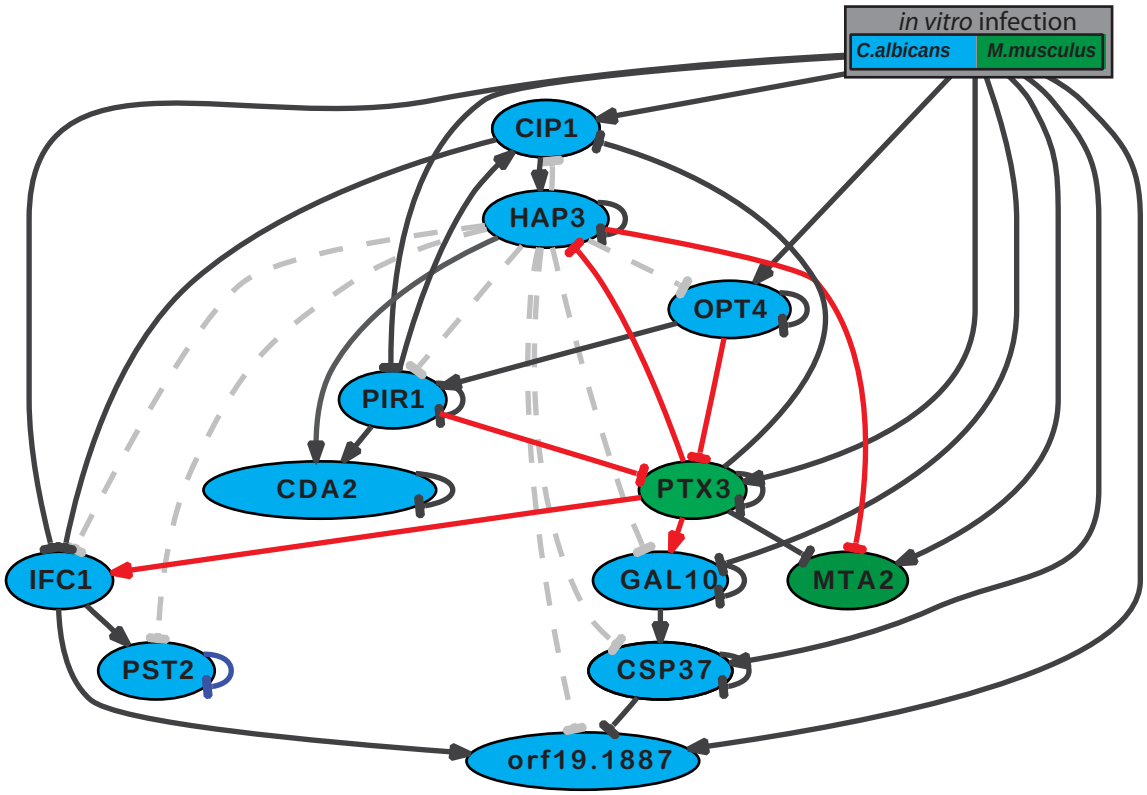


FIGURE A3 | Inferred network model between *C. albicans* and *M. musculus* using of candidate Hap3 effector genes. *C. albicans* (blue) and *M. musculus* genes (green) included in the model are shown. The following interactions are represented on the model: predicted interactions based on the RNA-seq data set from individual species where no prior knowledge exists (black) or predicted interactions between a *C. albicans* and *M. musculus* gene where no prior knowledge exists (red), or where prior knowledge exists but does not corresponds to expression data set (gray, dotted). Here, activation is shown as a pointed arrow and a repression a blunted arrow.

frontiers in
MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 14 February 2012
doi: 10.3389/fmicb.2012.00046

Infection strategies of bacterial and viral pathogens through pathogen–human protein–protein interactions

***Saliha Durmuş Tekir*[1]*, *Tunahan Çakır*[2] *and Kutlu Ö. Ülgen*[1]**

[1] *Biosystems Engineering Research Group, Department of Chemical Engineering, Boğaziçi University, İstanbul, Turkey*
[2] *Computational Systems Biology Group, Department of Bioengineering, Gebze Institute of Technology, Kocaeli, Turkey*

Edited by:
Reinhard Guthke, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany

Reviewed by:
Mirko Trilling, Heinrich-Heine-University Düsseldorf, Germany
Sylvia Schleker, Forschungszentrum Jülich, Germany

***Correspondence:**
Saliha Durmuş Tekir, Biosystems Engineering Research Group, Department of Chemical Engineering, Boğaziçi University, 34342 Bebek, İstanbul, Turkey.
e-mail: saliha.durmus@boun.edu.tr

Since ancient times, even in today's modern world, infectious diseases cause lots of people to die. Infectious organisms, pathogens, cause diseases by physical interactions with human proteins. A thorough analysis of these interspecies interactions is required to provide insights about infection strategies of pathogens. Here we analyzed the most comprehensive available pathogen–human protein interaction data including 23,435 interactions, targeting 5,210 human proteins. The data were obtained from the newly developed pathogen–host interaction search tool, PHISTO. This is the first comprehensive attempt to get a comparison between bacterial and viral infections. We investigated human proteins that are targeted by bacteria and viruses to provide an overview of common and special infection strategies used by these pathogen types. We observed that in the human protein interaction network the proteins targeted by pathogens have higher connectivity and betweenness centrality values than those proteins not interacting with pathogens. The preference of interacting with hub and bottleneck proteins is found to be a common infection strategy of all types of pathogens to manipulate essential mechanisms in human. Compared to bacteria, viruses tend to interact with human proteins of much higher connectivity and centrality values in the human network. Gene Ontology enrichment analysis of the human proteins targeted by pathogens indicates crucial clues about the infection mechanisms of bacteria and viruses. As the main infection strategy, bacteria interact with human proteins that function in immune response to disrupt human defense mechanisms. Indispensable viral strategy, on the other hand, is the manipulation of human cellular processes in order to use that transcriptional machinery for their own genetic material transcription. A novel observation about pathogen–human systems is that the human proteins targeted by both pathogens are enriched in the regulation of metabolic processes.

Keywords: pathogen–human protein–protein interactions, PHISTO, infection strategy, hub, bottleneck, gene ontology

INTRODUCTION

According to a report of World Health Organization (WHO), more than 20% of total deaths in the world are due to infectious diseases (World Health Organization, 2008). Different types of microorganisms (bacteria, fungi, protozoa, and viruses) act as pathogens for such diseases. The mechanism of infection is based on the interactions between the proteins of pathogen and host. Thanks to the developments in high-throughput protein interaction detection methods, it is possible to identify pathogen–host protein–protein interactions (PHIs) at large-scale. Infection strategies have been studied through intraspecies protein interactions of various pathogens (Flajolet et al., 2000; McCraith et al., 2000; Rain et al., 2001; LaCount et al., 2005; Uetz et al., 2006; Calderwood et al., 2007; Wang et al., 2010) as well as through interspecies protein interactions between pathogens and human (Filippova et al., 2004; Mogensen et al., 2006; Uetz et al., 2006; Calderwood et al., 2007; König et al., 2008). Notwithstanding these, a general overview of infection mechanisms of different types of pathogens is missing since there has been a lack of interspecies interactome data until very recent years.

A major step toward a complete picture of the pathogenesis of infectious diseases and consequently identifying drug targets is the cataloging of large-scale PHIs. There are few PHI-specific databases, which enable the access to PHI data for each type of pathogen from a single source (Driscoll et al., 2009; Kumar and Nanduri, 2010). Nevertheless these databases have not been updated since their first release, and miss lots of recently reported PHIs. Therefore, we have recently developed a pathogen–host interaction search tool (PHISTO), which serves as a centralized and up-to-date source for the entire available PHI data between various pathogen strains and human via a user-friendly and functional interface[1].

The systemic analysis of PHI data has so far focused mainly on virus-based infections due to the scarcity of data for other types of pathogenic organisms (Uetz et al., 2006; Calderwood et al., 2007; Dyer et al., 2008). We have enough bacterial PHIs to get statistically meaningful results, providing a good opportunity to get a

[1] http://www.phisto.boun.edu.tr

systemic picture of the pathogenesis of bacterial infections. In this work, we studied up-to-date PHI data reported in PHISTO with a specific focus on comparison between bacterial and viral infections of human. We constructed various sets of human proteins targeted by bacteria, fungi, protozoa, and viruses to pick out specific infection strategies of different pathogen types. On the other hand, the set of human proteins targeted by both bacteria and viruses were used to obtain common infection strategies of these pathogens. We computed degree (connectivity) and betweenness centrality distributions of the human protein sets targeted by bacteria and viruses to observe the network properties of targeted human proteins. Additionally, we computed gene ontology (GO; Ashburner et al., 2000) terms enriched in each above-mentioned protein set to find out attacked mechanisms in human. GO enrichment analysis was also performed for sets of human proteins targeted by each pathogen group included in our PHI data to decipher the pathogen group(s) manipulating specific processes in human.

MATERIALS AND METHODS

PATHOGEN–HOST INTERACTION SEARCH TOOL

We have developed PHISTO that presents experimentally verified pathogen–human protein interaction data in the most comprehensive and updated manner. The database provides the entirety of relevant information about the physical PHIs in a single non-redundant resource to researchers. It offers access via a user-friendly and functional web interface (see text footnote 1) with various searching, filtering, browsing, and extraction options. Results are displayed in a very clear and consistent presentation format. PHISTO enables the users to reach additional information easily by providing links in the search results to external databases. Proteins, pathogens, and publications listed in the search results are linked to UniProt, NCBI Taxonomy, and PubMed, respectively, offering users quick navigation in these informative databases.

We downloaded the pathogen–human PHIs from PHISTO in October 2011. The data cover 23,435 physical interactions occurring between 5,210 human proteins and 3,419 proteins of 257 pathogen strains of 72 pathogen groups (24 bacterial groups, 3 fungal groups, 2 protozoan groups, and 43 viral groups; **Table 1**). In PHISTO, pathogen strains are grouped to provide an option to present PHI results together for related strains. Bacterial groups are sets of strains of the same genus as the names of the groups are the names of the genuses. For viral groups, there are two definitions. Some viral groups are sets of strains of the same family as the names of the groups are the names coming from the families of the strains (e.g., papillomaviruses, herpesviruses, polyomaviruses). Some viral strains are grouped based on the related infections caused by them, as the names are generally coming from the diseases (e.g., HIV, hepatitis viruses, anemia viruses). Detailed PHI data for 72 groups are given in Data Sheets 1–4 in Supplementary Material.

HUMAN PPI DATA

To obtain degree and betweenness centrality values of pathogen-targeted proteins, the human protein–protein interaction (PPI) network was constructed by downloading 194,006 interactions between 13,015 human proteins from BioGRID (Stark et al., 2011), DIP (Salwinski et al., 2004), IntAct (Kerrien et al., 2012), Mint (Ceol et al., 2010), and Reactome (Matthews et al., 2009; Croft et al., 2011) in April 2011.

HUMAN PROTEIN SETS

A total of 10 sets of human proteins interacting with pathogens were constructed from PHI data to analyze the properties of targeted human proteins as follows: The sets targeted by bacterial pathogens (bacteria-targeted set), fungal pathogens (fungi-targeted set), protozoan pathogens (protozoa-targeted set), and viral pathogens (virus-targeted set) were analyzed for specific infection strategies of these different pathogen types. For a deeper comparison between bacterial and viral infections, human proteins interacting with at least two bacterial groups (two-bacteria-targeted set) and two viral groups (two-viruses-targeted set) and also human proteins interacting with at least three bacterial groups (three-bacteria-targeted set) and three viral groups (three-viruses-targeted set) were used. To obtain common infection strategies of pathogens, sets of human proteins targeted by all types of pathogens (pathogen-targeted set) and by both bacteria and viruses (bacteria–virus-targeted set) were also analyzed. Finally, 72 sets of human proteins each targeted by a pathogen group reported in **Table 1** were additionally used in GO enrichment analysis to investigate the human mechanisms attacked by each pathogen group in the PHI data. Totally, 82 human protein sets were constructed and analyzed.

DEGREE AND BETWEENNESS CENTRALITY CALCULATIONS

Degree of a protein within a network is defined as its number of connections. Betweenness centrality of a protein is equal to the number of shortest paths between any pairs passing through that protein. The degree and centrality values of proteins in interaction networks provide valuable information about the role of corresponding proteins in the network's functional organization using the topology of the interconnections. For instance, hubs (highly connected proteins) and bottlenecks (central proteins to many paths in the network) are critical players in the intraspecies protein networks for information flow (Barabasi and Oltvai, 2004; Yu et al., 2007).

The undirected human PPI network was represented as an adjacency matrix, and the degree and centrality values of each protein in the network were calculated in MATLAB environment. Betweenness centrality calculations were performed by freely available MATLAB BGL package developed by David Gleich[2]. The results were normalized by $(n-1)(n-2)$, where n is the number of all proteins in the PPI network. Self-interactions were not taken into account in these calculations.

GO ENRICHMENT ANALYSIS

Gene Ontology (Ashburner et al., 2000) enrichments of all 82 human protein sets were performed using BiNGO plugin (ver. 2.44) of Cytoscape (ver. 2.8.1; Maere et al., 2005). Significance level was set to 0.05 meaning that only terms enriched with a p-value of at most 0.05 were considered. All three GO terms (biological process, molecular function, and cellular component) were

[2]https://launchpad.net/matlab-bgl

Table 1 | Contents of pathogen–human PHI data.

Pathogen	Number of strains	Number of PHIs	Number of targeting pathogen proteins	Number of targeted human proteins
BACTERIA	**41**	**8,549**	**2,591**	**3,589**
Aeromonas	1	2	1	2
Bacillus	2	3,021	940	1,736
Campylobacter	1	3	1	3
Chlamydia	2	21	3	21
Citrobacter	1	1	1	1
Clostridium	3	47	9	10
Corynebacterium	1	1	1	1
Escherichia	4	30	14	27
Finegoldia	1	1	1	1
Francisella	1	1,338	346	986
Helicobacter	2	3	3	2
Klebsiella	1	1	1	1
Legionella	1	1	1	1
Listeria	1	4	4	3
Moraxella	1	1	1	1
Mycoplasma	1	2	1	2
Neisseria	1	17	1	17
Pseudomonas	1	12	4	10
Salmonella	1	5	5	5
Shigella	1	11	9	8
Staphylococcus	3	12	10	10
Streptococcus	5	15	12	9
Vibrio	1	1	1	1
Yersinia	4	3,999	1,221	2,120
FUNGI	**3**	**4**	**3**	**4**
Candida	1	1	1	1
Pneumocystis	1	1	1	1
Radiomyces	1	2	1	2
PROTOZOA	**4**	**9**	**5**	**9**
Plasmodium	3	8	4	8
Toxoplasma	1	1	1	1
VIRUS	**209**	**14,873**	**820**	**2,398**
Adenovirus	13	121	36	80
Anemia virus	6	8	6	4
ASFV	1	1	1	1
Bacteriophage	6	6	6	5
Coxsackie virus	1	1	1	1
Dengue virus	3	3	3	2
Ebola virus	1	1	1	1
Echo virus	2	3	3	1
Ectromelia virus	1	2	2	2
Encephalitis virus	1	2	1	2
Foamy virus	1	1	1	1
Hantaan virus	1	6	1	6
Hendra virus	1	1	1	1
Hepatitis virus	21	1,573	179	399
Herpesvirus	28	666	141	388
HIV	49	11,435	279	1,601
Influenza virus	9	523	27	182
Leukemia virus	3	10	3	10
Measles virus	3	10	4	4

(Continued)

Table 1 | Continued

Pathogen	Number of strains	Number of PHIs	Number of targeting pathogen proteins	Number of targeted human proteins
Molluscum virus	1	1	1	1
MPMV	1	1	1	1
Nipah virus	1	1	1	1
Nucleopolyhedrovirus	1	1	1	1
Orf virus	2	2	2	1
Papillomavirus	14	290	51	128
Parainfluenza virus	1	2	1	2
Parvo virus	1	1	1	1
Polio virus	2	3	2	2
Polyomavirus	4	64	10	45
Puumala virus	1	3	1	3
Rabies virus	1	1	1	1
Rhino virus	1	1	1	1
Rota virus	4	8	5	6
Sarcoma virus	5	15	6	11
SARS	1	4	3	4
Sendai virus	1	1	1	1
Seoul virus	1	4	1	4
SIV	2	3	2	3
Stomatitis virus	3	7	3	6
T-lymphotropic virus	3	38	7	35
Tula virus	1	2	1	2
Vaccinia virus	5	46	20	33
West Nile virus	1	1	1	1
TOTAL	**257**	**23,435**	**3,419**	**5,210**

See Data Sheets 1–4 in Supplementary Material for detailed information.

scanned to identify the terms having significant association with each human protein set studied.

RESULTS

PATHOGEN-TARGETED HUMAN PROTEINS

The distribution of 5,210 human proteins on their targeting pathogens are shown in the Venn diagram (**Figure 1**). Detailed properties of all pathogen-targeted human proteins including number and types of targeting pathogens together with degree and betweenness centrality values in the human PPI network are given in Data Sheet 5 in Supplementary Material. The most targeted human proteins are listed in **Table 2**. The top of this list, P53 (Tumor suppressor), DRA (HLA class II histocompatibility antigen, DR alpha chain), SUMO1 (Small ubiquitin-related modifier 1), JUN (Transcription factor AP-1), NPM (Nucleophosmin), ROA1 (Heterogeneous nuclear ribonucleoprotein A1), and UBC9 (SUMO-conjugating enzyme) and the following proteins have potential to give important insights about infections.

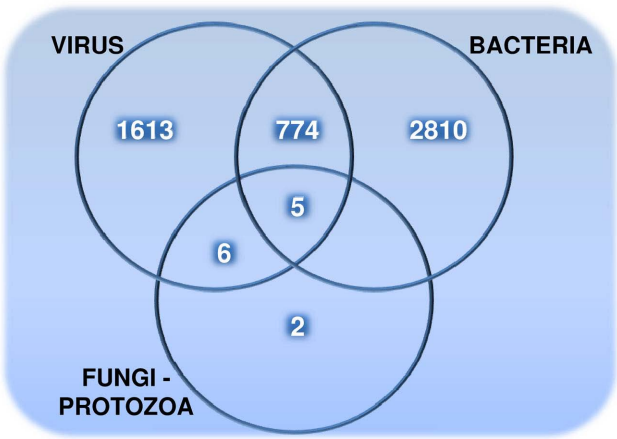


FIGURE 1 | The number of pathogen-targeted human proteins that are grouped based on their interactions with viruses, bacteria, and fungi – protozoa (targeted by fungi and/or protozoa).

DEGREE AND CENTRALITY DISTRIBUTIONS

Figure 2 displays the comparison between the degree distributions of non-targeted proteins in the human PPI network and bacteria and virus-targeted sets. For both cases of bacteria and virus-targeted sets, it is observed that pathogen-targeted human proteins have generally higher degree values than non-targeted ones. However a difference is observed in trends of degree distributions of multibacteria and multiviruses targeted sets. For bacteria-targeted cases, the increase in degree values of human proteins with increasing number of targeting pathogen groups is not as clear as those of virus-targeted cases (**Figure 2**). Very similar trends are obtained for centrality distributions of human proteins (**Figure 3**). In order

Table 2 | Highly targeted human proteins.

Protein	Degree	Betweenness centrality	Targeting bacterial groups	Targeting viral groups
P53	347	0.01547	*Bacillus, Escherichia, Francisella, Yersinia*	Adenovirus, Hepatitis virus, HIV, Papillomavirus, Polyomavirus, SIV
DRA	52	0.00003	*Bacillus, Francisella, Mycoplasma, Staphylococcus, Yersinia*	Herpesvirus, HIV, Influenza virus
SUMO1	103	0.00366	*Bacillus*	Herpesvirus, HIV, Papillomavirus, Puumala virus, SARS, Tula virus, Vaccinia virus
JUN	122	0.00335	*Bacillus, Francisella, Yersinia*	Hepatitis virus, HIV, Papillomavirus, Vaccinia virus
NPM	137	0.00166	*Bacillus, Francisella, Yersinia*	Adenovirus, Hepatitis virus, Herpesvirus, HIV
ROA1	246	0.00262	*Bacillus, Francisella, Yersinia*	Herpesvirus, HIV, Influenza virus, SARS
UBC9	134	0.00410	*Yersinia*	Hantaan virus, Herpesvirus, HIV, Influenza virus, Papillomavirus, Seoul virus
IGHG1	57	0.00219	*Bacillus, Francisella, Staphylococcus, Streptococcus, Yersinia*	Herpesvirus
RAC1	239	0.00279	*Bacillus, Clostridium, Pseudomonas, Salmonella, Yersinia*	HIV
CDC42	232	0.00405	*Bacillus, Francisella, Salmonella, Yersinia*	HIV, T-lymphotropic virus
DRB5	–	–	*Bacillus, Francisella, Streptococcus, Yersinia*	Herpesvirus, HIV
LCK	147	0.00202	*Bacillus, Francisella, Yersinia*	Hepatitis virus, Herpesvirus, HIV
XRCC6	131	0.00445	*Bacillus, Francisella, Yersinia*	Herpesvirus, HIV, Polyomavirus
KPYM	76	0.00041	*Bacillus, Francisella, Yersinia*	Hepatitis virus, Herpesvirus, Papillomavirus
ROA2	189	0.00069	*Bacillus, Francisella, Yersinia*	Herpesvirus, Influenza virus, Vaccinia virus
P85A	402	0.00914	*Bacillus, Francisella, Yersinia*	Anemia virus, HIV, Influenza virus
STAT3	77	0.00133	*Bacillus, Francisella, Yersinia*	Hepatitis virus, Herpesvirus, HIV
STAT1	71	0.00104	*Bacillus, Francisella, Yersinia*	Adenovirus, Herpesvirus, HIV
GBLP	93	0.00265	*Bacillus, Francisella, Yersinia*	Adenovirus, Herpesvirus, HIV
PARP4	1	0.00000	*Bacillus, Francisella, Yersinia*	Hepatitis virus, Herpesvirus, HIV
RB	149	0.00282	*Yersinia*	Adenovirus, Herpesvirus, HIV, Papillomavirus, Polyomavirus
SP1	103	0.00268	*Yersinia*	Adenovirus, Herpesvirus, HIV, Polyomavirus, T-lymphotropic virus
TAF1	58	0.00025	*Bacillus*	Adenovirus, Hepatitis virus, HIV, Papillomavirus, Polyomavirus
CDK2	151	0.00220	*Shigella*	Herpesvirus, HIV, Papillomavirus, Polyomavirus, T-lymphotropic virus
TF2B	69	0.00020	*Bacillus*	Hepatitis virus, Herpesvirus, HIV, Papillomavirus, Polyomavirus
EP300	123	0.00245	*Bacillus*	Adenovirus, Hepatitis virus, HIV, Papillomavirus, Polyomavirus
CBP	147	0.00304	*Yersinia*	Adenovirus, Hepatitis virus, HIV, Papillomavirus, Polyomavirus
TBP	147	0.00241	–	Adenovirus, Hepatitis virus, Herpesvirus, HIV, Papillomavirus, Polyomavirus

The targeting pathogenic proteins for each human protein can be obtained from Data Sheets 1–4 in Supplementary Material.

to justify these global trends, the same analyses was then repeated with human protein sets excluding the overrepresented pathogens, i.e., *Bacillus*, *Yersinia*, and HIV which target the largest number of human proteins (**Table 1**). Similar results are still obtained when major pathogen groups are eliminated (**Figure 4**).

GO ENRICHMENT ANALYSIS

All enriched GO terms for each human protein set are available in Data Sheets 6–10 in Supplementary Material for further detailed analyses. Special attention should be paid to the results of sets of human proteins interacting with three and more bacterial groups and three and more viral groups for a comparison between their infection strategies. The human proteins targeted by more pathogen groups reflect more specificity to infection mechanism of the corresponding pathogen (bacteria or virus). The enriched GO terms in human proteins interacting with both bacterial and viral pathogens are also important to highlight common infection mechanisms. The first 20 enriched GO process terms for three-bacteria-targeted-set, three-viruses-targeted-set, and bacteria–virus-targeted set are listed in **Tables 3–5** to point out the human processes that are attacked by pathogens.

DISCUSSION

In this study, we aim to provide a general overview of infection strategies used by different pathogens based on the comprehensive PHI data in PHISTO. Although large-scale pathogen–human protein interaction data have been identified in the last few years, the data for fungal and protozoan systems are still scarce (**Table 1**) to extract significant conclusions about their infection mechanisms. On the other hand, interspecies protein interaction networks for bacterial and viral pathogens with human have been identified,

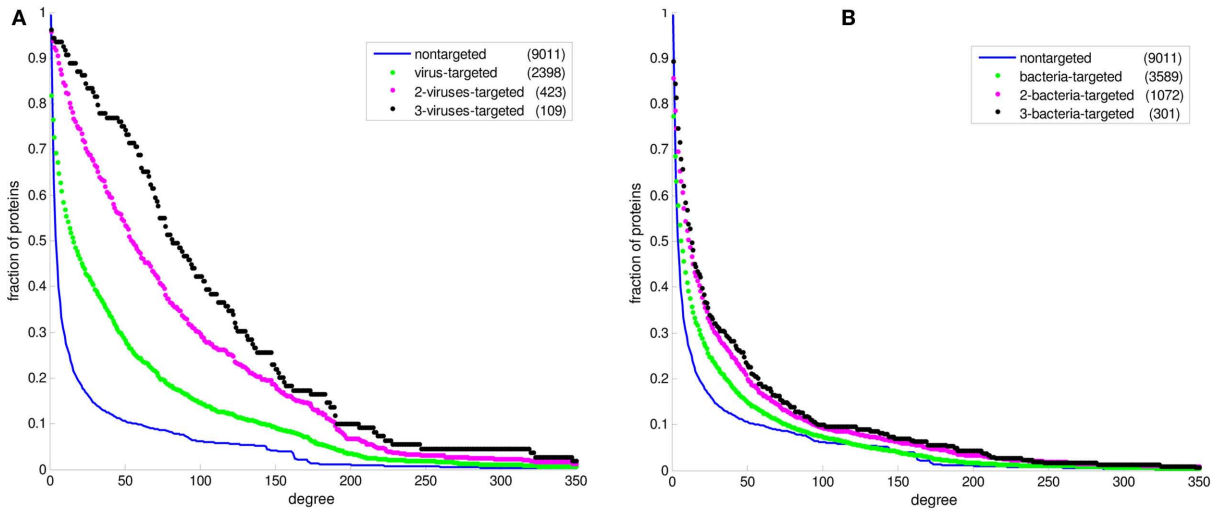


FIGURE 2 | The cumulative degree distributions of human protein sets. The distribution of all proteins in the PPI network is given in comparison with **(A)** the bacteria-targeted sets, and **(B)** the virus-targeted sets. The number of proteins in each set is given in the parentheses. The fraction of proteins at a particular value of degree is the number of proteins having that value and greater divided by the number of proteins in the set.

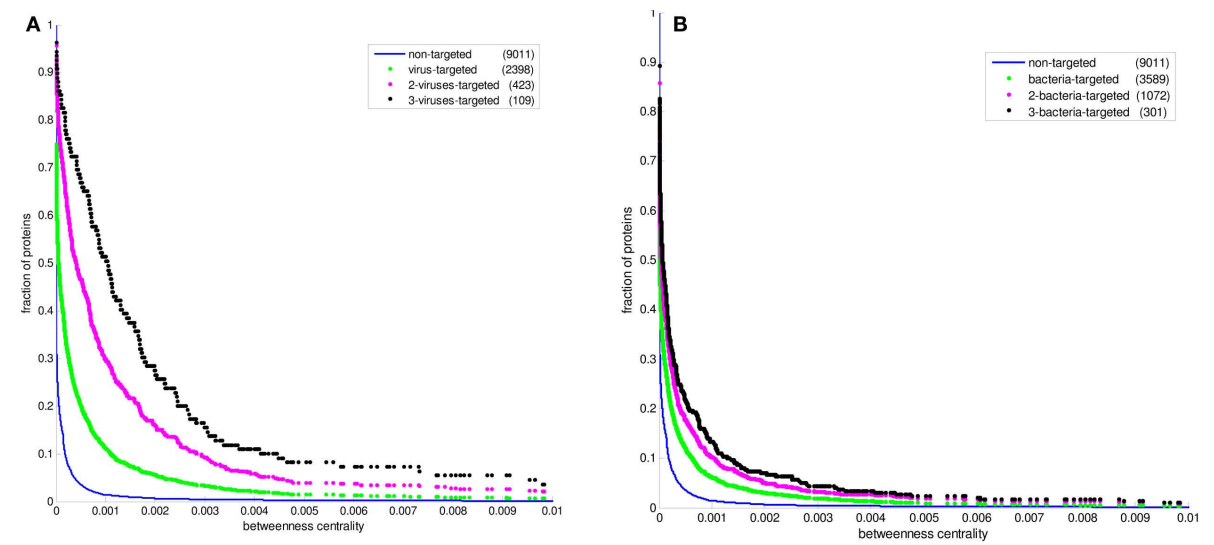


FIGURE 3 | The cumulative betweenness centrality distributions of human protein sets. The distribution of all proteins in the PPI network is given in comparison with **(A)** the bacteria-targeted sets, and **(B)** the virus-targeted sets. The number of proteins in each set is given in the parentheses. The fraction of proteins at a particular value of betweenness centrality is the number of proteins having that value and greater divided by the number of proteins in the set.

enough to provide some insights about their strategies to subvert human cellular processes during infection.

With increasing PHI data of bacterial and viral pathogens, studies have been performed to enlighten specific bacteria–human (Mogensen et al., 2006; Dyer et al., 2010) and virus–human (Filippova et al., 2004; Uetz et al., 2006; Calderwood et al., 2007; König et al., 2008) interaction systems. Although some studies provided global views of infection strategies of viruses (Dyer et al., 2008) and bacteria (Dyer et al., 2010) separately, they do not provide a direct comparison between bacterial and viral infections. In fact, only <2% of the PHI data of Dyer et al. (2008) are for bacteria–human interactions whereas it is more than 36% in our database of PHISTO. Hence, our study constitutes the first extensive comparison between bacteria–human and virus–human interspecies protein interaction networks to retrieve information about infection strategies specific to each system and then common to both systems. Our findings should be interpreted with caution since the protein interaction networks between pathogens and human are not complete yet.

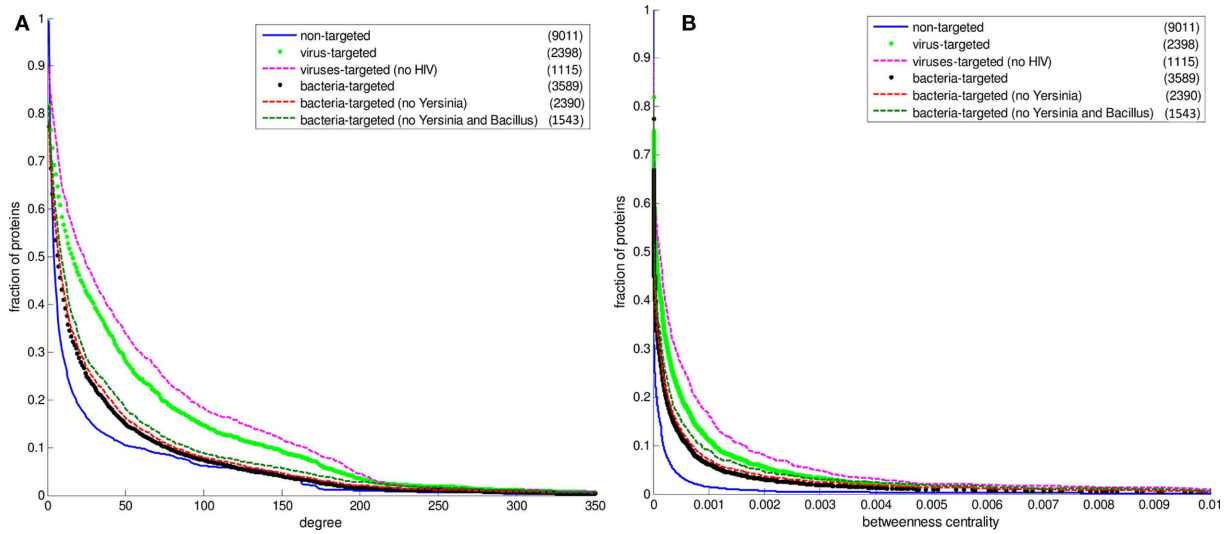


FIGURE 4 | The cumulative distributions of degree and betweenness centrality of human proteins excluding *Yersinia* and HIV data. The number of proteins in each set is given in the parentheses. **(A)** The degree distributions **(B)** the betweenness centrality distributions. The fraction of proteins at a particular value of degree is the number of proteins having that value and greater divided by the number of proteins in the set.

SPECIAL INFECTION STRATEGIES

In recent studies it has been suggested that viral proteins (Calderwood et al., 2007; Dyer et al., 2008; Itzhaki, 2011) and bacterial proteins (Dyer et al., 2010) have evolved to preferentially interact with hubs and bottlenecks in the human PPI network. The degree and betweenness centrality distributions of the bacteria-targeted and virus-targeted human protein sets are displayed in comparison with non-targeted proteins in the human PPI network in **Figures 2** and **3**. We observe that the degree and centrality values of human proteins increase with increasing number of targeting bacterial and viral groups, confirming the previous results with the most comprehensive PHI data. A novel finding by our comparative analysis is that the increase in degree and centrality values with increasing number of pathogen groups is much more pronounced in virus-targeted cases than bacteria-targeted cases (**Figures 2** and **3**). Therefore we can conclude that attacking to hub and bottleneck proteins in the human interaction network is more specific to viral infections.

In our PHI data, some pathogen groups are overrepresented with their larger number of reported interactions with human (**Table 1**). As most of these large-scale data have been produced with high-throughput detection methods, which are prone to experimental biases and errors, it was necessary to check whether the distributions of degree and centrality values of the pathogen-targeted human proteins would be same without the groups with large number of interacting partners of human proteins (i.e., *Bacillus*, *Yersinia*, and HIV). Hence, we performed the above-mentioned analyses with human protein sets excluding these major pathogen groups. 1,199 human proteins targeted by only *Yersinia* strains were excluded from the bacteria-targeted set, and 1,283 human proteins targeted by only HIV strains were excluded from the virus-targeted set to obtain the degree and betweenness centrality values of the remaining human proteins. We also analyzed the human proteins targeted by bacteria other than *Bacillus* and *Yersinia* to exclude the effect of large-scale data of the two. 1,199 only *Yersinia*-targeted and 847 only *Bacillus*-targeted human proteins were excluded from the bacteria-targeted set. The behavior of the remaining human proteins can be observed in **Figure 4** resulting in similar trends with the global case. Additionally, a direct comparison of the degree and centrality between bacteria and virus-targeted interaction partners with respect to non-targeted human proteins is also given in **Figure 4**. The difference in the behavior of the bacteria- and virus-targeted sets are clear especially in degree distributions (**Figure 4A**). The degree values of bacteria-targeted human proteins with or without *Bacillus* and *Yersinia* are nearly same. On the other hand, attack of viruses to more connected human proteins is more clear when HIV data are excluded.

From the enriched GO process terms in human proteins targeted by at least three bacterial groups (**Table 3**), we can conclude that bacteria may have adapted to attack proteins involved generally in human immunity pathways. Therefore, the most specific bacterial infection strategy is through evading or suppressing human immune responses as also concluded previously (Lai et al., 2001; Park et al., 2002; Zhang et al., 2005; Dyer et al., 2010). The human immune system is manipulated by bacterial proteins attacking human proteins functioning in innate and adaptive immunity (i.e., TLR4 and TLR7), inflammation (i.e., NF-κB and BCL6), and activation of T cells (i.e., CXCR4 and LCK; Zhang and Ghosh, 2000; Alonso et al., 2004; Oda and Kitano, 2006; Dyer et al., 2010). In our PHI data it is observed that *Yersinia* bacteria attack all these human defense mechanisms targeting all mentioned human proteins. Proteins of *Bacillus* and *Francisella* interact with NF-κB and LCK (Dyer et al., 2010) aiming to disrupt the mechanisms of inflammation and T cell responses. On the other hand, proteins of Chlamydia, *Escherichia*, and *Neisseria* interact with crucial players of innate and adaptive immunity, toll-like receptors (TLR4 and TLR7; Croft et al., 2011) to collapse the human immune system.

Table 3 | First 20 enriched GO process terms in human proteins targeted by at least three bacterial groups (three-bacteria-targeted set).

GO process term	*p*-Value
I-kappaB kinase/NF-kappaB cascade	9.64E−13
Regulation of biological process	9.69E−10
Biological regulation	2.59E−09
Negative regulation of biological process	4.89E−09
Positive regulation by organism of immune response of other organism involved in symbiotic interaction	6.64E−09
Modulation by organism of immune response of other organism involved in symbiotic interaction	6.64E−09
Modulation by symbiont of host immune response	6.64E−09
Positive regulation by symbiont of host immune response	6.64E−09
Response to immune response of other organism involved in symbiotic interaction	6.64E−09
Response to host immune response	6.64E−09
Positive regulation by organism of defense response of other organism involved in symbiotic interaction	6.64E−09
Positive regulation by symbiont of host defense response	6.64E−09
Positive regulation by organism of innate immunity in other organism involved in symbiotic interaction	6.64E−09
Modulation by organism of innate immunity in other organism involved in symbiotic interaction	6.64E−09
Pathogen-associated molecular pattern dependent modulation by organism of innate immunity in other organism involved in symbiotic interaction	6.64E−09
Modulation by organism of defense response of other organism involved in symbiotic interaction	6.64E−09
Pathogen-associated molecular pattern dependent induction by organism of innate immunity of other organism involved in symbiotic interaction	6.64E−09
Modulation by symbiont of host defense response	6.64E−09
Pathogen-associated molecular pattern dependent induction by symbiont of host innate immunity	6.64E−09
Modulation by symbiont of host innate immunity	6.64E−09

See Data Sheet 10 in Supplementary Material for the whole list and the human proteins corresponding to each GO term.

Table 4 | First 20 enriched GO process terms in human proteins targeted by at least three viral groups (three-viruses-targeted set).

Go process term	*p*-Value
Interspecies interaction between organisms	1.89E−40
Multi-organism process	1.19E−27
Positive regulation of cellular process	1.12E−17
Positive regulation of biological process	1.06E−16
Cellular macromolecule metabolic process	1.12E−15
Nucleic acid metabolic process	4.49E−14
Positive regulation of macromolecule metabolic process	4.60E−14
Cell cycle process	6.72E−14
Positive regulation of gene expression	1.49E−13
Cell cycle	2.06E−13
Positive regulation of metabolic process	3.79E−13
Positive regulation of transcription	4.37E−13
Macromolecule metabolic process	8.51E−13
Positive regulation of cellular metabolic process	3.89E−12
Cellular response to stimulus	6.61E−12
Positive regulation of nucleobase, nucleoside, nucleotide, and nucleic acid metabolic process	7.26E−12
Positive regulation of macromolecule biosynthetic process	1.32E−11
Positive regulation of nitrogen compound metabolic process	1.41E−11
Positive regulation of transcription, DNA-dependent	1.44E−11
Regulation of cell cycle	1.47E−11

See Data Sheet 10 in Supplementary Material for the whole list and the human proteins corresponding to each GO term.

There are several other bacteria-targeted human proteins involved in the immune system. Their interactions with bacterial proteins should be investigated carefully for a complete understanding of bacterial strategies targeting human defense mechanism during infection.

Viruses attack human cellular processes (**Table 4**) enabling themselves to proliferate in human during infection. All viruses use this mechanism since they need host's transcriptional machinery for viral genetic material transcription. Even the human proteins targeted by only one viral group are enriched in GO process terms relevant to regulation of cellular mechanisms (Data Sheet 10 in Supplementary Material). Viruses manipulate human cellular mechanisms by interacting with various proteins functioning in cell cycle (i.e., DLG1, PTMA, and EP300), with human transcription factors to promote their own genetic material transcription (i.e., E2F1 and TAF1), with key proteins controlling apoptosis (i.e., P53), and with nuclear membrane proteins for transporting their genetic material across the nuclear membrane (i.e., RAN, and SUMO1; Lechner and Laimins, 1994; Thompson et al., 1997; Carrillo et al., 2004; Thomas et al., 2005; Dyer et al., 2008). In our PHI data, Adenoviruses, HIV, Papillomaviruses, and Polyomaviruses are observed to target one or more proteins in each of four groups; cell cycle proteins, transcription factors, apoptosis regulator, and nuclear membrane proteins. Proteins of Hepatitis viruses interact with PTMA, EP300, TAF1, and p53 while proteins of Herpesviruses interact with PTMA and SUMO1. On the other hand, viral groups of Influenza, Puumala, Tula, SARS, and Vaccinia are observed to target nuclear membrane proteins. The other virus-targeted human proteins involved in cellular mechanisms should be investigated comprehensively for a complete understanding of viral strategies targeting human cellular mechanism.

We can conclude that the main infection strategies of bacteria and viruses are through attacking human immune system and cellular processes, respectively. However, there are some exceptions such that some bacterial groups target human proteins functioning in cellular mechanisms whereas some viral groups target human proteins functioning in defense mechanisms. In the case of bacteria, the difference might arise from the life-style, e.g., intracellular bacteria like Chlamydia, *Listeria*, and *Mycoplasma* are able to grow and reproduce only within the host cells just like viruses (Kaufmann, 1993). Therefore, human protein sets targeted by these intracellular bacterial groups are enriched in GO process terms related to the cellular mechanisms (e.g., regulation of cellular

Table 5 | First 20 enriched GO process terms in human proteins targeted by both bacterial and viral groups (bacteria–virus-targeted set).

GO process term	*p*-Value
Interspecies interaction between organisms	1.64E−52
Multi-organism process	1.01E−47
Positive regulation of biological process	5.62E−47
Regulation of biological process	1.32E−42
Biological regulation	2.66E−40
Positive regulation of cellular process	4.59E−40
Negative regulation of biological process	6.32E−37
Regulation of cellular process	2.00E−36
Negative regulation of cellular process	4.84E−32
Regulation of protein metabolic process	7.68E−30
Regulation of macromolecule metabolic process	3.21E−29
Regulation of cellular protein metabolic process	1.39E−28
Regulation of cell death	1.74E−28
Positive regulation of macromolecule metabolic process	1.91E−28
Positive regulation of cellular metabolic process	7.04E−28
Regulation of programmed cell death	1.13E−27
Cellular macromolecule metabolic process	1.94E−27
Positive regulation of metabolic process	2.92E−27
Negative regulation of macromolecule metabolic process	5.31E−27
Regulation of apoptosis	6.54E−27

See Data Sheet 10 in Supplementary Material for the whole list and the human proteins corresponding to each GO term.

processes, regulation of transcription) in addition to the immune sytem (Data Sheet 6 in Supplementary Material). On the other hand, viruses like herpes and pox (ectromelia, molluscum, orf, vaccinia) viruses as well as HIV have the ability to evade human immune system (Alcami and Koszinowski, 2000) as observed in our results (Data Sheet 9 in Supplementary Material).

For more specific infection strategies of pathogen groups, the results of GO enrichment analysis for the human protein sets targeted by each of the 72 groups in the PHI data can be used (Data Sheets 6–9 in Supplementary Material). Additionally, intranetworks of pathogenic proteins in each pathogen group should be analyzed for drug target identification after a thorough understanding of pathogenesis via interspecies protein interactions.

COMMON INFECTION STRATEGIES

In spite of the difference in the trends of distributions of degree and centrality values of human proteins in bacteria-targeted and virus-targeted sets, the tendency to attack human proteins that are highly connected (hubs) and central to shortest paths (bottlenecks) is common to all types of pathogens. We observed in our PHI data that the degree and centrality values of pathogen-targeted human proteins are generally greater than non-targeted ones. This infection strategy of pathogens, attacking more connected and central nodes in the human PPI network, is probably due to enabling themselves to control and disrupt essential complexes and pathways more easily. With scale-free nature, the human PPI network is robust to attacks on random nodes. However, the selective attacks to even a small number of nodes of high degree can dramatically change the topology and functionality of the network (Albert et al., 2000; Li et al., 2006).

Although bacteria and viruses have a tendency to interact with different human proteins (**Figure 1**), they together target those (779 human proteins) enriched in the regulation of metabolic processes in addition to cellular processes (**Table 5**). For instance, a pyruvate kinase isozyme, KPYM, functions in glycolysis catalyzing the transfer of a phosphoryl group from phosphoenolpyruvate (PEP) to ADP, generating ATP. This metabolic human protein is targeted by three bacterial (*Bacillus*, *Francisella*, and *Yersinia*) and three viral groups (Hepatitis, Herpesviruses, and Papillomaviruses) in the PHI data. Alpha-enolase is another bacteria–virus-targeted enzyme which functions in glycolysis, just before KPYM enzyme, converting 2-phospho-glycerate to PEP. A metabolic step operating again around lower glycolysis is the production of lactate from pyruvate. Both isoenzymes (LDHA, LDHB) are found to be a target for bacterial and viral groups. In addition to lower glycolysis, some enzymes functioning in lipid metabolism (ACSA, ACOT9, CPT1A) were identified as common targets of bacteria and viruses. Interestingly, two enzymes functioning for protection against oxidative-stress are in our common-target list: catalase (CATA) and glutathione peroxidase 3 (GPX3). These enzymes remove H_2O_2, which is a reactive oxygen species (ROS) harmful for the cell.

To our knowledge, the human proteins targeted by both bacteria and viruses have not been investigated in any previous study. Through our analyses using large-scale PHI data we can conclude that both bacteria and viruses attack to the proteins functioning in human metabolic processes as a common infection strategy. All bacteria–viruses-targeted human proteins involved in metabolic processes should be investigated carefully for a complete picture of commonalities in bacterial and viral infections.

ACKNOWLEDGMENTS

We thank Ali Semih Sayılırbaş for his invaluable contribution to the development of PHISTO database. The financial support for this research was provided by the Research Funds of Boğaziçi University and TÜBİTAK through projects 5554D and 110M428, respectively. The scholarship for Saliha Durmuş Tekir is provided by TÜBİTAK, is gratefully acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Microbial_Immunology/10.3389/fmicb.2012.00046/abstract

Data Sheets 1–4 | Pathogen–human PHIs.

Data Sheet 1 | Bacteria–human PHIs.

Data Sheet 2 | Fungi–human PHIs.

Data Sheet 3 | Protozoa–human PHIs.

Data Sheet 4 | Virus–human PHIs.

Data Sheet 5 | Properties of pathogen-targeted set.

Data Sheets 6–10 | Gene ontology enrichment results.

Data Sheet 6 | Gene ontology enrichment results of sets of human proteins targeted by each bacterial group.

Data Sheet 7 | Gene ontology enrichment results of sets of human proteins targeted by each fungal group.

Data Sheet 8 | Gene ontology enrichment results of sets of human proteins targeted by each protozoan group.

Data Sheet 9 | Gene ontology enrichment results of sets of human proteins targeted by each viral group.

Data Sheet 10 | Gene ontology enrichment results of other sets of human proteins (pathogen-targeted, virus-targeted, two-virus-targeted, three-virus-targeted, bacteria-targeted, two-bacteria-targeted, three-bacteria-targeted, fungi-targeted, protozoa-targeted, and bacteria–virus-targeted).

REFERENCES

Albert, R., Jeong, H., and Barabasi, A. L. (2000). Error and attack tolerance of complex networks. *Nature* 406, 378–382.

Alcami, A., and Koszinowski, U. H. (2000). Viral mechanisms of immune evasion. *Immunol. Today* 21, 447–455.

Alonso, A., Bottini, N., Bruckner, S., Rahmouni, S., Williams, S., Schoenberger, S. P., and Mustelin, T. (2004). Lck dephosphorylation at Tyr-394 and inhibition of T cell antigen receptor signaling by *Yersinia* phosphatase YopH. *J. Biol. Chem.* 279, 4922–4928.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, 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., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29.

Barabasi, A. L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113.

Calderwood, M. A., Venkatesan, K., Xing, L., Chase, M. R., Vazquez, A., Holthaus, A. M., Ewence, A. E., Li, N., Hirozane-Kishikawa, T., Hill, D. E., Vidal, M., Kieff, E., and Johannsen, E. (2007). Epstein–Barr virus and virus human protein interaction maps. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7606–7611.

Carrillo, E., Garrido, E., and Gariglio, P. (2004). Specific in vitro interaction between papillomavirus E2 proteins and TBP-associated factors. *Intervirology* 47, 342–349.

Ceol, A., Catr Aryamontri, A., Licata, L., Peluso, D., Briganti, L., Perfetto, L., Castagnoli, L., and Cesareni, G. (2010). MINT, the molecular interaction database: 2009 update. *Nucleic Acids Res.* 38, D532–D539.

Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., Caudy, M., Garapati, P., Gopinath, G., Jassal, B., Jupe, S., Kalatskaya, I., Mahajan, S., May, B., Ndegwa, N., Schmidt, E., Shamovsky, V., Yung, C., Birney, E., Hermjakob, H., D'Eustachio, P., and Stein, L. (2011). Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 39, D691–D697.

Driscoll, T., Dyer, M. D., Murali, T. M., and Sobral, B. W. (2009). PIG – the pathogen interaction gateway. *Nucleic Acids Res.* 37, D647–D650.

Dyer, M. D., Murali, T. M., and Sobral, B. W. (2008). The landscape of human proteins interacting with viruses and other pathogens. *PLoS Pathog.* 4, e32. doi:10.1371/journal.ppat.0040032

Dyer, M. D., Neff, C., Dufford, M., Rivera, C. G., Shattuck, D., Bassaganya-Riera, J., Murali, T. M., and Sobral, B. W. (2010). The human-bacterial pathogen protein interaction networks of *Bacillus anthracis*, *Francisella tularensis*, and *Yersinia pestis*. *PLoS ONE* 5, e12089. doi:10.1371/journal.pone.0012089

Filippova, M., Parkhurst, L., and Duerksen-Hughes, P. J. (2004). The human papillomavirus 16 E6 protein binds to Fas-associated death domain and protects cells from Fas-triggered apoptosis. *J. Biol. Chem.* 279, 25729–25744.

Flajolet, M., Rotondo, G., Daviet, L., Bergametti, F., Inchauspe, G., Tiollais, P., Transy, C., and Legrain, P. (2000). A genomic approach of the hepatitis C virus generates a protein interaction map. *Gene* 242, 369–379.

Itzhaki, Z. (2011). Domain-domain interactions underlying herpes virus human protein-protein interaction networks. *PLoS ONE* 6, e21724. doi:10.1371/journal.pone.0021724

Kaufmann, S. H. E. (1993). Immunity to intracellular bacteria. *Annu. Rev. Immunol.* 11, 129–163.

Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., Duesbury, M., Dumousseau, M., Feuermann, M., Hinz, U., Jandrasits, C., Jimenez, R. C., Khadake, J., Mahadevan, U., Masson, P., Pedruzzi, I., Pfeiffenberger, E., Porras, P., Raghunath, A., Roechert, B., Orchard, S., and Hermjakob, H. (2012). The IntAct molecular interaction database in 2012. *Nucleic Acids Res.* 40, D841–D846.

König, R., Zhou, Y., Elleder, D., Diamond, T. L., Bonamy, G. M., Irelan, J. T., Chiang, C. Y., Tu, B. P., De Jesus, P. D., Lilley, C. E., Seidel, S., Opaluch, A. M., Caldwell, J. S., Weitzman, M. D., Kuhen, K. L., Bandyopadhyay, S., Ideker, T., Orth, A. P., Miraglia, L. J., Bushman, F. D., Young, J. A., and Chanda, S. K. (2008). Global analysis of host-pathogen interactions that regulate early stage HIV-1 replication. *Cell* 135, 49–60.

Kumar, R., and Nanduri, B. (2010). HPIDB – a unified resource for host-pathogen interactions. *BMC Bioinformatics* 11, S16. doi:10.1186/1471-2105-11-S6-S16

LaCount, D. J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J. R., Schoenfeld, L. W., Ota, I., Sahasrabudhe, S., Kurschner, C., Fields, S., and Hughes, R. E. (2005). A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* 438, 103–107.

Lai, X. H., Golovliov, I., and Sjostedt, A. (2001). *Francisella tularensis* induces cytopathogenicity and apoptosis in murine macrophages via a mechanism that requires intracellular bacterial multiplication. *Infect. Immun.* 69, 4691–4694.

Lechner, M. S., and Laimins, L. A. (1994). Inhibition of p53 DNA binding by human papillomavirus E6 proteins. *J. Virol.* 68, 4262–4273.

Li, D., Li, J., Ouyang, S., Wang, J., Wu, S., Wan, P., Zhu, Y., Xu, X., and He, F. (2006). Protein interaction networks of *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster*: large-scale organization and robustness. *Proteomics* 6, 456–461.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21, 3448–3449.

Matthews, L., Gopinath, G., Gillespie, M., Caudy, M., Croft, D., de Bono, B., Garapati, P., Hemish, J., Hermjakob, H., Jassal, B., Kanapin, A., Lewis, S., Mahajan, S., May, B., Schmidt, E., Vastrik, I., Wu, G., Birney, E., Stein, L., and D'Eustachio, P. (2009). Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Res.* 37, D619–D622.

McCraith, S., Holtzman, T., Moss, B., and Fields, S. (2000). Genome-wide analysis of vaccinia virus protein–protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4879–4884.

Mogensen, T. H., Paludan, S. R., Kilian, M., and Østergaard, L. (2006). Live *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Neisseria meningitidis* activate the inflammatory response through toll-like receptors 2, 4, and 9 in species-specific patterns. *J. Leukoc. Biol.* 80, 267–277.

Oda, K., and Kitano, H. (2006). A comprehensive map of the toll-like receptor signaling network. *Mol. Syst. Biol.* 2, 2006.0015.

Park, J. M., Greten, F. R., Li, Z.-W., and Karin, M. (2002). Macrophage apoptosis by anthrax lethal factor through p38 MAP kinase inhibition. *Science* 297, 2048–2051.

Rain, J. C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schächter, V., Chemama, Y., Labigne, A., and Legrain, P. (2001). The protein–protein interaction map of *Helicobacter pylori*. *Nature* 409, 211–215.

Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U., and Eisenberg, D. (2004). The database of interacting proteins: 2004 update. *Nucleic Acids Res.* 32, D449–D451.

Stark, C., Breitkreutz, B. J., Chatr-Aryamontri, A., Boucher, L., Oughtred, R., Livstone, M. S., Nixon, J., Van Auken, K., Wang, X., Shi, X., Reguly, T., Rust, J. M., Winter, A., Dolinski, K., and Tyers, M. (2011). The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res.* 39, D698–D704.

Thomas, M., Massimi, P., Navarro, C., Borg, J. P., and Banks, L. (2005). The hScrib/Dlg apico-basal control complex is differentially targeted by HPV-16 and HPV-18 E6 proteins. *Oncogene* 24, 6222–6230.

Thompson, D. A., Belinsky, G., Chang, T. H.-T., Jones, D. L., Schlegel, R., and Münger, K. (1997). The human papillomavirus-16 E6 oncoprotein decreases the vigilance of mitotic checkpoints. *Oncogene* 15, 3025–3035.

Uetz, P., Dong, Y., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., Rajagopala, S., Roupelieva, M., Rose, D., Fossum, E., and Haas, J. (2006). Herpesviral protein networks and their interaction with the human proteome. *Science* 311, 239–242.

Wang, Y., Cui, T., Zhang, C., Yang, M., Huang, Y., Li, W., Zhang, L., Gao, C., He, Y., Li, Y., Huang, F., Zeng, J., Huang, C., Yang, Q., Tian, Y., Zhao, C., Chen, H., Zhang, H., and He, Z. G. (2010). Global protein-protein interaction network in the human pathogen *Mycobacterium tuberculosis* H37Rv. *J. Proteome. Res.* 9, 6665–6677.

World Health Organization. (2008). *The Global Burden of Disease: 2004 Update*. Geneva: WHO Press.

Yu, H., Kim, P. M., Sprecher, E., Trifonov, V., and Gerstein, M. (2007). The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* 3, e59. doi:10.1371/journal.pcbi.0030059

Zhang, G., and Ghosh, S. (2000). Molecular mechanisms of NF-κB activation induced by bacterial lipopolysaccharide through toll-like receptors. *J. Endotoxin Res.* 6, 453–457.

Zhang, Y., Ting, A. T., Marcu, K. B., and Bliska, J. B. (2005). Inhibition of MAPK and NF-kappa B pathways is necessary for rapid apoptosis in macrophages infected with *Yersinia*. *J. Immunol.* 174, 7939–7949.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 December 2011; accepted: 30 January 2012; published online: 14 February 2012.

Citation: Durmuş Tekir S, Çakır T and Ülgen KÖ (2012) Infection strategies of bacterial and viral pathogens through pathogen–human protein–protein interactions. Front. Microbio. ***3****:46. doi: 10.3389/fmicb.2012.00046*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 22 March 2012
doi: 10.3389/fmicb.2012.00104

Population dynamics of *Borrelia burgdorferi* in Lyme disease

Sebastian C. Binder*[1], *Arndt Telschow*[2] *and Michael Meyer-Hermann*[1,3]

[1] *Department of Systems Immunology, Helmholtz Centre for Infection Research, Braunschweig, Germany*
[2] *Institute for Evolution and Biodiversity, Westfälische Wilhelms-Universität, Münster, Germany*
[3] *Bio Center for Life Sciences, University of Technology Braunschweig, Braunschweig, Germany*

Edited by:
Marc Thilo Figge, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany

Reviewed by:
Franziska Mech, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany
Ute Neugebauer, Institute of Photonic Technology, Germany

***Correspondence:**
Michael Meyer-Hermann, Department of Systems Immunology, Helmholtz Centre for Infection Research, Inhoffenstr. 7, 38124 Braunschweig, Germany.
e-mail: michael.meyer-hermann@helmholtz-hzi.de

Many chronic inflammatory diseases are known to be caused by persistent bacterial or viral infections. A well-studied example is the tick-borne infection by the gram-negative spirochaetes of the genus *Borrelia* in humans and other mammals, causing severe symptoms of chronic inflammation and subsequent tissue damage (Lyme Disease), particularly in large joints and the central nervous system, but also in the heart and other tissues of untreated patients. Although killed efficiently by human phagocytic cells *in vitro*, *Borrelia* exhibits a remarkably high infectivity in mice and men. In experimentally infected mice, the first immune response almost clears the infection. However, approximately 1 week post infection, the bacterial population recovers and reaches an even larger size before entering the chronic phase. We developed a mathematical model describing the bacterial growth and the immune response against *Borrelia burgdorferi* in the C3H mouse strain that has been established as an experimental model for Lyme disease. The peculiar dynamics of the infection exclude two possible mechanistic explanations for the regrowth of the almost cleared bacteria. Neither the hypothesis of bacterial dissemination to different tissues nor a limitation of phagocytic capacity were compatible with experiment. The mathematical model predicts that *Borrelia* recovers from the strong initial immune response by the regrowth of an immune-resistant sub-population of the bacteria. The chronic phase appears as an equilibration of bacterial growth and adaptive immunity. This result has major implications for the development of the chronic phase of *Borrelia* infections as well as on potential protective clinical interventions.

Keywords: *Borrelia burgdorferi*, Lyme disease, mathematical model, immunology, macrophages, mouse model

INTRODUCTION

Originally described in 1977 (Steere et al., 1977), Lyme disease is today recognized as the most common vector-borne disease in the United States and Europe, with approximately 65,000 cases reported annually in Europe (Rizzoli et al., 2011). Its causative agent, *Borrelia burgdorferi* and several closely related species commonly referred to as *Borrelia burgdorferi* sensu lato (sl), is transmitted by hard ticks of the genus *Ixodes* (Lane et al., 1991). *Borrelia burgdorferi* sl (BB) is spreading geographically (Hubalek, 2009), and the number of reported cases is increasing, although this is possibly partly due to increased diagnostic sensitivity.

BB belongs to the spirochaetes and shares characteristic traits of this group, including the spiral shape and a remarkable motility system enabling them to move efficiently through dense material like connective tissues (Tilly et al., 2008). Their natural reservoir are mainly small rodents, birds and reptiles that can harbor BB and transmit them to uninfected ticks feeding on these animals. The spirochaete is transmitted to humans mainly by *Ixodes ricinus* and, to a lesser extent, by *I. persulcatus* (Lane et al., 1991).

The disease symptoms in humans can be roughly divided into three infection stages. In most patients, the first noticeable symptom consists of a slowly growing, circle-shaped skin rash known as *Erythema migrans (EM)* around the tick bite appearing shortly after infection. This early, localized infection symptom is often accompanied by mild fever and flu-like symptoms.

Several days to weeks after infection, the bacteria begin to disseminate to different tissues, where they cause strong inflammatory reactions and tissue damage. Tissues commonly affected include the heart, joints and the central and peripheral nervous system as well as skin tissue where the pathogen can cause multiple EM lesions. In the U.S., approximately 60% of untreated patients develop an inflammation of the synovial tissue, particularly in the large joints, referred to as Lyme Arthritis (Steere and Glickstein, 2004). Among other possible routes of infection, e.g., along peripheral nerves, hematogenous dissemination of BB seems to occur frequently at least in patients in the USA (Rupprecht et al., 2008). This stage of the infection lasts up to 6 months.

If left untreated, symptoms can occur more than 6 month up to years after disease onset; in fact, some symptoms have been shown to persist for more than 10 years (Åsbrink and Hovmark, 2009; Stanek et al., 2011), presumably by persistent infection (Steere and Glickstein, 2004). At any of these stages, the disease can be treated by antibiotics with a high degree of success (Nau et al., 2009), although Lyme Arthritis fails to resolve in a minority of patients despite apparently effective antibiotic treatment, possibly due to an autoimmune reaction triggered by the infection (Steere et al.,

2006). The different origins and exact nature of the chronic disease are, however, under debate (Feder et al., 2007).

When entering the mammalian host, BB is confronted first with a strong innate immune response. The first lines of defense include the attack of the spirochaetes by the complement system that opsonizes the spirochaetes and attracts, among other leukocytes, macrophages and neutrophils that have been shown to kill BB efficiently *in vitro* (Montgomery et al., 2002). When confronted with three times larger spirochaete concentrations in an *in vitro* assay, human polymorphonuclear cells and monocytes were both shown to ingest and kill the bacteria rapidly: after 60 min, 41–51% of the spirochaetes were phagocytized, and more than 50% of the ingested cells were dead within 30 min (Peterson et al., 1984). Additionally, mice and humans can produce antibodies against surface proteins of BB (Sadziene and Barbour, 1996).

The fact that despite the strong and effective early immune reaction, mammalian hosts frequently fail to clear the infection and that BB can often disseminate throughout the body remains striking. Even more surprising is the large population size that the bacteria reach at the infection site after an initial immune response has almost cleared the infection (see **Table 1**; Åsbrink and Hovmark, 2009). Given that dissemination of the bacteria has been shown as early as 2 days after experimental infection in mice (Barthold et al., 1991), active emigration from the infection site might aid the bacteria in escaping the host's early immune response. Furthermore, BB has been shown to actively alter the expression of several outer surface proteins in reaction to immune responses of the host (de Silva and Fikrig, 1997; Hodzic et al., 2003).

In this study, a deterministic model is developed to understand the interactions between BB and its mammalian host in the first stages of the infection. The model is used to elucidate how the bacteria evade the immune response, i.e., whether active emigration from the infection site, limited phagocytic capacity, or molecular adaptations of BB are critical for surviving the first line of defense by the innate immune response and to allow for a second growth phase reaching even higher bacterial loads. In this context, the importance of the adaptive immune system for controlling the infection is investigated.

Table 1 | Spirochaete counts per 10^6 mouse tissue cells as determined by qPCR (Pahl et al., 1999).

Time p.i.	Sample no.	Infected foot	Contralateral foot
1 h	1	12,592	0
	2	751	0
3 h	1	16,306	0
	2	12,265	0
6 h	1	11,999	0
	2	ND	0
12 h	1	287	0
	2	ND	0
24 h	1	415	0
	2	642	0
72 h	1	880	0
	2	ND	513
8 days	1	28,000	0
	2	146,067	0
15 days	1	6,667	343
	2	5,429	13,889
55 days	1	5,278	2,200
	2	8,644	159

Only the development in C3H/HeJ mice is shown here; spirochaete populations in BALB/c mice develop similarly (Pahl et al., 1999).

RESULTS

Our model is based on a study with mice infected experimentally with BB by Pahl et al. (1999). The spirochaete burden was measured using a quantitative polymerase chain reaction (qPCR) over the infection course. To analyze the kinetics of dissemination, two different mouse strains, disease-resistant BALB/c mice and susceptible C3H/HeJ mice were infected with BB and the spirochaete load was quantified at different tissues. Although the concentration of spirochaetes reached in these experiments differed between the two strains, the overall pattern was similar between the two strains. This modeling approach focuses on bacterial dynamics in C3H/HeJ mice, as they develop symptoms resembling Lyme disease in humans and are commonly used as a model organism for Lyme disease. **Table 1** summarizes spirochaete counts per 10^6 mouse cells as measured by Pahl et al. Out of the six tissue sites analyzed experimentally, we consider the infection site at the hind foot and the corresponding tissue site at the contralateral hind foot that serves as a measure of disseminated infection.

The bacterial populations in C3H/HeJ mice show remarkable dynamics at the infection site: spirochaetes are present in high numbers at three and six hours post infection (p.i.), but the population diminishes drastically after 12 h and remains at low levels at the following days; however, 8 days p.i., the BB population not only recovers, but reaches very high concentrations far above the level measured shortly after infection. At the contralateral site, the bacterial population reaches its maximum later and without the high intermediate population size observed at the infection site. Although the bacterial populations in BALB/c mice reach only lower levels, the dynamics are qualitatively similar. These data can be interpreted as an immediate immune response that controls the early infection effectively, but for unknown reasons fails to clear the infection and is overwhelmed at later time points. In our mathematical model, we attempt to explain this failure to clear the infection by three different hypotheses: (1) the immune response might be effective at the infection site, but dissemination and later re-migration to the infection site could aid the spirochaetes in escaping from this immune response, (2) the immune response at the infection site could be effective only in the very first hours and bacteria might overcome it due to limitations of innate immune responses, and (3) molecular adaptations of BB might be necessary to evade the host's immune response.

HEMATOGENOUS MIGRATION OF BACTERIA

Our first model (depicted in **Figure 1**) attempts to explain the persistence of the bacteria at the infection site by migration of BB to different tissues. Hence, we consider two tissue sites potentially harboring spirochaetes and describe bacterial population dynamics and the immune response at each of these sites. The first tissue

site described by this model is the infection site, i.e., the site at the hind foot where the mouse has been infected, the second one is a site at the contralateral hind foot. The bacteria live in the extracellular space where they can actively move through the extracellular matrix. Additionally, we also include bacterial dynamics in the blood as a third compartment connecting the two sites.

In each of these three compartments, the spirochaetes replicate at a maximal growth rate β. They are assumed to grow according to a logistic growth model with a carrying capacity (per 10^6 mouse cells) K. Hence, the growth of a bacterial population B_i in each of the compartments can be formulated as

$$\beta B_i \left(1 - \frac{B_i}{K}\right) \tag{1}$$

The maximal growth rate of BB in *in vitro* assays can be estimated at $0.06\,\mathrm{h}^{-1}$ (Barbour, 1983). The maximum bacterial concentration has been estimated at $1–4{\cdot}10^8$ BB cells per milliliter liquid medium in the same study; however, since spirochaete densities are given per 10^6 mouse cells and it is difficult to relate the *in vitro* result to the *in vivo* situation in the tissue, we take the maximal bacterial density measured in the experiments by Pahl et al. (1999), approximately $1–5{\cdot}10^5$ cells per 10^6 mouse cells, as the bacterial carrying capacity K.

In human patients, hematogenous migration of the bacteria is associated with disease symptoms and dissemination to tissues (Wormser et al., 2005). Whereas other routes of dissemination like the lymph system or direct migration within connective tissues cannot be excluded, their significance for the disseminated infection remains unclear. Hence, it is assumed that BB primarily disseminates hematogenously to the contralateral site and that bacteria can migrate from each of the tissues to the blood and vice versa. These migration rates, ν_i and μ_i are difficult to measure experimentally and thus have to be estimated.

Phagocytic cells are considered as an important early immune response. They are assumed to directly remove the bacteria. However, since the capacity of a single cell to kill bacteria is limited, we approximate phagocytosis of bacteria by a function saturable in B_i:

$$\frac{\rho B_i}{C_\rho + B_i} \cdot M_i, \tag{2}$$

where ρ is the maximal rate of phagocytosis, C_ρ is the bacterial number where phagocytosis reaches 50% of its maximum and M_i is the number of phagocytes in the compartment.

Bacterial dynamics for the populations at the infection site (B_1), the contralateral hind foot (B_2) and the blood (B_B) are described by the ordinary differential equations

$$\frac{dB_1}{dt} = \beta B_1 \left(1 - \frac{B_1}{K}\right) - \nu_1 B_1 + \mu_1 B_B - \frac{\rho B_1}{C_\rho + B_1} \cdot M_1, \tag{3}$$

$$\frac{dB_2}{dt} = \beta B_2 \left(1 - \frac{B_2}{K}\right) - \nu_2 B_2 + \mu_2 B_B - \frac{\rho B_2}{C_\rho + B_2} \cdot M_2, \tag{4}$$

$$\frac{dB_B}{dt} = \beta B_B \left(1 - \frac{B_B}{K}\right) + \nu_1 B_1 + \nu_2 B_2 - (\mu_1 + \mu_2) B_B - \frac{\rho B_B}{C_\rho + B_B} \cdot M_B. \tag{5}$$

The dynamics of the mice' phagocytic cells must not be neglected in this model. First, the production and death of phagocytic cells have to be considered. Furthermore, since phagocytic cells encountering BB produce strong inflammatory signals (Hirschfeld et al., 1999), they are assumed to recruit more phagocytic cells to the infected tissue site. Production of phagocytes by the bone marrow is assumed to occur at a constant rate δ. In addition to this, an infection-induced leukocyte production that depends on and is saturable with the bacterial load in the tissues, can be denoted as

$$\frac{\sigma(B_1 + B_2)}{C_\sigma + (B_1 + B_2)}, \tag{6}$$

with σ describing the maximal rate of BB-induced leukocyte production. Both processes are assumed to increase the number of leukocytes in the blood in our model.

Similarly, there is a physiological migration of leukocytes from the blood to the tissues (ϕ) and an additional infection-induced recruitment to the tissues (ψ):

$$\frac{\psi B_i}{C_\psi + B_i} M_i M_B. \tag{7}$$

Note that this recruitment also depends on the phagocytic cells already present in the tissue, because they are assumed to be the main producers of chemokines, attracting further leukocytes, and pro-inflammatory signals like TNF-α. The average lifetime of all phagocytes is assumed to be limited by a death rate θ in all three compartments.

The dynamics of phagocytic cell populations at the infection site M_1, at the contralateral foot M_2, and in the blood M_B are described by the differential equations

$$\frac{dM_1}{dt} = \phi M_B + \frac{\psi B_1}{C_\psi + B_1} M_1 M_B - \theta M_1, \tag{8}$$

$$\frac{dM_2}{dt} = \phi M_B + \frac{\psi B_2}{C_\psi + B_2} M_2 M_B - \theta M_2, \tag{9}$$

$$\frac{dM_B}{dt} = \delta + \frac{\sigma(B_1 + B_2)}{C_\sigma + (B_1 + B_2)} - \theta M_B - 2\phi M_B - \frac{\psi B_1}{C_\psi + B_1} M_1 M_B - \frac{\psi B_2}{C_\psi + B_2} M_2 M_B. \tag{10}$$

Figure 1 gives an overview of this model including both phagocytosis and bacterial dynamics at all three sites.

Although the host's immune response to BB is limited by different mechanisms in this model and bacterial migration is included here, it fails to reproduce the bacterial dynamics observed at the infection site under biologically reasonable conditions. With this model, the immune control at the infection site can be overwhelmed due to bacterial migration after a good initial success in diminishing bacteria that leaves less than 20% of the initial bacterial population alive (**Figure 2**). In this simulation the bacterial load saturates on a high level. However, the intermediate peak of the number of bacteria as found in experiments could not be reproduced. The bacterial population always approaches the saturation level from below. Very high numbers of bacteria in the

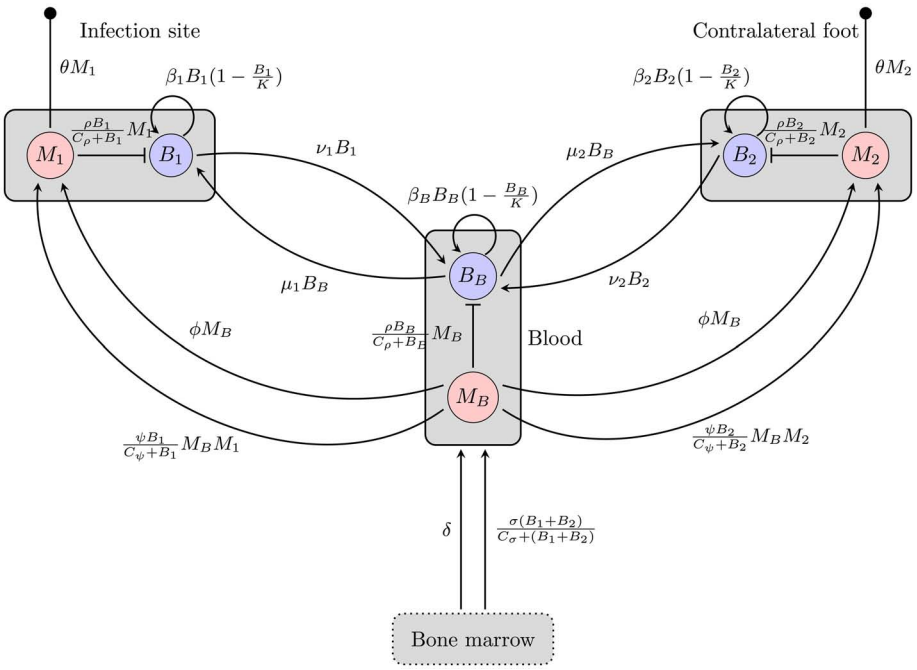


FIGURE 1 | Overview of the BB infection dynamics model based on the hematogenous migration of the bacteria. Boxes represent different compartments of the organism, circles are the considered quantities, and arrows/bars depict a stimulating/inhibiting influence of one quantity on another. The small filled circles represent death of the phagocytes. B, M: bacteria and phagocyte populations in the respective compartments, μ, ν: migration rates between the blood compartment and the two tissue sites, β: bacterial replication rates, K: carrying capacity for bacteria, ρ: phagocytosis rate, C_ρ: bacterial concentration with half-maximal phagocytosis rate, δ: physiological phagocyte production rate, σ: phagocyte production in response to infection, C_σ: bacterial concentration with half-maximal phagocyte production, ϕ: physiological phagocyte migration to tissue, ψ: phagocyte recruitment in response to infection, C_ψ: bacterial concentration with half-maximal phagocyte recruitment, θ: phagocyte death rate.

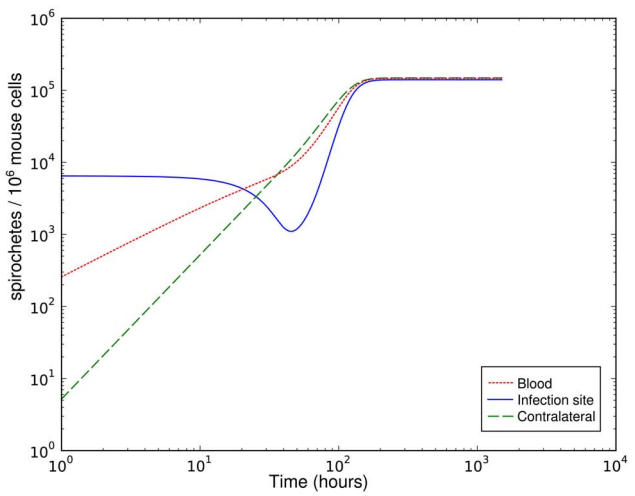


FIGURE 2 | Simulation of spirochetal dynamics at the infection site, the contralateral hind foot and in the blood with a parameter set leading to an effective early immune response that is overwhelmed due to bacterial migration later.

blood are required to achieve long-term persistence of bacteria at both sites, leading to a very high migration rate from the blood back into the tissue. Although BB can be cultured from the blood of many infected patients, this seems unlikely given that the blood was consistently culture-positive for BB only after day 10 p.i. in experimentally infected mice (Barthold et al., 1991). Within the concept of two sites and the blood compartment, no solution to this contradiction could be identified. The contradictions turned out to be a robust property of the model within physiologically relevant parameter values for bacterial replication and migration rates.

SINGLE-SITE MODEL WITHOUT MIGRATION AND LIMITED PHAGOCYTIC CAPACITY

Having dismissed bacterial migration as a likely reason for failure of early control at the infection site, an alternative explanation might be a limited killing capacity of the phagocytic cells: uptake of the bacteria might be faster than clearance, leading to fully occupied phagocytic cells that can take up no more bacteria. For this model, only the infection site was considered, because the timing of the infection rules out that the infection of the second infection site influences the immune control of the primary infection site. Early control of the infection at the infection site is crucial for the later outcome. If the site of the tick bite is excised during the first 48 h, dissemination of the bacteria is not observed (Shih et al., 1993). Hence, migration processes are ignored in this model. **Figure 3** depicts an overview of this model.

The average natural death of the phagocytic cells is again described as a rate θ. To describe the limitation of bacterial

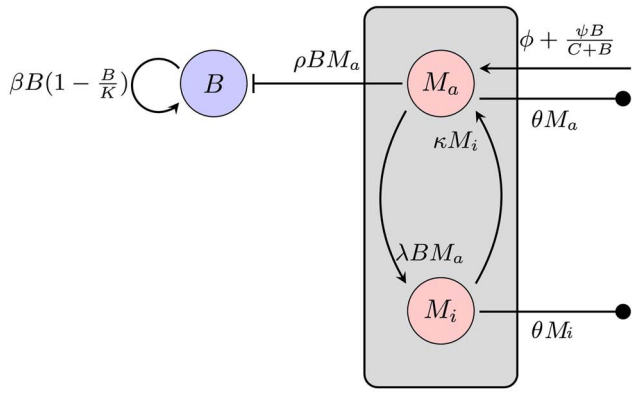


FIGURE 3 | Model of the bacterial dynamics at the infection site with limited phagocytic capacity of the immune cells. The phagocyte population is subdivided into an active M_a and an inactive M_i compartment. B: bacterial population at the infection site, β: bacterial replication rate, K: carrying capacity, ρ: phagocytosis rate, ϕ: physiological phagocyte migration into infection site, ψ: phagocyte recruitment as infection response, C: bacterial concentration yielding half-maximal recruitment of phagocytes, θ: phagocyte death rate, λ: inactivation rate of phagocytes, κ: rate of re-activation of phagocytes.

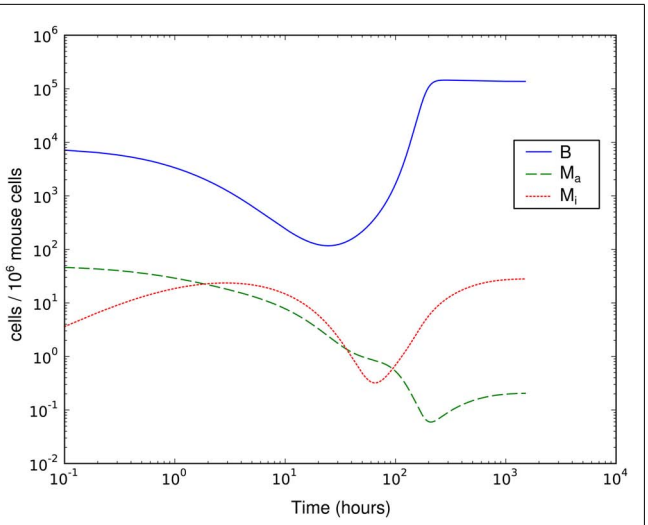


FIGURE 4 | Simulation of spirochetal dynamics in an infection at a single-site. In this model, the phagocytic capacity of the phagocytes limits the immune reaction and prevents control of the infection. B: bacteria population, M_a: active phagocytes, M_i: inactive phagocytes.

uptake, the population of phagocytic cells is subdivided into two compartments, one population M_a being active and able to phagocytose bacteria and the other one, M_i, in a saturated, inactive state, unable to take up more bacteria. The transition to the inactive state is modeled proportional to the number of bacteria at the infection site at a rate λ. Inactive phagocytic cells are allowed to return to the active state at a rate κ.

Again, bacterial replication is assumed to follow a logistic growth model with the same parameters used in the previous approach. However, since a limitation of the phagocytic capacity is modeled explicitly here, the phagocytosis rate is modeled as simply being proportional to the number of phagocytes and bacteria in the compartment:

$$\rho BM_a, \tag{11}$$

where M_a is the number of active phagocytes that are able to take up more bacteria. Hence, the bacterial dynamics can be described by the differential equation

$$\frac{dB}{dt} = \beta B\left(1 - \frac{B}{K}\right) - \rho BM_a. \tag{12}$$

Recruitment of leukocytes is also modeled similarly to the previous model. A low physiological, bacteria-independent migration ϕ into the infection site is combined with a bacteria-dependent log-sigmoidal recruitment

$$\frac{\psi B}{C+B}M, \tag{13}$$

with M being the sum of all macrophages. The phagocyte dynamics in the model can be described by the differential equations

$$\frac{dM_a}{dt} = \phi + \frac{\psi B}{C+B}(M_a + M_i) - \lambda BM_a + \kappa M_i - \theta M_a, \tag{14}$$

$$\frac{dM_i}{dt} = \lambda BM_a - \kappa M_i - \theta M_i. \tag{15}$$

Adjustment of the ratio of the phagocytosis rate to the transition to the inactive state, $\frac{\rho}{\lambda}$, regulates the phagocytic capacity of the immune cells. **Figure 3** gives an overview of this second model.

This much simpler model is able to explain the recovery of the bacterial population between day 3 and day 8 p.i. (**Figure 4**). However, it requires that the phagocytic cells are saturated quickly and return to the active state only very slowly – the phagocytes have to operate at the very limit of their uptake capability for several days. The clearance half-time of ingested BB is approximately 20 min (Montgomery and Malawista, 1996), making limited phagocytic capacity as an explanation for the failure to clear the early infection seem unlikely. In addition, as in the hematogenous migration model, with this model we could not identify any possibility to reproduce the measured second peak together with the second reduction of the bacterial load.

SINGLE-SITE MODEL WITH BACTERIAL ADAPTATION AND ADAPTIVE IMMUNE RESPONSE

In a third model, it is hypothesized that molecular adaptation of the spirochaetes is critical for BB to survive the first days to weeks of the infection. It is known that BB alters the expression of several outer surface proteins in response to immune reactions of its host and employs a variety of different immune evasion strategies (Tsao, 2009).

Considering the different mechanisms to hide from the host's immune mechanism and the critical role of the innate immune response early in BB infection, especially the fact that there are

adaptations to the innate immune response, it seems necessary to include this in the mathematical model. However, the impact and exact nature of these molecular adaptations is far less understood. It is difficult to estimate to what extent a molecular adaptation protects the bacteria from an immune response. We approximate this task by assuming that there are two different varieties of BB: one is infecting the mouse and is susceptible to the innate immune response (B_t), the second is resistant and is not directly phagocytosed (B_m; **Figure 5**).

In this model, we focus again on a single infection site. Bacterial growth is described by a logistic growth model using the same parameters as in the previous models. However, there are now two different BB compartments. The first, B_t, has not seen the host environment in the mouse and is thus not adapted to the mammalian immune response. This population is susceptible to phagocytosis, but adapts to the host environment at a rate α, not being recognized by the phagocytes after this adaptation. The host environment is constantly present as a stimulus inducing the transition from the susceptible to the resistant state. This is reflected in the model by an irreversible transition to the robust state. The dynamics in the two different BB compartments can be described by the following differential equations:

$$\frac{dB_t}{dt} = \beta B_t \left(1 - \frac{B_t + B_m}{K}\right) - \alpha B_t - \rho B_t M, \tag{16}$$

$$\frac{dB_m}{dt} = \beta B_m \left(1 - \frac{B_t + B_m}{K}\right) + \alpha B_t. \tag{17}$$

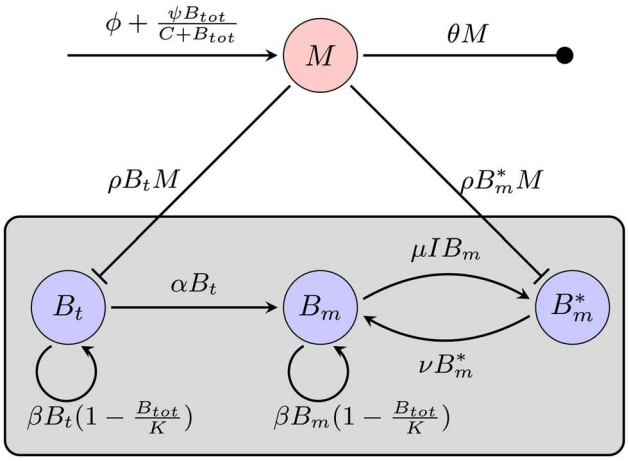


FIGURE 5 | Model of the bacterial dynamics at the infection site with molecular adaptations of the pathogen and a simple representation of the adaptive immunity. Vulnerable BB B_t enter the infection site and adapt at a rate α to a state where they are not recognized by phagocytes anymore. These adapted bacteria (B_m) can be bound by antibody I (not shown), making them susceptible again. B_t: phagocytosis-susceptible bacteria, B_m: adapted bacteria, B_m^* : antibody-bound bacteria, *beta*: bacterial replication rate, K: carrying capacity, ρ: phagocytosis rate, ϕ: physiological phagocyte migration into infection site, ψ: phagocyte recruitment as infection response, C: bacterial concentration yielding half-maximal recruitment of phagocytes, θ: phagocyte death rate, α: bacterial adaptation rate (transition to resistant state), μ: antibody binding rate, ν antibody dissociation rate, I antibody concentration.

Choosing α appropriately high, a complete transition from the susceptible to the robust state is achieved. However, with all bacteria now being resistant against the only immune response modeled so far, it is obvious that the infection cannot be controlled in this model. C3H mice develop antibody titers to BB as early as 7 days p.i. (Barthold et al., 1991). Opsonized bacteria can be seen as susceptible to phagocytosis again. We introduce a new quantity I into the model, reflecting the antibody concentration. As including antibody production and degradation is far beyond the scope of this model, we describe the antibody concentration simply as a time-dependent function reflecting the antibody titers measured in experimentally infected mice:

$$I(t) = \begin{cases} 0 & t \leq \Delta t_{ig} \\ \frac{I_{\max}(t-\Delta t_{ig})^n}{t_{1/2}^n + (t-\Delta t_{ig})^n} & t > \Delta t_{ig} \end{cases}, \tag{18}$$

where Δt_{ig} represents the time when the antibody production starts, $t_{1/2}$ denotes the time when the antibody concentration reaches its half-maximal value and $I_{\max}$ is the maximal antibody concentration reached. As this is a phenomenological description, parameters have to be estimated based on experimental data. The antibody production term is based on IgM titers measured by Barthold et al. (1991). **Figure 6** shows the fit of the function $I(t)$ to IgM titers against BB normalized to 1. The measured IgM titers, in contrast to our fit, show a decline at the last measurement >400 h. However, since IgG antibodies at that time started to show significant concentrations in the same studies, we do not consider this decline in our model, as these antibodies complement the IgM response. To capture the adaptive immune response at this time more accurately, the model would have to be extended to reflect the difference in diffusibility and efficacy of the different immune strategies as well as the adaptations of the bacteria to the presence of IgG antibody, which is beyond the scope of the present study.

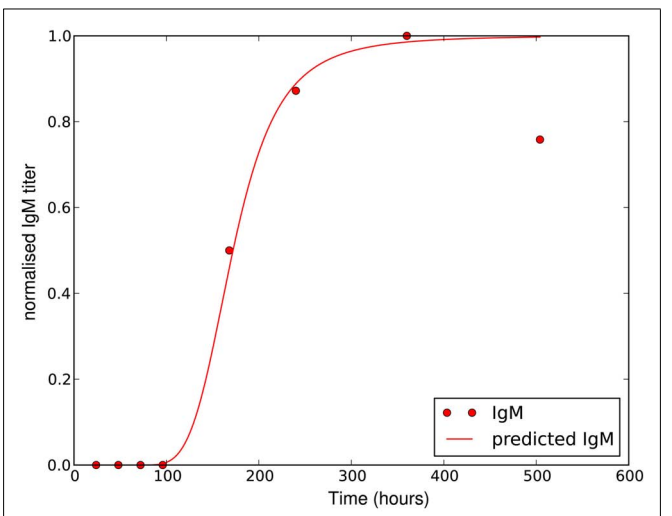


FIGURE 6 | Experimentally measured IgM titers (Barthold et al., 1991), normalized to 1, and the corresponding function given by equation 18 with $I_{max} = 1$, $\Delta t_{ig} = 72$, $t_{1/2} = 100$ and $n = 4$.

The maximal antibody concentration in the tissue is difficult to estimate, as the concentration in the serum is most likely higher than in the tissue. However, since to our knowledge, the antibody concentration in the tissue has not been determined experimentally and also varies between different mice and pathogen strains, we take the IgM serum concentration measured by Wooten et al. (2002) 2 weeks p.i. as approximation of the antibody concentration in the tissue.

The antibody can now bind reversibly to the phagocytosis-resistant spirochaetes B_m, leading to a third, opsonized bacterial population B_m^*. Binding and dissociation of the antibody can be described by

$$\frac{dB_m}{dt} = -\mu I B_m + \nu B_m^*, \tag{19}$$

$$\frac{dB_m^*}{dt} = \mu I B_m - \nu B_m^*, \tag{20}$$

where μ denotes the binding rate and ν the corresponding dissociation rate. Note that the antibody consumption by the binding process is ignored in this approach. Combining equations 17 and 19/20 and considering the antibody-bound bacteria B_m^* to be susceptible to phagocytosis, we can express the bacterial dynamics at the infection site as

$$\frac{dB_m}{dt} = \beta B_m \left(1 - \frac{B_t + B_m}{K}\right) + \alpha B_t - \mu I B_m + \nu B_m^*, \tag{21}$$

$$\frac{dB_m^*}{dt} = \mu I B_m - \nu B_m^* - \rho B_m^* M. \tag{22}$$

Since we do not consider the limited phagocytic capacity to be critical for the survival of the pathogen, the distinction between active and saturated phagocytes is obsolete. Recruitment and death of phagocytes are modeled as before. The differential equation describing their dynamics is thus straightforward:

$$\frac{dM}{dt} = \phi + \frac{\psi B}{C + B} M - \theta M. \tag{23}$$

Numerical simulation of the system shows that the model is able to reproduce the experimental data (**Figure 7A**). The parameter set used for this simulation is shown in **Table 2**. The three individual sub-populations of BB show a quick transition from the initial, vulnerable state to the "phagocytosis-resistant" state (**Figure 7B**). Already 1 day p.i., most bacteria have adapted to the immune response in our model; 2 days p.i., the initial, vulnerable population is almost extinct. This fast adaptation is achieved by regulating the parameter α and seems reasonable given the fact that *in vivo*, already 2 days p.i. BB has completely downregulated the tick-stage specific *ospA* gene (Hodzic et al., 2002). Although the leukocyte recruitment is limited in the model, this limitation alone could not rescue the bacterial population, as relatively low phagocyte numbers in the beginning are sufficient to drastically reduce the bacterial population size (**Figure 8**).

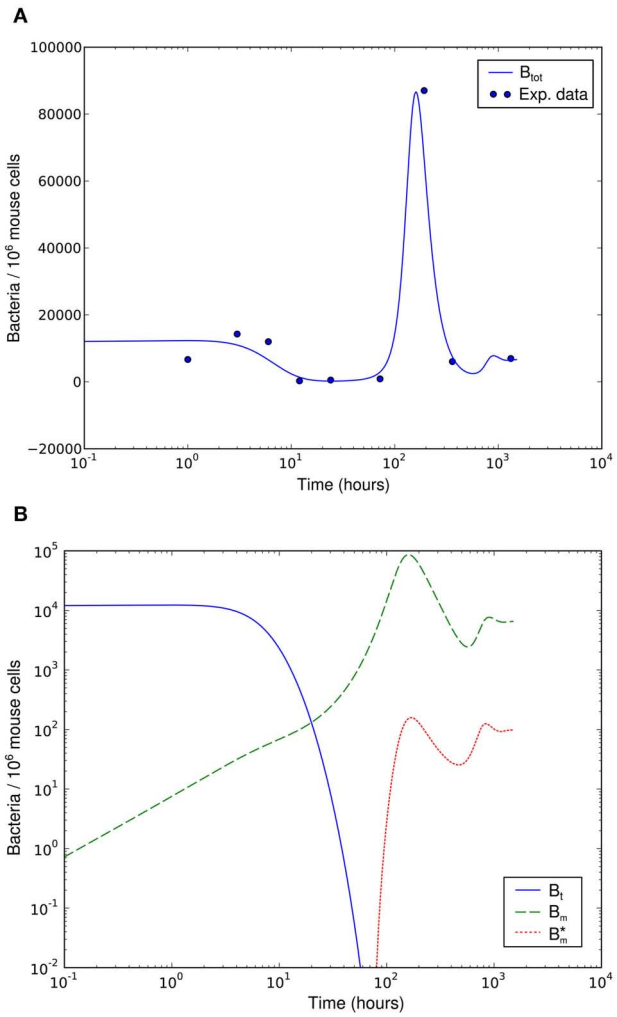


FIGURE 7 | Numerical simulation of the BB infection model with adaptation of the spirochaetes and humoral immune response. (A) Total number of spirochaetes per 10^6 mouse cells in experiment (Pahl et al., 1999) and simulation. The parameters used are described in **Table 2**. The simulation result is shown as the sum of all three BB sub-populations. **(B)** *Borrelia* sub-populations. B_t is the initial, vulnerable BB population, B_m the adapted, resistant one, and B_m^* antibody-bound spirochaetes.

DISCUSSION

Our three modeling approaches presented here try to give three different explanations for the survival of BB in the first days of the immune response. All three models can explain this, but in the first two approaches, this is only possible in a narrow range of parameters that is biologically unlikely, e.g., with extremely high bacterial replication and migration rates or with a fast saturation of the phagocytes (i.e., low phagocytic capacity) and a slow digestion rate within the phagocytic cells. The high population size that the spirochaetes reach at the infection site after the infection was almost cleared is striking and can only be explained by the third model presented here, which includes bacterial adaptations to the immune response of the host.

The first model tries to explain the failure to clear the infection by spirochetal escape to other tissues in the body. The bacteria might escape to other tissues, replicate there and then migrate back to the blood and from there into the infection site. However,

Table 2 | Parameters used for the numerical simulations in Figure 7A.

Parameter	Unit	Value	Derived from	Description
β	h^{-1}	0.06	Barbour (1983)	Bacterial growth rate
K	cells	150,000	Max. value from Pahl et al. (1999)	Bacterial carrying capacity
α	h^{-1}	0.0006	Estimate	Bacterial adaptation rate (transition to resistant state)
ρ	cells h^{-1}	0.005	Estimate	Phagocytosis rate
I_{max}	$\mu g\,ml^{-1}$	8.5	Wooten et al. (2002)	Antibody concentration
Δt	h	72	Barthold et al., 1991; phenomenology)	Time until antibody response starts
$t_{1/2}$	h	100	Barthold et al., 1991; phenomenology)	Time until antibody response reaches half-maximal value
n	–	4	Barthold et al., 1991; phenomenology)	Hill coefficient in antibody response function
μ	ml μg^{-1} h^{-1}	0.0085	Estimate	Antibody binding rate
ν	h^{-1}	1	Estimate	Antibody dissociation rate
ϕ	cells h^{-1}	0.1	Estimate	Physiological phagocyte migration to tissue
ψ	h^{-1}	0.001	Estimate	Phagocyte recruitment (infection response)
C	cells	1000	Estimate	Bacterial population size that yields half-maximal phagocyte recruitment
θ	h^{-1}	0.0076	van Furth and Diesselhoff-den Dulk (1970)	Phagocyte death rate

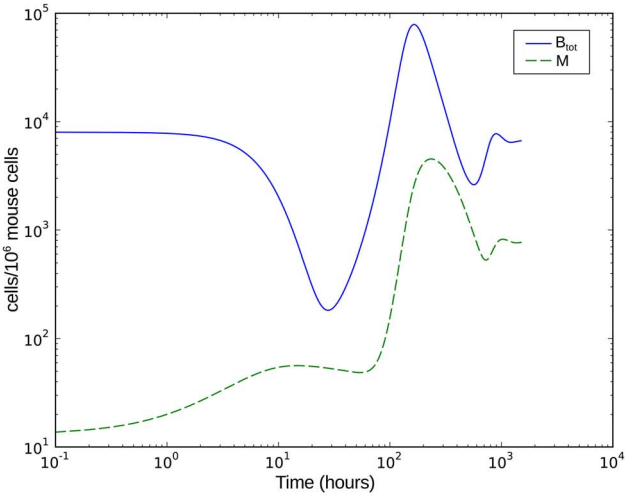


FIGURE 8 | Numerical simulation of the BB infection model with adaptation of the spirochaetes and humoral immune response. B_{tot} is the total bacterial population size, M the number of phagocytic cells.

this process would have to be very fast to explain the dynamics measured at the infection site: to overcome the immune response, a high bacterial concentration in the blood is required. A large amount of bacteria has to migrate to the infection site to lead to a high concentration of bacteria that is in excess of the number of bacteria that can be digested by the phagocytes. Considering previous studies on experimentally infected mice of the same strain (Barthold et al., 1991), this seems unlikely. Blood samples of these mice were not consistently tested positive on BB before day 10 p.i., indicating that the average spirochaete load in the blood was probably not very high in the first days. This is confirmed by later results where culture-positive on BB were only found 2 weeks p.i. (Hodzic et al., 2003) and even at this time only in one out of five tested mice. One factor that contributes to this fact is that there is a strong dilution effect for the bacteria in the blood: the spatially limited tissue area can yield a high concentration of bacteria. A fraction of these bacteria now migrates through the endothelial barriers into the blood stream, where it is dispersed in a much larger volume. Furthermore, although bacteria can move in the extracellular space, migration through endothelial barriers is an active process involving bacterial movement through dense tissues and binding of host factors, likely resulting in a slow process. In contrast, at the infection site, the bacterial population recovers at earlier times from the innate immune response. It seems thus unlikely, that the recovery of the spirochetal populations from the initial immune response at the infection site results from the re-immigration of bacteria from the blood.

The second model neglects migration processes of the bacteria and tries to explain the failure of the immune response by a limited phagocytic capacity. Phagocytes might take up bacteria, but fail to digest them fast enough. The phagocytic cells would engulf many bacteria, but stop to phagocytose more due to physical restrictions, becoming inactive until the bacteria are digested. In our model, this is expressed by introducing a population of inactive phagocytic cells. However, we could not reproduce the dynamics measured at the infection site using this model. It is possible to achieve a recovery of the bacterial population in the model, but this requires a large fraction of the phagocytic cells to be in an inactive state. The short clearance time of BB in macrophages (Peterson et al., 1984) and the generally high capacity of phagocytic cells to engulf particles (Cannon and Swanson, 1992) makes this assumption unlikely from a biological point of view. In addition, the final reduction of the bacterial population at day 7 p.i. could not be reproduced by this model.

The third model does not rely on phagocytes reaching the limit of bacteria uptake. Instead, a very simple phagocytosis model is employed; phagocytosis is limited only by the amount of available phagocytic cells in this model which is determined by the recruitment of cells. There is solid evidence for a number of different adaptations of BB (Zhang and Norris, 1998; Hodzic et al., 2003; Grimm et al., 2004; Palmer et al., 2009; Tsao, 2009) to different

factors of the host environment, including immune responses. A classical example of these adaptations is the reciprocal regulation of the surface proteins OspA and OspC. The former is a tick specific protein that is exclusively expressed when the spirochaete is in the tick's midgut, but downregulated in the mammalian host environment. OspC is only expressed inside the mammalian host, and is crucial for dissemination and survival of the spirochaetes in the early stages. The downregulation of OspA has been shown to be suppressed in reaction to the host's immune response, which is also true for other surface proteins (Hodzic et al., 2003). These adaptations are included in the third modeling approach.

It was assumed in this third model that BB becomes completely resistant against phagocytosis by immune cells. This is a strong assumption and the situation is probably different *in vivo*, as the bacteria are recognized and attacked via different pathways (Schröder et al., 2008). However, this simplification allows to estimate the influence of a molecular adaptation to a host response without making further assumptions about different innate immune processes that are not completely resolved. The transition of the whole population to the resistant state happens fast. But, surprisingly, the influence of the rate α on the speed of this transition is low. The transition is facilitated by removing bacteria susceptible to phagocytosis and by this allowing the resistant population to replicate faster within the limits imposed by the carrying capacity K for bacteria in the logistic growth function.

After the peak on day 8 p.i., the bacterial population decreases again. This coincides with an early IgM antibody response that may start to control the infection at this time. The antibody modeled in this approach binds to the spirochaetes and is assumed to make them susceptible to phagocytosis again. The affinity of these antibodies has not been investigated to our knowledge. Also, antibody responses to BB have different targets, making it difficult to determine the antibody affinity that is represented by the ratio of the dissociation and binding rates of the antibody $\frac{\nu}{\mu}$ in our model.

As the dissociation constant widely determines the asymptotic bacterial load (**Figure 9**), these rates were estimated to fit the bacterial levels measured experimentally at 55 days p.i. The ratio $\frac{\nu}{\mu}$ used in our simulations, corresponding to a dissociation constant $K_D = 1.3 \cdot 10^{-7}$ M, can be considered to be in a realistic range of IgM affinities. For example, human IgM antibodies have been shown to bind to different human antigens with dissociation constants between 10^{-6} and 10^{-8} M (Diaw et al., 1997). The presented results are robust within this experimental range of dissociation constants. However, the resulting asymptotic bacterial load is quantitatively lifted up and down in dependence on its exact value (see **Figure 9**).

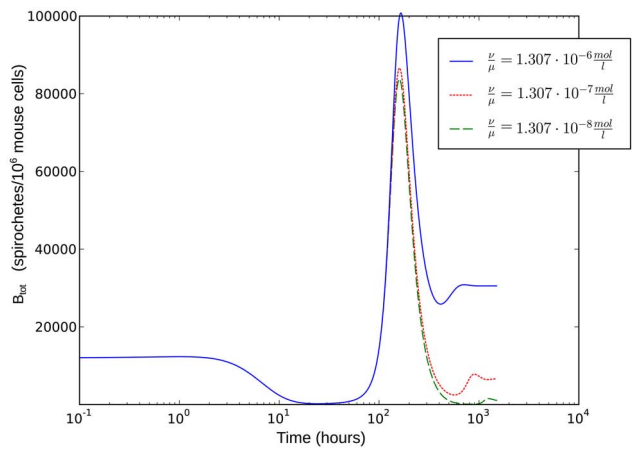


FIGURE 9 | Single-site model with bacterial adaptation and adaptive immune reaction. Simulations were done with different values for $\frac{\nu}{\mu}$ as shown in the figure.

To evaluate the robustness of this last model, single parameters like the phagocytosis rate ρ or reasonable combinations of parameters like $\frac{\nu}{\mu}$ were varied. The fact that there is a population of bacteria in our model that is not attacked by the phagocytes at all before the IgM response is visible makes the overall behavior of the system robust to strong variations in different parameters; the parameters ρ and ψ can, however, change the behavior qualitatively, since setting them arbitrarily low leads to uncontrolled replication of the bacteria in the early stages of the infection. For the later development, the antibodies' dissociation constant $\frac{\nu}{\mu}$ determines the bacterial population size in the late infection. This population size, however, cannot be interpreted as a real equilibrium, since the biological situation is more complicated due to different adaptive immune strategies, adaptations of the bacteria to these immune strategies and the presence of bacteria in immunoprivileged niches like the synoviae.

Our model is only intended to capture the first few weeks of the infection, specifically why the spirochaetes are not removed from the infection site. More sophisticated models of the specific immune response would be required to describe later stages and investigate the persistence of BB. For example, the IgG response that reaches its maximum significantly later would have to be included. Furthermore, for long-term persistence of the spirochaetes, the migration to different tissues might still be important even though the present analysis showed that its role is not important in the early stage of the infection. BB aligns with collagenic fibers in many tissues by binding to different host factors which might provide a protected niche for spirochetal survival (Steere and Glickstein, 2004; Rupprecht et al., 2008).

The model could also be extended to better describe the *in vivo* situation. Currently, the model is based on experimental data from experimentally infected mice, where both the effect of the tick environment before infection and the influence of the tick bite itself are missing, as the mice are infected intradermally with a syringe. In particular, immunosuppressive components of tick saliva can be expected to have an influence on the infection dynamics and control of the disease. Such effects are currently embedded into the phagocytosis rates without being explicit.

Our approach has shown that the representation of different mechanisms of bacteria-host-interaction in mathematical models allows to disentangle their relevance in different phases of a bacterial infection. In the case of BB, the measured dynamics of the bacterial load provides sufficient constraints to exclude specific mechanisms and to favor others. However, the relevance derived from the modeling is valid only in the early stage of the infection. Thus, this model for the BB infection dynamics can be used to

simulate and optimize therapeutical approaches in the first days after infection.

MATERIALS AND METHODS

Numerical simulations of the model were done using the LSODA solver for Ordinary Differential Equations implemented in the library deSolve for the R statistics language and in our own implementation of a fifth order Dormand-Prince method in C++.

Parameter values were derived from literature data where available and otherwise estimated in an iterative approach. First, the phagocytosis rate ρ, the phagocyte recruitment ψ and its half-maximal bacterial concentration, and the physiological phagocyte migration ϕ were approximated by fitting the decline of the total bacterial population size to the sharp initial decline measured experimentally for the first days. Second, the parameter α was adjusted to allow for recovery of the bacterial population according to measured data by a transition from the vulnerable to a resistant state of the bacteria. Setting α in this way also shows an adaptation of most of the bacterial population after the first 3 days p.i., reflecting the time frame in which reciprocal regulation of *ospA* and *ospC* as an example for molecular adaptations to a changed host environment can be found *in vivo* (Hodzic et al., 2003) and that is hence considered as a realistic time required for the majority of a bacterial population to adapt to a host's immune response. For this second step, the experimental data up to the second peak in the bacterial population size were taken into account. In the last step, the binding affinity of the antibody, $\frac{\mu}{\nu}$, was estimated to match the size that the bacterial population approaches at the end of the experimental measurements. The sum of squared errors was used as measure of discrepancy between data and predictions. When estimating parameters without prior knowledge about the exact values, parameters were set to physiologically reasonable values and the parameter space was scanned within a maximum of biologically plausible range; e.g., for the antibody dissociation constant, $\frac{\nu}{\mu}$, published values lie within 10^{-8} to 10^{-6} M, so $\frac{\nu}{\mu}$ was varied from 10^{-9} to 10^{-5} M and the fit optimized within this range. The parameters used for our final model are given in **Table 2** and explained in the text.

ACKNOWLEDGMENTS

Michael Meyer-Hermann was supported by HFSP and BMBF within the GerontoSys initiative (projects GerontoMitoSys and GerontoShield). Sebastian Binder was supported by the Helmholtz International Graduate School for Infection Research. This research was supported by a fellowship of the Volkswagen Foundation awarded to Arndt Telschow.

REFERENCES

Åsbrink, E., and Hovmark, A. (2009). Successful cultivation of spirochetes from skin lesions of patients with erythema chronicum migrans Afzelius and acrodermatitis chronica atrophicans. *Acta Pathol. Microbiol. Immunol. Scand. B* 93B, 161–163.

Barbour, A. G. (1983). Isolation and cultivation of Lyme disease spirochetes. *Yale J. Biol. Med.* 57, 521–525.

Barthold, S. W., Persing, D. H., Armstrong, A. L., and Peeples, R. A. (1991). Kinetics of *Borrelia burgdorferi* dissemination and evolution of disease after intradermal inoculation of mice. *Am. J. Pathol.* 139, 263–273.

Cannon, G. J., and Swanson, J. A. (1992). The macrophage capacity for phagocytosis. *J. Cell. Sci.* 101 (Pt 4), 907–913.

de Silva, A. M., and Fikrig, E. (1997). Arthropod- and host-specific gene expression by *Borrelia burgdorferi*. *J. Clin. Invest.* 99, 377–379.

Diaw, L., Magnac, C., Pritsch, O., Buckle, M., Alzari, P. M., and Dighiero, G. (1997). Structural and affinity studies of IgM polyreactive natural autoantibodies. *J. Immunol.* 158, 968.

Feder, H. M., Johnson, B. J. B., O'Connell, S., Shapiro, E. D., Steere, A. C., and Wormser, G. P. (2007). A critical appraisal of "chronic Lyme disease." *N. Engl. J. Med.* 357, 1422–30.

Grimm, D., Tilly, K., Byram, R., Stewart, P. E., Krum, J. G., Bueschel, D. M., Schwan, T. G., Policastro, P. F., Elias, A. F., and Rosa, P. A. (2004). Outer-surface protein C of the Lyme disease spirochete: a protein induced in ticks for infection of mammals. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3142–3147.

Hirschfeld, M., Kirschning, C. J., Schwandner, R., Wesche, H., Weis, J. H., Wooten, R. M., and Weis, J. J. (1999). Cutting edge: inflammatory signaling by *Borrelia burgdorferi* lipoproteins is mediated by toll-like receptor 2. *J. Immunol.* 163, 2382–2386.

Hodzic, E., Feng, S., Freet, K., Borjesson, D., and Barthold, S. (2002). *Borrelia burgdorferi* population kinetics and selected gene expression at the host-vector interface. *Infect. Immun.* 70, 3382.

Hodzic, E., Feng, S., Freet, K. J., and Barthold, S. W. (2003). *Borrelia burgdorferi* population dynamics and prototype gene expression during infection of immunocompetent and immunodeficient mice. *Infect. Immun.* 71, 5042–5055.

Hubalek, Z. (2009). Epidemiology of Lyme borreliosis. *Curr. Probl. Dermatol.* 37, 31–50.

Lane, R., Piesman, J., and Burgdorfer, W. (1991). Lyme borreliosis: relation of its causative agent to its vectors and hosts in North America and Europe. *Annu. Rev. Entomol.* 36, 587–609.

Montgomery, R. R., Lusitani, D., de Boisfleury Chevance, A., and Malawista, S. E. (2002). Human phagocytic cells in the early innate immune response to *Borrelia burgdorferi*. *J. Infect. Dis.* 185, 1773–1779.

Montgomery, R. R., and Malawista, S. E. (1996). Entry of *Borrelia burgdorferi* into macrophages is end-on and leads to degradation in lysosomes. *Infect. Immun.* 64, 2867–2872.

Nau, R., Christen, H.-J., and Eiffert, H. (2009). Lyme disease – current state of knowledge. *Dtsch. Arztebl. Int.* 106, 72–81; quiz 82, I.

Pahl, A., Kuehlbrandt, U., Brune, K., Roellinghoff, M., and Gessner, A. (1999). Quantitative detection of *Borrelia burgdorferi* by real-time PCR. *J. Clin. Microbiol.* 37, 1958–1963.

Palmer, G. H., Bankhead, T., and Lukehart, S. A. (2009). "Nothing is permanent but change" – antigenic variation in persistent bacterial pathogens. *Cell. Microbiol.* 11, 1697–1705.

Peterson, P. K., Clawson, C. C., Lee, D. A., Garlich, D. J., Quie, P. G., and Johnson, R. C. (1984). Human phagocyte interactions with the Lyme disease spirochete. *Infect. Immun.* 46, 608–611.

Rizzoli, A., Hauffe, H., Carpi, G., Vourc'h, G., Neteler, M., and Rosa, R. (2011). Lyme borreliosis in Europe. *Euro Surveill.* 16, 1–8.

Rupprecht, T. A., Koedel, U., Fingerle, V., and Pfister, H.-W. (2008). The pathogenesis of Lyme neuroborreliosis: from infection to inflammation. *Mol. Med.* 14, 205–212.

Sadziene, A., and Barbour, A. (1996). Experimental immunization against Lyme borreliosis with recombinant Osp proteins: an overview. *Infection* 24, 195–202.

Schröder, N. W. J., Eckert, J., Stübs, G., and Schumann, R. R. (2008). Immune responses induced by spirochetal outer membrane lipoproteins and glycolipids. *Immunobiology* 213, 329–340.

Shih, C.-M., Telford, S. R., Pollack, R. J., and Spielman, A. (1993). Rapid dissemination by the agent of Lyme disease in hosts that permit fulminating infection. *Infect. Immun.* 61, 2396–2399.

Stanek, G., Fingerle, V., Hunfeld, K.-P., Jaulhac, B., Kaiser, R., Krause, A., Kristoferitsch, W., O'Connell, S., Ornstein, K., Strle, F., and Gray, J. (2011). Lyme borreliosis: clinical case definitions for diagnosis and management in Europe. *Clin. Microbiol. Infect.* 17, 69–79.

Steere, A. C., and Glickstein, L. (2004). Elucidation of Lyme arthritis. *Nat. Rev. Immunol.* 4, 143–152.

Steere, A. C., Klitz, W., Drouin, E. E., Falk, B. A., Kwok, W. W., Nepom, G. T., and Baxter-Lowe, L. A. (2006). Antibiotic-refractory Lyme arthritis is associated with HLA-DR molecules that bind a *Borrelia burgdorferi* peptide. *J. Exp. Med.* 203, 961–971.

Steere, A. C., Malawista, S. E., Snydman, D. R., Shope, R. E., Andiman, W. A., Ross, M. R., and Steele, F. M. (1977). An epidemic of oligoarticular arthritis in children and adults in three Connecticut communities. *Arthritis Rheum.* 20, 7–17.

Tilly, K., Rosa, P. A., and Stewart, P. E. (2008). Biology of infection with *Borrelia burgdorferi*. *Infect. Dis. Clin. North Am.* 22, 217–234.

Tsao, J. I. (2009). Reviewing molecular adaptations of Lyme borreliosis spirochetes in the context of reproductive fitness in natural transmission cycles. *Vet. Res.* 40, 36.

van Furth, R., and Diesselhoff-den Dulk, M. (1970). The kinetics of promonocytes and monocytes in the bone marrow. *J. Exp. Med.* 132, 813.

Wooten, R. M., Ma, Y., Yoder, R. A., Brown, J. P., Weis, J. H., Zachary, J. F., Kirschning, C. J., and Weis, J. J. (2002). Toll-like receptor 2 is required for innate, but not acquired, host defense to *Borrelia burgdorferi*. *J. Immunol.* 168, 348–355.

Wormser, G., McKenna, D., Carlin, J., Nadelman, R., Cavaliere, L., Holmgren, D., Byrne, D., and Nowakowski, J. (2005). Brief communication: hematogenous dissemination in early Lyme disease. *Ann. Intern. Med.* 142, 751–755.

Zhang, J.-R., and Norris, S. J. (1998). Genetic variation of the *Borrelia burgdorferi* gene vlsE involves cassette-specific, segmental gene conversion. *Infect. Immun.* 66, 3698–3704.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 December 2011; accepted: 01 March 2012; published online: 22 March 2012.

Citation: Binder SC, Telschow A and Meyer-Hermann M (2012) Population dynamics of Borrelia burgdorferi in Lyme disease. Front. Microbio. ***3****:104. doi: 10.3389/fmicb.2012.00104*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 26 April 2012
doi: 10.3389/fmicb.2012.00129

Agent-based modeling approach of immune defense against spores of opportunistic human pathogenic fungi

Christian Tokarski[1]*, Sabine Hummert[1,2], Franziska Mech[2], Marc Thilo Figge[2], Sebastian Germerodt[1], Anja Schroeter[1] and Stefan Schuster[1]

[1] Department of Bioinformatics, Friedrich Schiller University, Jena, Germany
[2] Research Group Applied Systems Biology, Leibniz Institute for Natural Product Research and Infection Biology–Hans-Knoell-Institute, Jena, Germany

Edited by:
Jörg Linde, Leibniz-Institute for Natural Product Research and Infection Biology–Hans-Knoell-Institute, Germany

Reviewed by:
Hubertus Haas, Innsbruck Medical University, Austria
Dirk Dittmer, University of North Carolina at Chapel Hill, USA
Guoku Hu, Creighton University, USA

***Correspondence:**
Christian Tokarski, Department of Bioinformatics, Friedrich Schiller University, Ernst – Abbe – Platz 2, 07743 Jena, Germany.
e-mail: christian.tokarski@uni-jena.de

Opportunistic human pathogenic fungi like the ubiquitous fungus *Aspergillus fumigatus* are a major threat to immunocompromised patients. An impaired immune system renders the body vulnerable to invasive mycoses that often lead to the death of the patient. While the number of immunocompromised patients is rising with medical progress, the process, and dynamics of defense against invaded and ready to germinate fungal conidia are still insufficiently understood. Besides macrophages, neutrophil granulocytes form an important line of defense in that they clear conidia. Live imaging shows the interaction of those phagocytes and conidia as a dynamic process of touching, dragging, and phagocytosis. To unravel strategies of phagocytes on the hunt for conidia an agent-based modeling approach is used, implemented in NetLogo. Different modes of movement of phagocytes are tested regarding their clearing efficiency: random walk, short-term persistence in their recent direction, chemotaxis of chemokines excreted by conidia, and communication between phagocytes. While the short-term persistence hunting strategy turned out to be superior to the simple random walk, following a gradient of chemokines released by conidial agents is even better. The advantage of communication between neutrophilic agents showed a strong dependency on the spatial scale of the focused area and the distribution of the pathogens.

Keywords: agent-based modeling, individual-based modeling, host-pathogen interaction, immune defense, opportunistic pathogenic fungi, chemotaxis, video analysis of life cell imaging

1. INTRODUCTION

The immune system of healthy humans successfully defends them against ubiquitous fungi like *Aspergillus fumigatus*, *Candida albicans*, and *Cryptococcus neoformans*. A weakened immune defense, due to diseases like infection with human immunodeficiency virus (HIV) or in the course of medical treatment, is compromised in the ability to repel those opportunistic fungal pathogens (Richardson, 2005; Karkowska-Kuleta et al., 2009). The recent medical progress especially in immune therapy, transplantation therapy, and life-prolonging measures leads to a growing number of susceptible patients (Richardson, 2005; Karkowska-Kuleta et al., 2009). Invasive mycoses caused by *A. fumigatus* lead to mortality rates of 60–90% (Karkowska-Kuleta et al., 2009). Thus, understanding infection dynamics and the response of the immune system to invading fungi is an essential step on the way to stop the opportunistic pathogens from taking over.

Inhaled conidia of *A. fumigatus* that are not repelled by respiratory tract mucociliary defenses lead to a complex response of the immune system (Shoham and Levitz, 2005). Here, we focus on the role of neutrophil granulocytes (neutrophils). These phagocytes are recruited by chemokines like IL-8 to the site of infection (Shoham and Levitz, 2005). In case of getting in contact with fungal conidia, phagocytes can engulf, and degrade them. The recruitment of neutrophils occurs within 4–8 h of intratracheal conidial infection, the same time *A. fumigatus* needs for conidial germination and hyphal formation (Hohl, 2009). Hyphae can invade pulmonary tissues, enter the bloodstream, and disseminate to remote tissues, leading to life-threatening systemic infections. Thus, clearance of fungal conidia should occur rapidly. To elucidate the dynamics of the interaction of *A. fumigatus* conidia and neutrophils live imaging data have been recorded and analyzed (Behnsen et al., 2007).

The goal of the present paper is to make a first step in analyzing these data by computer simulations. Computational immunology as a method to complement wet lab immunology comprises several mathematical methods (for a review see reference Forrest and Beauchemin, 2007). While game-theoretic models of immunological aspects are focused on states of equilibrium (Hummert et al., 2010), agent-based models (ABMs) are well-suited to model spatially heterogeneous, dynamic processes like interactions of immune cells with conidia. The main feature of ABMs is that a complex system's behavior is modeled by actions and interactions of entities called agents which themselves behave according to rules.

In theoretical immunology, models are composed of agents belonging to immune defense, e.g., phagocytes, as well as agents representing evading pathogens, e.g., fungal or bacterial cells. Behavioral rules of agents are based on experimental findings complemented by hypothetical properties. Thus, the impact of hypothetical properties on the system's behavior can be tested,

predictions can be made and for validation, the system's behavior *in silico* can be compared with the experimentally observed one. For reviews on ABMs in theoretical immunology see references Bauer et al. (2009), Chavali et al. (2008).

We implemented an ABM in NetLogo (Wilensky, 1999) to test the impact of different hunting strategies of neutrophils on their phagocytosis efficiency. A central question tackled in this work is whether chemical communication and chemotaxis of neutrophils improve the clearing efficiency. Germinating conidia of *A. fumigatus* activate the complement system and induce neutrophil chemotaxis (Waldorf and Diamond, 1985). Further, it is known that neutrophils have the ability to recruit other neutrophils (Scapini et al., 2000).

2. MATERIALS AND METHODS

2.1. VIDEO ANALYSIS OF LIFE CELL IMAGING

There exist well established techniques to visualize fungal infections in hosts, such as confocal laser scanning microscopy (Paddock, 1999) and bright-field microscopy (Pluta, 1988). The latter was used in this approach. The majority of produced data are images with immune cells and conidia, either as bright-field or fluorescence images. Important quantities that characterize the dynamic cellular behavior are the cell classification, cell counting, shape and extent measurements, cell-cell interactions, and, for time-lapse data, cell motility, and velocity. However, the image analysis forms the bottleneck in these studies. Automation of this process is the goal of biological image analysis. This was already realized for the two-dimensional case of phagocytosis assays (Mech et al., 2011). However, a specific application to temporally resolved phagocytosis assays of *A. fumigatus* conidia and phagocytes in a fully automated fashion is not available, today.

We analyzed video material (15 sequences) of the first 90 min of an *in vitro* interaction of neutrophils and conidia of *A. fumigatus* (videos have been provided by the research group of Matthias Gunzer). Every 30 s a frame was recorded which corresponds to 180 frames for sequences of 90 min length.

The analysis comprises three steps: segmentation, classification, and tracking. The segmentation of frames was done for each frame separately similar to the approach used in (Mech et al., 2011). Erroneous segmentation results due to highly heterogeneous cell morphologies and uneven illumination artifacts were corrected manually. With that, each cell was recognized as an object. Afterwards, neutrophilic objects were semi-automatically tracked. Each neutrophilic object was tracked from frame to frame using overlap in two consecutive frames or, if no overlap existed, by searching for the nearest neutrophilic object in the subsequent frame. During the tracking an unique identifier was assigned to each neutrophilic object as well as the changes in positions as velocities. Then, the number of conidia which are freely moving, ingested, or dragged were counted manually for each time step. The numbers of dragged and ingested conidia were assigned to the respective neutrophils. Thus, the phagocytosis rate of a neutrophil is the number of conidia which a neutrophil ingested over time. Furthermore, adherence of conidia may lead over time to ingestion by or dissociation from neutrophils. In the former case respecting conidia were categorized as dragged. Hence, the dragging rate depicts the number of adherent conidia which did not dissociate from a neutrophil.

After the time-lapse data analysis the number of neutrophils, the velocities, and number of associated and ingested conidia per neutrophil as well as the number of conidia categorized as free, ingested, or dragged, were available.

2.2. THE MODEL

The "hunt" of neutrophils for conidia is modeled by an ABM. The behavior of neutrophilic agents as well as conidial agents is defined by rules which is a key element of ABMs. Here, the rules are based on properties extracted from the given video material of life cell imaging (see subsection 2.1.) as described in subsequent sections. We follow the ODD-protocol for describing ABMs (Grimm et al., 2006, 2010).

2.3. PURPOSE OF THE ABM

Clearing conidia before they germinate can be seen as an action enhancing the fitness of the organism. Different strategies of neutrophils in tracking and disposing conidia vary in their efficiency. In this ABM approach we tested the impact of different tracking modes of neutrophilic agents on their efficiency in reducing the amount of free conidial agents. Simple random walk as well as a short-term persistence in keeping direction, both without sensing of chemotactic molecules were tested. Then, perception of excreted metabolites of germinating conidia or signal-molecules from complement system, thus all chemokines, was considered as well as positive feedback activation via chemical communication between neutrophilic agents. In a large-scale modeling approach representing a part of lung tissue we additionally tested the efficiency of chemotactic communication between neutrophilic agents for clustered distributions of conidial agents simulating an infection scenario.

2.4. ENTITIES, STATE VARIABLES, AND SCALES

The ABM was initialized on a grid of 41×35 (=1,435) discrete cells. The width of a grid cell fits the size of a conidium which is about 3 μm in diameter. The evaluation of the life imaging videos showed immigration and emigration of the motile cells out of the scope. For the model we assumed an equal rate of this events and therefore implemented periodical boundary conditions. One time-step in the simulation refers to a 30 s interval and matches the 2 frames per min pattern of the evaluated life imaging videos. Simulations run over the first 90 min of neutrophil-conidia interactions.

For a large-scale simulation a grid of 401×341 (=136,741) cells was built. Thus, an area of more than 1 mm^2 was covered. The simulation was evaluated for 2,880 simulated time-steps, which corresponds to 24 h after the initial infection. The model comprises three types of entities: grid cells, conidial agents, and motile neutrophilic agents. See **Tables A1–A3** in Appendix for state variables and parameters.

2.5. DESIGN CONCEPTS

We tested and combined several modes of neutrophilic movement to explain the *in vitro* efficiency of a population of neutrophils in reducing the amount of free conidia. In the model approach that incorporates diffusion of chemokines, excreted by the conidial agents, neutrophilic agents adapt their movement to this signal.

This assumption is supported by the observation of Waldorf and Diamond (1985) that germinating conidia of *A. fumigatus* activate the complement system and induce neutrophil chemotaxis. Scapini et al. (2000) report the ability of neutrophils to recruit other neutrophils. Thus, as a further hunting strategy communication between neutrophilic agents that orientate on gradients of signals was tested.

In our model we do not focus on the fitness functions of single agents, but on the immune system as a whole, where the population of neutrophils form a first line defense of a multi-cellular organism. Here, the efficient reduction of free conidia, which potentially germinate, infect the organism, and may decrease its overall fitness, can be seen as the fitness function. Costs of different hunting strategies have not been taken into account. None of the agents did change their behavior due to any kind of increasing experience or learning effects. In the model approaches with diffusion of chemokines excreted by conidial agents and chemical communication between neutrophilic agents, latter were able to sense and follow the gradients of these substances. Prediction: neutrophilic agents adjusted their movement in direction of the highest amount of these substances on the eight neighboring grid cells (Moore neighborhood) which indirectly increased the probability to encounter either a conidial agent (see sub-model diffusion of chemokines in section 2.8.1.) or another neutrophilic agent, which was activated by at least one dragged conidial agent (see sub-model communication between neutrophils in section 2.8.1.) – which can be seen as a positive feedback loop. The interaction of neutrophilic and conidial agents was separated into two major types: dragging of conidial agents and intake of them, corresponding to the video material (see sub-model interaction of agents in section 2.8.2. and **Figure A1** in Appendix). Stochasticity: Each simulation started with a random distribution of neutrophilic and conidial agents on the grid. The direction of movement of neutrophilic agents was chosen randomly for the random walk model. The velocities of neutrophilic agents were chosen either to form a normal distribution or a log-normal distribution, which was derived from life cell imaging data. We stored the fraction of free conidial agents, neither dragged or phagocytized by neutrophilic agents, for a fixed set of simulated time-steps, which correspond to the evaluated timeframes of the video material.

2.6. INITIALIZATION

Small-scale simulation: Each simulation was initialized with a random distribution and orientation of conidial agents and neutrophilic agents. Starting numbers of both agents (110 conidial agents, 25 neutrophilic agents) were taken from one of 15 given *in vitro* life cell imaging sequences (videos have been provided by the research group of Matthias Gunzer). All experimental samples contained between 68 and 127 conidia (mean: 100).

Large-scale simulation: 600 conidial agents were clustered in 4 randomly located spots on the grid. Cluster-size and number of conidial agents per cluster were equal for all clusters at initialization. 100 neutrophilic agents were randomly placed onto the grid.

2.7. INPUT DATA

From one of the videos we derived a density-distribution of the neutrophils' velocity. A histogram and the fitted log-normal density-distribution of the velocities is shown in **Figure 1**.

2.8. SUB-MODELS

2.8.1. Diffusion of chemokines by conidial agents and positive feedback activation of neutrophilic agents

Two types of chemical signaling determining the movement behavior of neutrophilic agents were considered in our modeling approach. We tested the impact of chemokines emitted by conidial agents which could be recognized by neutrophilic agents (**Figure 2**). Further, we allowed chemical communication between neutrophilic agents, assuming that the contact with a conidium can activate and recruit other immune cells (positive feedback; **Figure 2**). In our ABM the process of diffusion was expressed by two distinct parameters: chemokine-diffusion-rate and number of diffusion steps per time-step (repetition-of-chemokine-diffusion). Chemokine-diffusion-rate gives the fraction of molecules in the source grid cell that is equally divided and transferred to the eight neighboring cells (Moore neighborhood) in a single diffusion step. Both parameters determine the chemical gradient (see **Figure 3**).

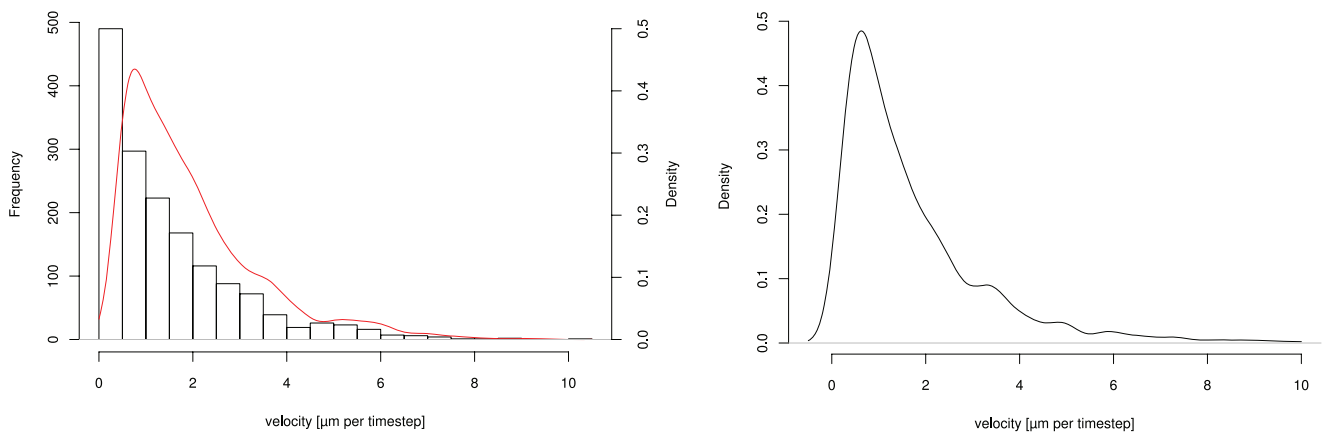


FIGURE 1 | Plots of (left) histogram and fitted density-distribution of neutrophils' velocity derived from life cell imaging (right) fitted continuous log-normal density-distribution function.

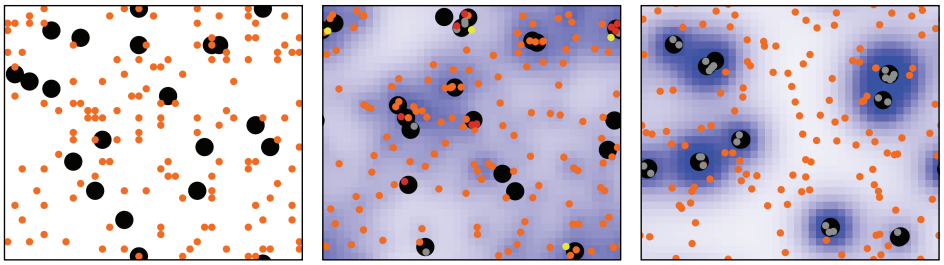


FIGURE 2 | ABM with (left) random walk of neutrophilic agents (middle) diffusion of chemokines excreted by conidial agents, and (right) with communication between neutrophilic agents, which causes a positive feedback loop in activating the immune defense and an aggregation of neutrophilic agents (see Appendix for sample videos of each movement type).

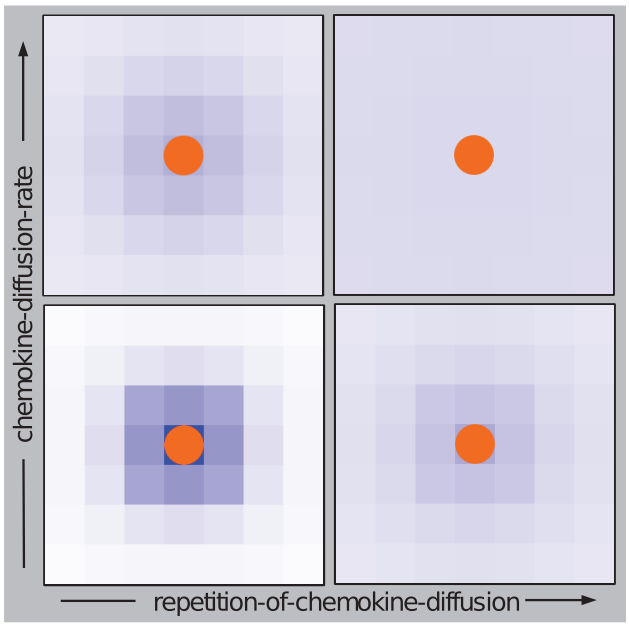


FIGURE 3 | Effects of the parameters "repetition-of-chemokine-diffusion" and "chemokine-diffusion-rate" on the process of diffusion of chemokines excreted by conidial agents on the grid cells for a fixed number of simulation-steps. A high value "repetition-of-chemokine-diffusion" leads to flat gradient of chemokines, while a high "chemokine-diffusion-rate" leads to a wider and faster diffusion of the chemical signal.

2.8.2. Interaction of agents

From life cell imaging we derived four main states of a conidial agent: it can either be free, dragged, phagocytized, or lysed by a neutrophilic agent (see **Figure A1** in Appendix). Free conidial agents can be dragged by neutrophilic agents with a certain probability. This linkage can vanish with a certain probability. Then, the conidial agent is free again, or the dragged conidial agent can be phagocytized with a certain probability. Once a conidial agent is ingested it gets digested by the neutrophilic agent after a fixed amount of time-steps (see **Figure 4**). The progress of a typical simulation run is shown in **Figure 5**.

3. RESULTS

For the interpretation of interactions of neutrophils with conidia of *A. fumigatus* recorded by *in vitro* life cell imaging we tested in an ABM several modes of movement of the neutrophilic agents in their hunt for conidial agents. The remaining amount of free conidial agents after a defined timespan was taken as a measure for the efficiency of the applied hunting strategy of neutrophilic agents. Different strategies were tested in a small-scale model covering about 10,000 μm^2 as well as on a larger scale representing about 1 mm^2 of lung tissue.

3.1. SMALL-SCALE SIMULATION

The simulation was conducted over the first 90 min of interactions, identical to the timespan of observation in life cell imaging experiments, and was repeated 250 times for each parameter combination.

3.1.1. Random walk, normally distributed velocity of neutrophilic agents

The simplest way for neutrophilic agents to search conidial agents while having no clue of their whereabouts is by searching randomly, i.e., random walk. The velocity of neutrophilic agents was assumed to be normally distributed. An increasing mean resulted in a significantly decreased amount of free conidial agents at the end of the simulations, while the impact of the SD was low (see **Figure 6**).

3.1.2. Neutrophilic agents keep direction for a short timespan (short-term persistence in direction)

Is it better to keep direction on the search and thus, to reduce the probability of moving back where the area is already cleared? Indeed, the efficiency of neutrophilic agents to phagocytize conidial agents increased significantly with a higher probability to hold the recent direction in the next time-step, which is called short-term persistence (STP; see **Figure 7**). The gain in clearing efficiency increases with the passing of time. The STP strategy pays off especially at low amounts of remaining free conidial agents.

3.1.3. Diffusion of chemokines

Does following traces of the presence of conidia lead to a higher clearing efficiency? Metabolic activities of germinating conidia as well

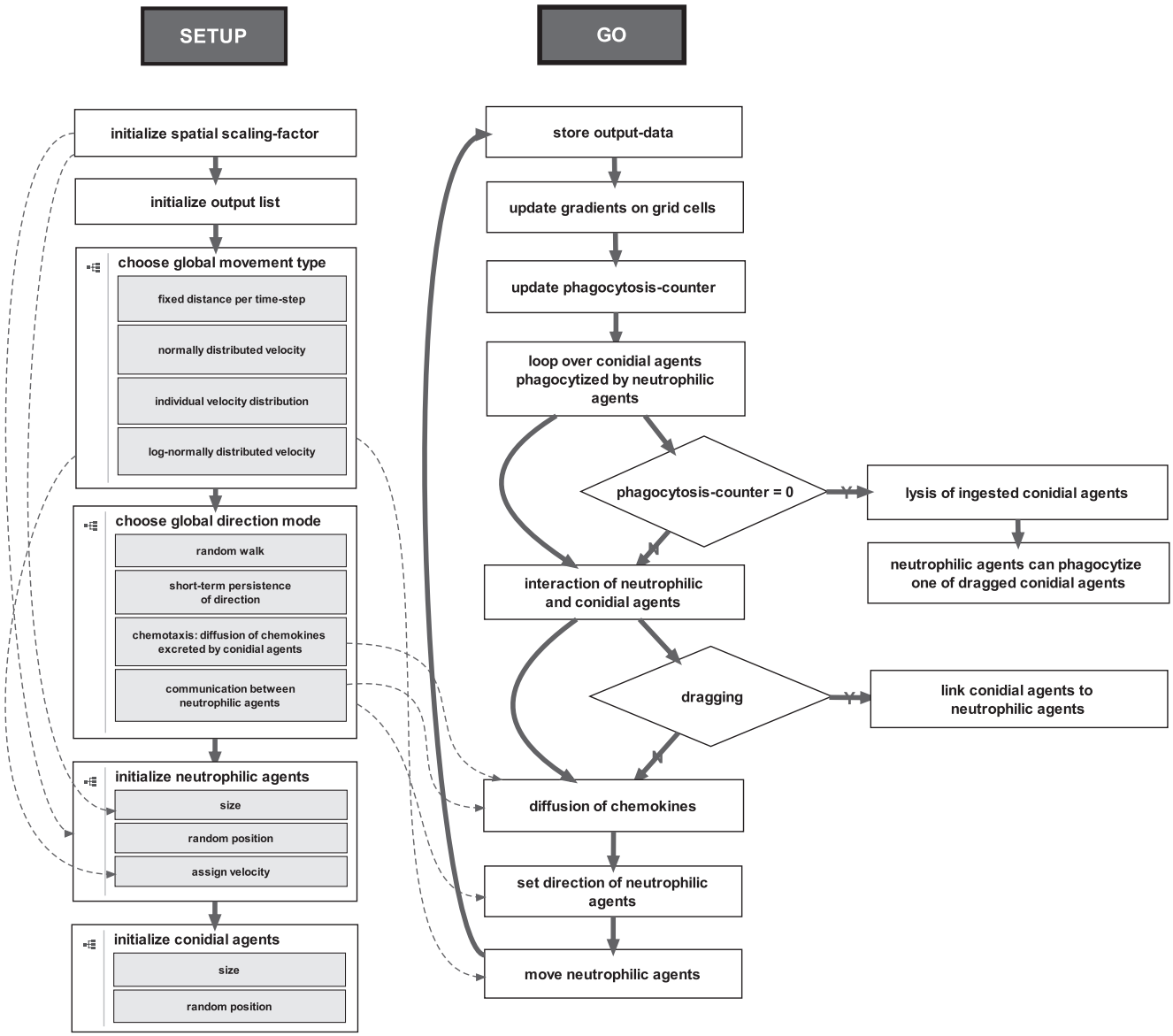


FIGURE 4 | Two main procedures of the ABM. The setup-procedure initializes the environment, the agents and the lists for storing the output-data. The go-procedure is a for-loop over 180 time-steps, which corresponds to the first 90 min of neutrophil-conidia interaction observed by live cell imaging.

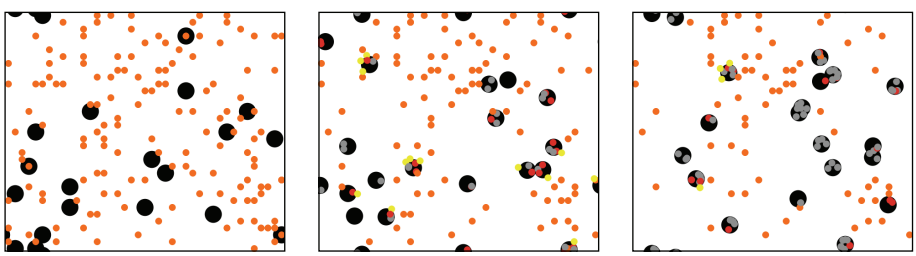


FIGURE 5 | Progress of a typical simulation run. Neutrophilic agents (black) move on the grid randomly or search for free conidial agents (orange), which they may drag (yellow), or phagocytize (red). Conidial agents, which have been phagocytized already (gray), remain in the neutrophilic agent for reasons of visualization and do not further contribute to the simulation run.

as recognition by the complement system result in certain molecules diffusing away from a conidium. Acting as a chemokine they could be sensed by a neutrophilic agent that follows the gradient. In **Figure 8** the clearing efficiency is shown in dependence on the diffusion parameters. If there is only one diffusion step per time-step the chemokine-diffusion-rate does not matter since the chemokines have only diffused to the next neighboring grid cells. The chemokines have not reached far and thus, the clearing efficiency is relatively low. For a high repetition-of-chemokine-diffusion the gradient is too flat if the chemokine-diffusion-rate is not very small. Optimal sets of parameters are a low repetition-of-chemokine-diffusion (but higher than one) with an arbitrary chemokine-diffusion-rate and a higher repetition-of-chemokine-diffusion in combination with a low chemokine-diffusion-rate.

3.1.4. Communication between neutrophils

Is it useful to call other neutrophils if a neutrophil was successful in its hunt? In this model neutrophilic agents secrete communication molecules if they ingested a conidial agent. Other neutrophilic agents follow the gradient and help clearing conidial agents. Surprisingly, at first glance, nearly any communication disturbed the search for conidial agents (see **Figure 9**). At a closer look on the initialization setup it becomes clear that the conidial agents are distributed randomly. Thus, a found conidial agent

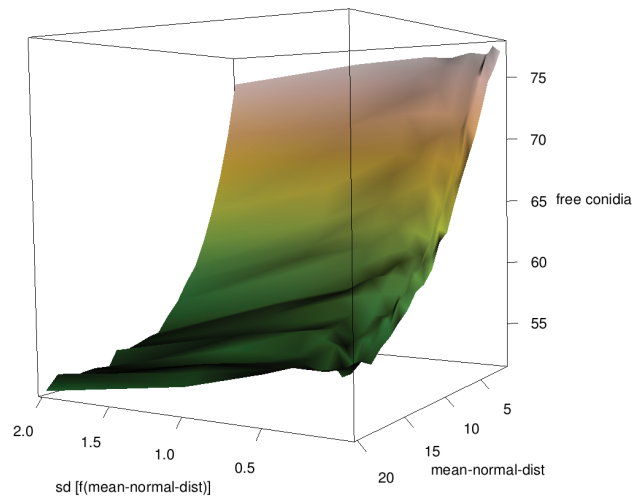


FIGURE 6 | Influence of mean and SD of neutrophilic agents' velocity on clearing efficiency (SD as function of mean).

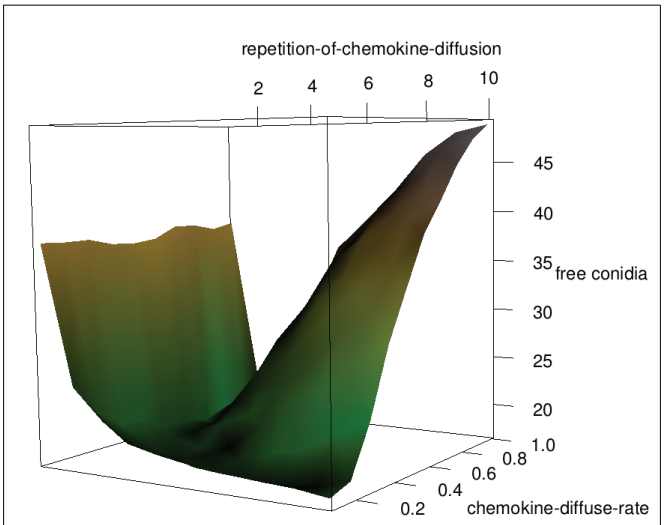


FIGURE 8 | Influence of diffusion parameters on clearing efficiency.

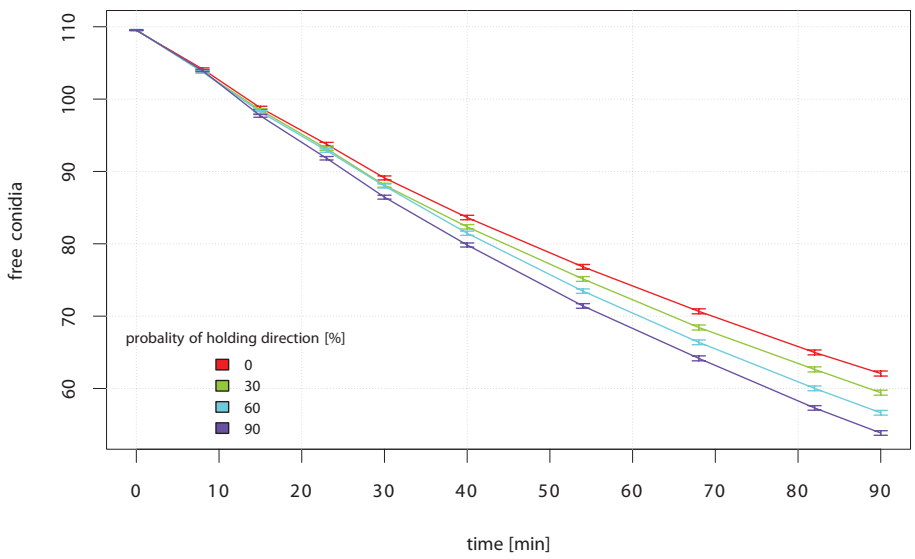


FIGURE 7 | Mean amount of free conidial agents during simulation-time which corresponds to 90 min of *in vivo* interaction. The simulation was repeated for each STP strategy 250 times.

does not imply that other conidial agents are located nearby. Communication between neutrophils leads to an accumulation of neutrophilic agents on locations where conidial agents already have been cleared. Hence, it is a better strategy to search at places where no other neutrophilic agents are and thus, to retain the initial random distribution.

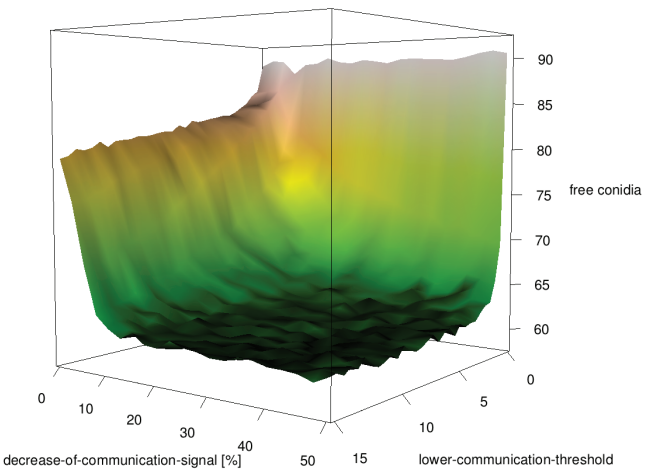


FIGURE 9 | Influence of intra-neutrophilic communication on clearing efficiency of randomly distributed conidial agents.

3.2. SIMULATION OF THE INTERACTION OF NEUTROPHILS AND CLUSTERS OF CONIDIA ON A LARGER SCALE

In a large-scale simulation-approach we tested whether communication between neutrophilic agents increases their efficiency to phagocytize conidial agents. In this setup the distribution of conidial agents was clustered in several randomly located spots on the grid. We evaluated the first 24 h after infection and documented the number of conidial agents which were not yet dragged or phagocytized by neutrophilic agents. In this approach we focused on the major parameters of a possible positive feedback stimulation of neutrophils by other neutrophils, which were already in contact with infectious conidia. Simulation results show that communication between neutrophilic agents significantly increases the efficiency compared to the random walk model (**Figures 10** and **11**). A high repetition-of-communication-signal-diffusion had a positive effect on the clearing efficiency. Here, the chemokines have reached areas far away from the source and thus, could activate neutrophilic agents that are far away from cluster of conidial agents.

A repetition-of-communication-signal-diffusion of 3 seems to produce as good results as higher repetition-of-communication-signal-diffusion (see **Figure 10**) while computation time is shorter. Hence, this parameter remained fixed for the next study while upper-communication-threshold was varied. At the beginning of the clearance process neutrophilic agents with a higher upper threshold of communication signal perception were more successful. Later on, the inverse situation occurred.

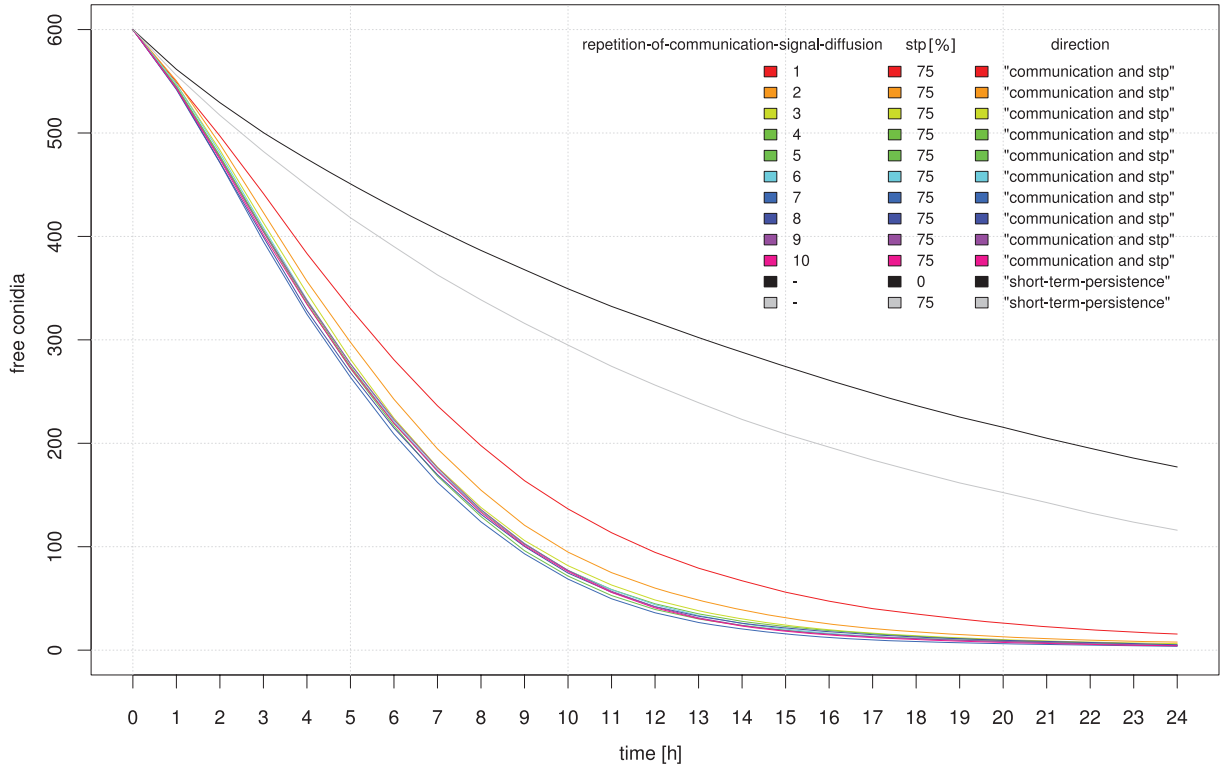


FIGURE 10 | Influence of the diffusion parameter "repetition-of-communication-signal-diffusion" on the clearing efficiency dynamics. For comparison simulation results based on random movement of neutrophilic agents (black) and STP with a 75% probability of holding direction (gray) are shown.

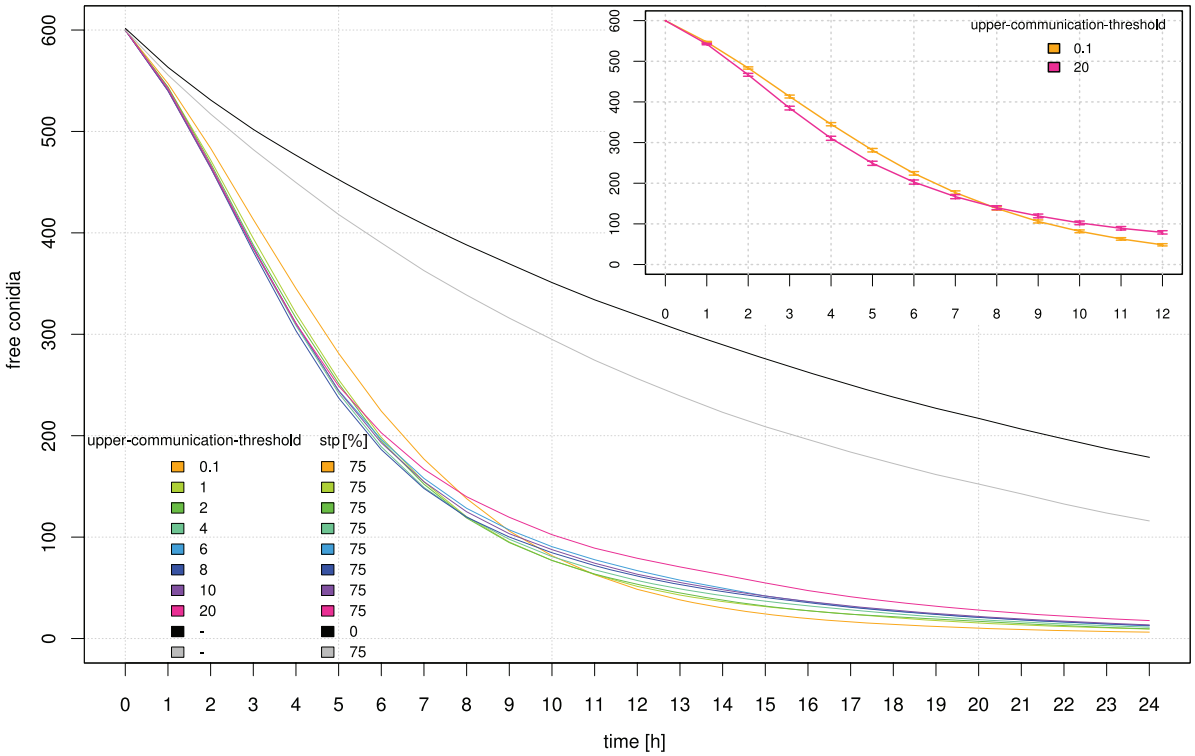


FIGURE 11 | Influence of the signal perception parameter "upper-communication-threshold" on the clearing efficiency dynamics. For comparison, simulation results based on random movement of neutrophilic agents (black) and STP with a 75% probability of holding direction (gray) are shown. The inset shows a zoom into the graph where a crossover of the clearing efficiency for the highest (pink) and lowest (orange) "upper-communication-threshold" occurs which can be interpreted as two opposing clearing strategies. SE are indicated.

A high upper threshold of communication signal perception is useful during the first hours where still many conidial agents are present near to calling neutrophilic agents. At a later time, when only few conidial agents remain in a cluster, a better strategy is to move to the vicinity of successful neutrophilic agents, but not nearby where the area is already cleared. This is realized by a small upper threshold of communication signal perception.

4. DISCUSSION

Neutrophils play an essential role in the elimination of *A. fumigatus* (Balloy and Chignard, 2009; Mircescu et al., 2009). Nevertheless, the detailed mechanisms how these immune effector cells protect the human host are still a matter of debate (Bruns et al., 2010). One effector mechanism is phagocytosis of conidia (Behnsen et al., 2007).

Different potential movement strategies of neutrophils in tracking infectious conidia have been tested in an ABM to explain the *in vitro* efficiency of a population of neutrophils in reducing the amount of free conidia. Simple random walk, random walk with a tendency to keep the former direction (short-term persistence) and chemotactic movement triggered by chemokines excreted by conidia or communication signals sent out by other neutrophils that already have found conidia have been considered. The short-term persistence hunting strategy turned out to be superior to the simple random walk. A similar result has been found in computer simulations on the effectiveness of various search strategies at finding habitat patches (Zollner and Lima, 1999). Following a gradient of chemokines released by conidia is even better. Then, the success in clearing conidia depends on the diffusion parameters of the sensed molecules.

Neutrophils comprise both tracking strategies, short-term persistence to direction and chemotaxis (Tranquillo et al., 1988). In the absence of a gradient of a chemoattractant, they keep direction over the scale of minutes.

A central question tackled in this work is whether chemical communication and chemotaxis of neutrophils improve the clearing efficiency. We found that the answer depends on the spatial distribution of conidia, so that two cases can be distinguished: (a) If conidia are distributed randomly, communication does not pay off because it leads to an aggregation of neutrophils, so that many conidia are unaffected from neutrophilic attacks. (b) If conidia are clustered, communication of neutrophils results in a higher clearing efficiency, since attracting other neutrophils for a faster cooperative clearing pays off.

Calenbuhr and Deneubourg (1992) achieved similar results with a chemotaxis-diffusion model that describes collective hunting strategies using chemical communication in animals. It is found that collective hunting is more efficient at low prey densities whereas individual strategies are more efficient at high prey densities. Clustering of conidia can be seen as decreasing the density of the searched-for objects.

The communication model includes two parameters for cytokine recognition. A lower threshold of perception, and an upper threshold for sensing communication signals between immune cells. The large-scale simulation results showed a non-linear effect of this upper threshold of cytokine recognition on the efficiency of neutrophilic agents in clearing the tissue from infectious conidia. If the upper-communication-threshold was high, neutrophilic agents, activated by another "calling" neutrophilic agent, directly followed the gradient of cytokines until they reached the position of the calling agent. With lower values for the upper-communication-threshold activated neutrophilic agents mainly headed to the whole cluster of conidial agents, which significantly increased their overall efficiency. Scanning through possible values for this upper-communication-threshold revealed two opposing strategies. For a high threshold of cytokine recognition the overall efficiency of neutrophilic agents to clear the tissue was fast, but after a simulated time of 24 h after infection this strategy was not that efficient as simulation runs with a low threshold. We concluded that the overall efficiency of a population of neutrophilic agents can be either "fast but sloppy" possessing a high threshold for communication signal perception or, "slow but thoroughly" having a low threshold.

In an evolutionary context, we assume that selection would favor hosts with a "fast but sloppy" immune-response by neutrophils against pathogens, like *A. fumigatus*. An infection is a race against time and we expect that the immune system of hosts are selected for a fast clearance to prevent conidial germination, even if this might not be the most thorough strategy in the long-run.

However, the possibility remains that neutrophils vary their threshold in signal perception according to the current situation comprising both strategies like they control chemotaxis via different concentrations of chemoattractants by G-protein signaling (Zhang et al., 2009).

ACKNOWLEDGMENTS

The authors are thankful to Matthias Gunzer for providing experimental data on *Aspergillus fumigatus* and murine neutrophils.

The work was supported by the "Pakt für Forschung und Innovation" of the BMBF (http://www.bmbf.de/) and TMBWK (http://www.thueringen.de/de/tmbwk/) to Sabine Hummert, Franziska Mech, and Marc Thilo Figge, BMBF "Virtual Liver" to Christian Tokarski, and DFG (Germany) within the excellence program "Jena School for Microbial Communication" to Anja Schroeter and Stefan Schuster.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Microbial_Immunology/10.3389/fmicb.2012.00129/abstract

REFERENCES

Balloy, V., and Chignard, M. (2009). The innate immune response to *Aspergillus fumigatus*. *Microbes Infect.* 11, 919–927.

Bauer, A. L., Beauchemin, C. A. A., and Perelson, A. S. (2009). Agent-based modeling of host-pathogen systems: the successes and challenges. *Inf. Sci. (Ny)* 179, 1379–1389.

Behnsen, J., Narang, P., Hasenberg, M., Gunzer, F., Bilitewski, U., Klippel, N., Rohde, M., Brock, M., Brakhage, A. A., and Gunzer, M. (2007). Environmental dimensionality controls the interaction of phagocytes with the pathogenic fungi *Aspergillus fumigatus* and *Candida albicans*. *PLoS Pathog.* 3, e13. doi:10.1371/journal.ppat.0030013

Bruns, S., Kniemeyer, O., Hasenberg, M., Aimanianda, V., Nietzsche, S., Thywissen, A., Jeron, A., Latgé, J.-P., Brakhage, A. A., and Gunzer, M. (2010). Production of extracellular traps against *Aspergillus fumigatus* in vitro and in infected lung tissue is dependent on invading neutrophils and influenced by hydrophobin roda. *PLoS Pathog.* 6, e1000873. doi:10.1371/journal.ppat.1000873

Calenbuhr, V., and Deneubourg, J. (1992). *Biology and Evolution of Social Insects, Chapter Pattern Formation Via Chemical Communication: Collective and Individual Strategies*. Leuven: Leuven University Press, 343–349.

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

Forrest, S., and Beauchemin, C. (2007). Computer immunology. *Immunol. Rev.* 216, 176–197.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giskec, J., Goss-Custard, J., Grand, T., Heinz, S., Huse, G., Huth, A., Jepsen, J., Jørgensen, C., Mooij, W., Müller, B., Pe'er, G., Piou, C., Railsback, S., Robbins, A., Robbins, M., Rossmanith, E., Rüger, N., Strand, E., Souiss, S., Stillman, R., Vabø, R., Visser, U., and DeAngelis, D. (2006). A standard protocol for describing individual-based and agent-based models. *Ecol. Modell.* 198, 115–126.

Grimm, V., Berger, U., DeAngelis, D., Polhill, J., Giske, J., and Railsback, S. (2010). The ODD protocol: a review and first update. *Ecol. Modell.* 221, 2760–2768.

Hohl, T. M. (2009). Stage-specific innate immune recognition of *Aspergillus fumigatus* and modulation by echinocandin drugs. *Med. Mycol.* 47, 192–198.

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

Karkowska-Kuleta, J., Rapala-Kozik, M., and Kozik, A. (2009). Fungi pathogenic to humans: molecular bases of virulence of *Candida albicans*, *Cryptococcus neoformans* and *Aspergillus fumigatus*. *Acta Biochim. Pol.* 56, 211–224.

Mech, F., Thywissen, A., Guthke, R., Brakhage, A. A., and Figge, M. T. (2011). Automated image analysis of the host-pathogen interaction between phagocytes and *Aspergillus fumigatus*. *PLoS ONE* 6, e19591. doi:10.1371/journal.pone.0019591

Mircescu, M. M., Lipuma, L., van Rooijen, N., Pamer, E. G., and Hohl, T. M. (2009). Essential role for neutrophils but not alveolar macrophages at early time points following *Aspergillus fumigatus* infection. *J. Infect. Dis.* 200, 647–656.

Paddock, S. W. (1999). Confocal laser scanning microscopy. *BioTechniques* 27, 992–996, 998–1002, 1004.

Pluta, M. (1988). *Advanced Light Microscopy: Principles and Basic Properties*, Vol. 1. New York: Elsevier.

Richardson, M. D. (2005). Changing patterns and trends in systemic fungal infections. *J. Antimicrob. Chemother.* 56(Suppl. 1), i5–i11.

Scapini, P., Lapinet-Vera, J. A., Gasperini, S., Calzetti, F., Bazzoni, F., and Cassatella, M. A. (2000). The neutrophil as a cellular source of chemokines. *Immunol. Rev.* 177, 195–203.

Shoham, S., and Levitz, S. M. (2005). The immune response to fungal infections. *Br. J. Haematol.* 129, 569–582.

Tranquillo, R. T., Zigmond, S. H., and Lauffenburger, D. A. (1988). Measurement of the chemotaxis coefficient for human neutrophils in the under-agarose migration assay. *Cell Motil. Cytoskeleton* 11, 1–15.

Waldorf, A. R., and Diamond, R. D. (1985). Neutrophil chemotactic responses induced by fresh and swollen *Rhizopus oryzae* spores and *Aspergillus fumigatus* conidia. *Infect. Immun.* 48, 458–463.

Wilensky, U. (1999). *NetLogo. Center for Connected Learning and Computer-Based Modeling*. Evanston, IL: Northwestern University.

Zhang, H., Sun, C., Glogauer, M., and Bokoch, G. M. (2009). Human neutrophils coordinate chemotaxis by differential activation of Rac1 and Rac2. *J. Immunol.* 183, 2718–2728.

Zollner, P. A., and Lima, S. L. (1999). Search strategies for landscape-level interpatch movements. *Ecology* 80, 1019–1030.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 December 2011; accepted: 19 March 2012; published online: 26 April 2012.

*Citation: Tokarski C, Hummert S, Mech F, Figge MT, Germerodt S, Schroeter A and Schuster S (2012) Agent-based modeling approach of immune defense against spores of opportunistic human pathogenic fungi. Front. Microbio. **3**:129. doi: 10.3389/fmicb.2012.00129*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

APPENDIX

Table A1 | List of agents.

State variables	Brief description
GRID CELLS	
Size	Corresponds to the mean conidium size of *A. fumigatus* of 2.5 μm
Radius of neutrophil	2.5 grid cells
Conidia-chemokines	Amount of chemokines segregated by conidial agents
Neutrophil chemokines	Amount of chemokines segregated by neutrophilic agents
CONIDIAL AGENTS	
Size	Same size as a grid cell
In-zone	Stores the identity number of the neutrophilic agent in whose radius a conidial agent finds itself
Dragged	– By which it is dragged
Caught	– By which it is caught
Phagocytosis-counter	Counts the time until a caught conidial agent is digested
NEUTROPHILIC AGENTS	
Identity number	Unique number for each neutrophilic agent
Size	Size-ratio of neutrophilic and conidial agents is 2.5
Velocity	Depends on movement mode
Direction	Depends on direction mode
Catching	Number of caught conidial agents
Dragging	Number of dragged conidial agents
Phagocytized	Number of phagocytized conidial agents

Table A2 | List of global parameters.

Parameter	Brief description
SETUP	
Initial-number-of-conidia	Initial population size of conidial agents
Initial-number-of-neutrophils	Initial population size of neutrophilic agents
Velocities-mode-of-neutrophils	Setup of neutrophilic agents' movement options
Direction-mode-of-neutrophils	Setup of neutrophilic agents' direction options
Move-conidia	Conidial agents are moved randomly around their initial position
INTERACTION	
Catch%	Neutrophilic agents' probability to: catch a free conidial agent
Drag%	– Drag a free conidial agent
Drag-to-release%	– Release a dragged conidial agent
Drag-to-catch%	– Phagocytize a dragged conidial agent
Phagocytosis-capacity	Maximum number of conidial agents which can be phagocytized by a neutrophilic agent
Phagocytosis-time	Duration of phagocytosis

Table A3 | Neutrophilic agents' movement modes.

Parameter	Brief description
Random walk	Random walk of neutrophilic agents
Mean-neutros	Option for normally-distributed mean of neutrophilic agents' velocity (with fixed standard deviation)
Mean-normal dist	Option for normally-distributed mean of neutrophilic agents' velocity
SD-normal dist	Option for normally-distributed standard deviation of neutrophilic agents' velocity
Short-term persistence (STP)	Neutrophilic agents hold their actual direction with a certain probability
Hold-direction	Probability to hold given direction at the next step
Diffusion of conidia-chemokines	Activation of neutrophilic agents through conidia-chemokines, neutrophilic agents follow chemokine gradients
Amount-of-chemokines	Amount of chemokines spread by the free conidial agents per time-step
Chemokine-perception-threshold	Neutrophilic agents' lower threshold of chemokine perception
Chemokine-diffusion-rate	Degree of diffusion
Repetition-of-chemokine-diffusion	Velocity of diffusion
Communication between neutrophils	Neutrophilic agents follow chemokine gradients segregated by activated neutrophilic agents (positive feedback-activation)
Activated neutrophil	Attracts other neutrophilic agents (positive feedback-activation)
Communication-signal	Signal strength of chemical communication spread by an activated neutrophilic agent
Decrease-of-communication-signal	Option for reducing the strength of communication signal
Lower-communication-threshold	Neutrophilic agents' lower threshold of communication signal perception
Communication-over-chemotaxis	Option to rank priority of chemokine perception and communication signal perception of neutrophilic agents
Communication-signal-diffusion-rate	Degree of diffusion
Repetition-of-communication-signal-diffusion	Velocity of diffusion
LARGE-SCALE GRID WITH CLUSTERS OF CONIDIA	
Number-of-conidia-clusters	Initial number of spots of infection
Size-of-clusters	Size of spot of infection
Density-of-clusters	Number of conidial agents per spot of infection
Upper-communication-threshold	Neutrophilic agents' upper threshold of communication signal perception

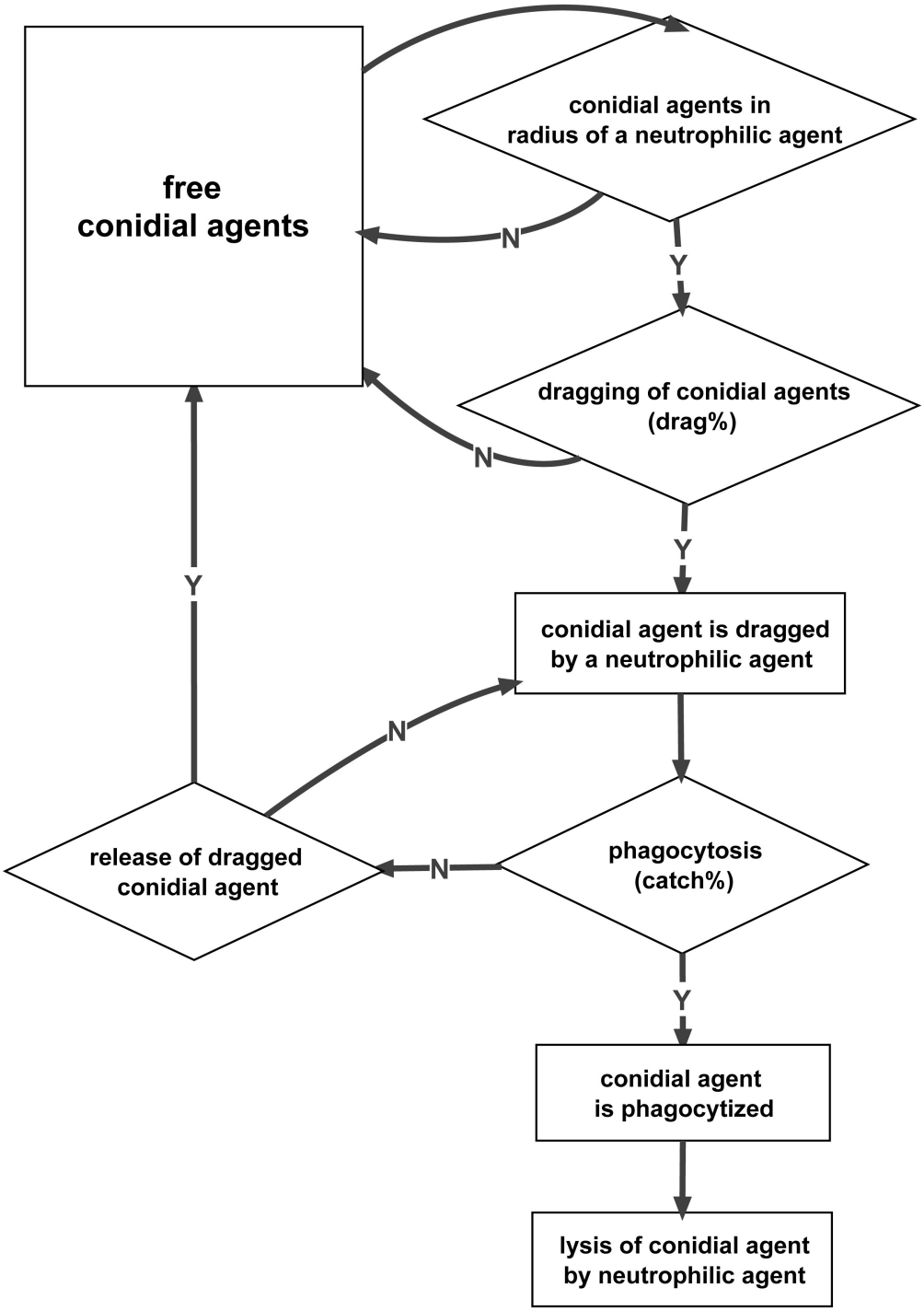


FIGURE A1 | Interaction of conidial and neutrophilic agents.

frontiers in
MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 16 February 2012
doi: 10.3389/fmicb.2012.00051

Genome-wide scale-free network inference for *Candida albicans*

Robert Altwasser[1]*, Jörg Linde[1], Ekaterina Buyko[2], Udo Hahn[2] and Reinhard Guthke[1]

[1] Research Group Systems Biology/Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knoell Institute, Jena, Germany
[2] Jena University Language and Information Engineering Lab, Friedrich Schiller University, Jena, Germany

Edited by:
Franziska Mech, Hans Knöll Institute, Germany

Reviewed by:
Anke Meyer-Baese, Florida State University, USA
Steffen Rupp, Fraunhofer Gesellschaft, Germany

***Correspondence:**
Robert Altwasser, Research Group Systems Biology/Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knoell Institute, Beutenbergstr. 11a, 07743 Jena, Germany.
e-mail: robert.altwasser@hki-jena.de

Discovery of essential genes in pathogenic organisms is an important step in the development of new medication. Despite a growing number of genome data available, little is known about *C. albicans*, a major fungal pathogen. Most of the human population carries *C. albicans* as commensal, but it can cause systemic infection that may lead to the death of the host if the immune system has deteriorated. In many organisms central nodes in the interaction network (hubs) play a crucial role for information and energy transport. Knockouts of such hubs often lead to lethal phenotypes making them interesting drug targets. To identify these central genes via topological analysis, we inferred gene regulatory networks that are sparse and scale-free. We collected information from various sources to complement the limited expression data available. We utilized a linear regression algorithm to infer genome-wide gene regulatory interaction networks. To evaluate the predictive power of our approach, we used an automated text-mining system that scanned full-text research papers for known interactions. With the help of the compendium of known interactions, we also optimize the influence of the prior knowledge and the sparseness of the model to achieve the best results. We compare the results of our approach with those of other state-of-the-art network inference methods and show that we outperform those methods. Finally we identify a number of hubs in the genome of the fungus and investigate their biological relevance.

Keywords: network inference, linear regression, LASSO, reverse engineering, scale-free, *Candida albicans*, hubs, prior knowledge

1. INTRODUCTION

Candida albicans is the most important human-pathogenic fungus (D'Enfert and Hube, 2007). Most of the time, it lives as a commensal in the microbial flora of the host. However, if the immune system of the host is impaired, it can switch to an aggressive pathogen that can cause systematic infections with a high mortality rate (Wilson et al., 2002). An important prerequisite of *C. albicans* virulence is its ability to react upon environmental changes such as temperature shifts, pH value changes, or nutrition supply. *C. albicans* can react to these environmental conditions by altering its gene expression pattern. These alteration can create phenotype changes, like switching from typical yeast-like ovoid to hyphal growth form (Hube, 2004). These changes in morphology are a crucial part of the infectious ability of *C. albicans*. Understanding how these gene expression alterations change the morphology of the fungus can uncover new therapeutic methods to counter fungal infections.

Gene expression regulation is primarily mediated by transcription factors but also by post-translational modification or other mechanisms. Reverse engineering of such mechanisms is an important part of systems biology (Hecker et al., 2009a). It aims to uncover essential interactions within the genome of the organism. This research is facilitated by the growing number of expression data available (Edgar et al., 2002).

Network inference approaches have been successfully applied in order to infer small-scale networks and to predict gene interactions for pathogenic fungi (Guthke et al., 2005, 2007; Linde et al., 2010). Such networks investigate certain aspects of regulatory processes and provide valuable information regarding specific gene interactions. However, the number of genes that can be considered using such approaches is limited. Topological analysis of the full genome is beyond the scope of this approach.

Different methods for the reverse engineering of genome-wide inferences have been developed. A common approach is the use of information-theoretic principles. Some define interactions between genes as statistical dependencies between gene expression profiles (Margolin et al., 2006). The idea is that statistical dependencies, that can not be explained as artifacts of other dependencies in the network, are likely to identify direct regulatory interactions. These methods are also called *mutual information*. Common representatives are ARACNE (Margolin et al., 2006), MRNET (Meyer et al., 2007), and CLRNET (Faith et al., 2007). Due to the nature of these methods, mutual information networks are primarily undirected, e.g., the network does not discriminate between source and target gene of an interaction.

In this work, we use a system of linear equations to model the regulatory interactions between genes. The idea of this approach is to model the expression of one gene as the weighted sum of the

expression of other genes and external perturbations (Gustafsson et al., 2004). The advantage of this approach is, that it can describe gene interaction in a quantitative way that takes the direction of the interaction into account, i.e., it discriminates between the source and the target gene of an interaction. Topological motives like feedback loops can be described as well as dynamic processes within gene regulatory interactions (Hecker et al., 2009b). A commonly used algorithm is the so-called LASSO (Tibshirani, 1994). It works well under the condition, that there are more genes than samples, which is mostly the case in biological data. This approach has already been implemented for pathogenic fungi like *C. albicans* (Linde et al., 2011). However, these models have not been scale-free.

One of the most severe problems researchers face while defining network inferences for fungi is the dearth of available information. So far, there are only few data sets for pathogenic fungi available, mostly from microarray experiments. This problem becomes even more serious when modeling networks including a large number of genes. One approach is the use of proposed gene interactions called prior knowledge, taken from data sources different from gene expression data. This concept has been successfully implemented in earlier approaches (Linde et al., 2011) and was used in this work as well.

Topological analysis of large-scale networks can unravel interesting interactions and regulatory genes with a high number of interaction partners called *hubs*. Hubs are essential for the viability of the organism since they are a central part of the interaction network architecture. Because of the large number of interactions, it is very likely to destroy an essential interaction by knocking out a hub (Han et al., 2004; He and Zhang, 2006). This property makes hubs interesting drug targets. Frequently, genome-wide models do not meet the requirement of scale-freeness, i.e., the distribution of connections between nodes does not follow a power-law. However, scale-freeness is a pre-condition for topological analysis and the detection of hubs because most biological networks exhibit such a power-law distribution (Barabási and Oltvai, 2004).

In this study, we combine the LASSO with the ridge regression, a method of regularization, as proposed by Gustafsson et al. (2004), to infer scale-free networks. We extend this approach to our gene data by implementing different sources of prior knowledge to our gene expression data. We use an automatic relation extraction system to scan 9,000 research papers in order to get a compendium of currently known interactions to compare and evaluate our networks. We then perform topological analysis on these networks to identify hubs. We investigate these hubs for their biological function. We also compare our algorithm to state-of-the-art methods.

2. MATERIALS AND METHODS

2.1. DATA

2.1.1. Gene expression data set

We took genome-wide gene expression data of *C. albicans* from a collection of Ihmels et al. (2005). The data set consists of transcription data of 6,167 open reading frames (ORF) under 198 conditions ranging from drug application, via stress exposition to response to mating pheromone. The set contains transcriptional profiles of cells growing as yeast or hyphal cells taken from four independent microarray designs. 16.7% of the data are missing. Four hundred eleven ORFs have more than 50% missing values. We tested different imputation methods to complete the data set and applied the best performing method LLS, since the used network inference method requires complete observations. We applied the Local Least Squares (LLS) imputation method as provided by the pca Method (Stacklies et al., 2007) package for R (R Development Core Team, 2009).

2.1.2. Gold standard

We evaluated the performance of the network inference approaches with emphasis on the reliability of the predicted interactions. The data set on which this evaluation was based was generated using text-mining technology. Accordingly, we automatically extracted information about gene regulatory interactions from full-text research articles in order to collect a set of known interactions published in the literature. Text mining was based on JReX (Buyko et al., 2011), a high-performance machine-learning relation extraction system. JReX identifies pairs of genes as interaction pairs exploiting rich syntactic and semantic information. Using this system, we harvested gene regulation information from about 9,000 open-access research papers about *C. albicans*. The resulting collection contains 509 genes and 1,016 interactions between them. We are very much aware of the fact that this procedure has inherent limitations (e.g., f-scores ranging between 50 and 60% are consistently reported for such approaches (Kim et al., 2011)), but in the absence of a comprehensive manually generated gold standard, we used this automatically built gold standard to evaluate the networks inferred using different methods and parameter settings. Only 503 genes of the gold standard are part of our gene expression data set. Therefore, these 503 *gold genes* were used to optimize different parameters.

2.2. NETWORK INFERENCE

To infer a regulatory network in *C. albicans*, we used a modeling approach based on linear regression. This approach describes the expression of a gene x_i under condition m as the weighted sum of the expression of the other genes under this condition:

$$x_i(m) = \sum_{\substack{j=1,\\ j\neq i}}^{N} \beta_{i,j} x_j(m) \tag{1}$$

N is the number of genes and $x_j = x_j(1),\ldots,x_j(M)$ describes the expression of gene j under the condition 1 to M. $\beta_{i,j}$ is the coefficient that describes the influence of gene x_j on gene x_i. The strength of the interaction is represented by the absolute value of the coefficient. This coefficients can be positive or negative, representing activating or inhibiting relations, respectively. A coefficient equal to zero means there is no interaction between these genes.

The equation system, defined in (1), has more variables than equations, i.e., more genes than samples. To cope with this problem and to enhance the interpretability of the inferred network, we followed the idea of sparseness (Yeung et al., 2002; Leclerc, 2008). This concept tries to maximize the number of zeros in the interaction matrix $B = \beta_{i,j}$. To solve this task, (Tibshirani, 1994) proposed the Least Absolute Shrinkage and Selection Operator (LASSO) algorithm. It applies the L^1-norm shown in equation (3) on the interaction matrix B and assigns many weights zero. To find

the model that fits best to the expression data, we minimized the *residual sum of squares* (RSS):

$$\hat{\beta}_{i,\cdot} = \underset{\beta_{i,\cdot}}{\arg\min} \sum_{m=1}^{M} \left(x_i(m) - \sum_{\substack{j=1,\\ j\neq i}}^{N} \beta_{i,j} x_j(m) \right)^2 \tag{2}$$

$$\text{subject to} \sum_{\substack{j=1,\\ j\neq i}}^{N} |\beta_{i,j}| \leq \mu_i \text{ for } i = 1, \ldots, N \tag{3}$$

where μ_i is a parameter limiting the absolute sum of all $\beta_{i,\cdot}$. To account for the varying reliability of the prior knowledge, we introduce an additional weight parameter $\omega_{i,j}$, denoting the reliability of interaction $\beta_{i,j}$. Hereby we follow a knowledge-driven approach and extend the equation (3) as presented by Zou (2006):

$$\sum_{\substack{j=1,\\ j\neq i}}^{N} \omega_{i,j} |\beta_{i,j}| \leq \mu_i \tag{4}$$

By default, all interactions $\omega_{i,j}$ have a value of 1. A small value of $\omega_{i,}j$ means that the interaction is reliable, while larger $\omega_{i,j}$ indicate questionable interactions. Setting $\omega_{i,j} = 0$ means that we trust $x_{i,j}$ unconditionally.

The prior knowledge was incorporated by the creation of an $N \times N$ penalty matrix Ω. The component $\omega_{i,j}$ of the matrix Ω is multiplied by $\beta_{i,j}$ during the computation of the threshold shown in equation (4). If a source of prior knowledge predicts an interaction between two edges i and j, the penalty of this interaction is $\omega_{i,j} = \epsilon^n$ where n is the number of prior knowledge sources that support the interaction. If an interaction is not supported by any prior knowledge, then $\omega_{i,j} = 1$.

To determine the optimal value for μ_i, we follow the approach suggested by Gustafsson et al. (2004, 2005). This approach first minimizes the L^2-norm:

$$\mu_i^{(2)} = \left(\sum_{\substack{j=1,\\ j\neq i}}^{N} \left(\omega_{i,j} \beta_{i,j} \right)^2 \right)^{\frac{1}{2}} \tag{5}$$

and set $\mu_i = c\mu_i^{(2)}$. The networks created using this method were proved to be scale-free.

The inference of genome-wide networks is computationally intensive. However, the calculation of the regression for one gene is independent from the regression of other genes. This way, the network inference factorizes and we used parallel computing to speed up the inference.

3. RESULTS

3.1. PARAMETER ESTIMATION AND NETWORK ASSESSMENT

The result of the inference depends on different parameters, that need to be estimated. The parameter ϵ defines the influence of the prior knowledge. It is too time consuming to perform an exhaustive search over this parameter exploiting the whole expression data set. Therefore, we only selected the expression data of genes, that are included in the *gold standard*. This subset contained the expression data from 503 genes, called *gold genes*. With this subset we investigated the influence of the prior knowledge by using a search over ten equidistant values each within the intervals 0.01, ..., 0.1 and 0.1, ..., 1 and calculated the F-measure (Van Rijsbergen, 1979) of the inferred networks. The F-measure incorporates the trade off between the recall (completeness of the identified interactions within the *gold standard*) and the precision (ratio of correctly identified interactions).

$$F = 2 * \frac{precision * recall}{precision + recall} \tag{6}$$

The second parameter to optimize determines the size of the network, i.e., the number of inferred interactions. LASSO works with constraint introduced by the parameter μ_i. As suggested by (Gustafsson et al., 2004), we first calculate the parameter$\mu_i^{(2)}$ via equation (5) and define $\mu_i = c\mu_i^{(2)}$. Gustafsson et al. fixed c at 0.1 and stated that deviating from this value does not result in large changes in the selected interactions and still leads to a scale-free network. Nevertheless, we performed a grid search over 24 different steps for c ranging from 0.00001 to 0.5. We calculated the corresponding F-measure with regard to the *gold standard* and degree of scale-freeness for all inferred networks.

Results show that smaller values for ϵ, i.e., more influence of the prior knowledge, yield higher F-measures. For the BIND prior knowledge, the result of different values of ϵ is depicted in **Figure 1**. Because of these results, we choose $\epsilon = 0.1$ for the network construction for all known interactions.

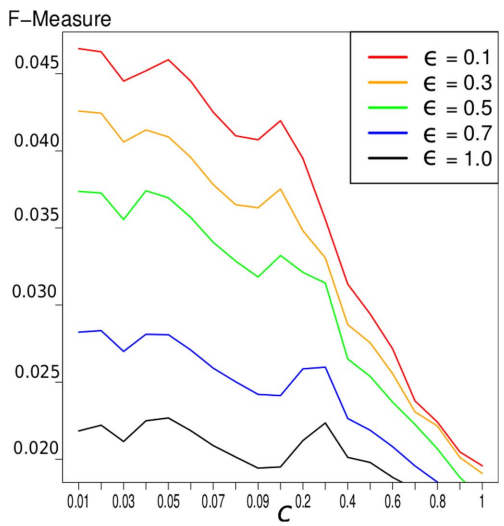


FIGURE 1 | F-measure of the LASSO inference for the 503 *gold genes* in which the *gold standard* and the expression data overlap. We exploited the BIND prior knowledge. The different graphs represent different values of ϵ and therefore different weighting of prior knowledge. It indicates that a higher influence of prior knowledge yields better results concerning the F-measure.

To study the influence of the different prior knowledge sources, we first constructed a genome-wide network without including prior knowledge in the model. Subsequently, we constructed networks involving all four sources of prior knowledge individually. After that, we also created one network that used all available prior knowledge to infer a network.

In the following, we took the full-genomic network that was supported by all prior knowledge sources (ALL). Since we computed this network for different network sizes, by variation of parameter c, we selected the one with the highest F-measure, which was $c = 0.2$, as illustrated in **Figure 2**.

In order to compare our approach to state-of-the-art methods, we also inferred genome-wide networks based on mutual information, like ARACNE (Margolin et al., 2006), MRNET (Meyer et al., 2007), and CLRNET (Faith et al., 2007). The results of the inferred networks can be seen in **Table 1**.

The results of these tests are shown in **Figure 3**. Comparing the LASSO-based networks without or with different prior knowledge sources, we found that the implementation of prior knowledge clearly improves the performance of the inference, especially when exploiting the BIND set of prior knowledge results in a high F-measure compared to the *gold standard*. All LASSO-based inferences outperform the networks constructed using mutual information. The inferred networks differ remarkably in size. While the LASSO-based networks are comparably sparse, having around 6,200–6,900 interactions, the network inferred by ARACNE has around 40,000 interactions. CLRNET and MRNET inferred networks contain about 15,000,000 interactions.

All of the networks inferred by LASSO are scale-free, as can be seen in **Figure 4**. We calculated the correlation of the degree distribution to the power-law distribution using *Cytoscape* (Smoot et al., 2011). The LASSO network that implemented all prior knowledge sources has a correlation coefficient of 0.88. This was the lowest correlation of all LASSO-based methods. In contrast, none of the mutual information networks are scale-free, see **Figure 5**.

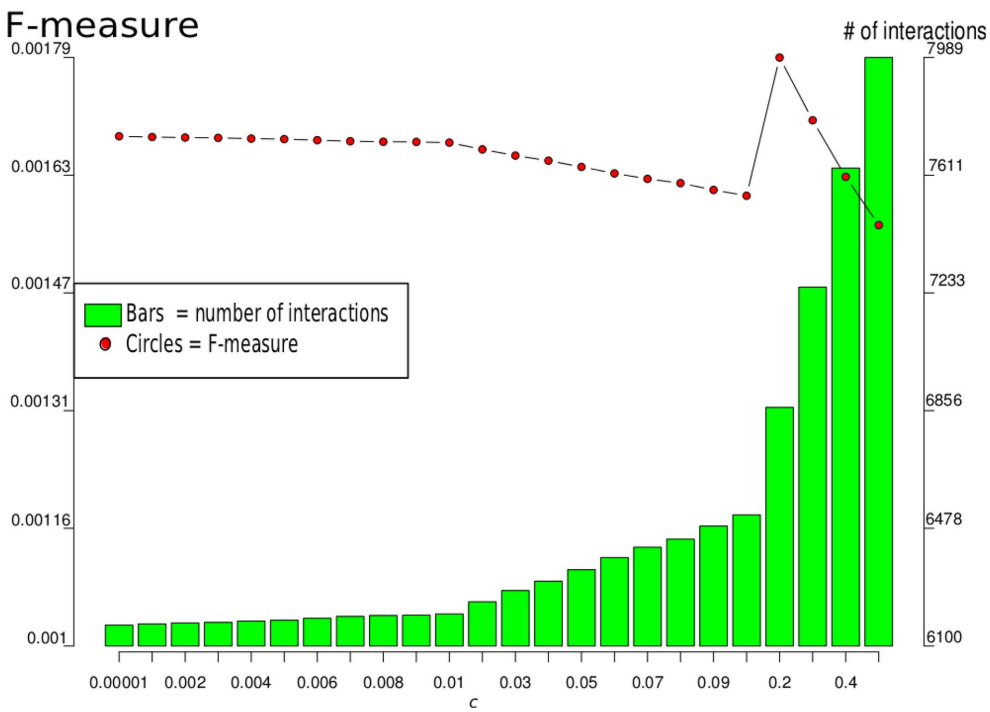


FIGURE 2 | F-measures and number of interactions for different values of *c* for the LASSO-based genome-wide inference that used all genes and all available prior knowledge. The maximum F-measure 0.0018 is reached at $c = 0.2$ with a network of 6866 interactions between 6,167 genes.

Table 1 | Results of the genome-wide network inference.

	LASSO	LASSO + FAC	LASSO + PPI	LASSO + TRANS	LASSO + BIND	LASSO + ALL	CLRNET	MRNET	ARACNE
F-measure	0.0014	0.0015	0.0053	0.0058	0.0067	0.0018	0.00006	0.00006	0.0009
No. of interactions	6,167	6,167	6,167	6,167	6,167	6,866	15,686,064	15,329,450	39,986

The first six rows show the results for LASSO and LASSO with different prior knowledge sources. The sixth row shows the LASSO inference with ALL four sources of prior knowledge. The last three rows show the results for the mutual information-based networks.

Figure 6 shows that there is little overlap between the *gold standard*, which we extracted from literature concerning *C. albicans*, and the prior knowledge, extracted from data bases where *C. albicans* is underrepresented. Besides BIND and PPI, none of the prior knowledge sources have a large overlap. Also the prior knowledge sources and the *gold standard* barely overlap with each other. FAC is by far the smallest of the prior knowledge sources (249 interactions) and only 14 of them are also part of the *gold standard*. Therefore, it is not surprising, that the network inferred exploiting FAC yields the smallest improvement concerning the F-measure over the network inferred without prior knowledge (**Figure 3**). The LASSO without the use of prior knowledge reaches a F-measure of 0.0014 and the use of the FAC improves this result to 0.0015. With the information of PPI, LASSO reaches a F-measure of 0.0053, with TRANS 0.0058 and 0.0067 with BIND.

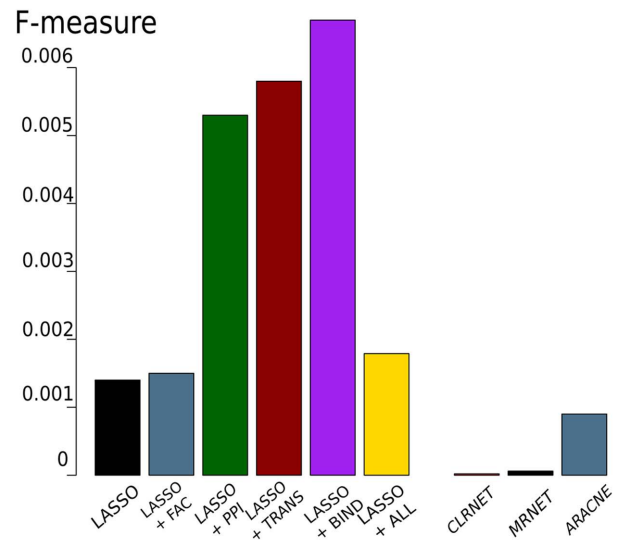


FIGURE 3 | F-measure obtained by LASSO-based genome-wide network inferences (left) with or without prior knowledge (FAC, PPI, TRANS, BIND) and with all four prior knowledge sources (ALL). The three bars on the right show the results of the mutual information-based networks.

3.2. CENTRAL GENES

This study aims at identifying *hubs*, i.e., genes with high influence on other genes. (Han et al., 2004) propose that hubs should have at least six interactions with other genes. Since our networks have more nodes than those by Han, we considered an out degree of seven or more to be reasonable. We found 126 genes with an out degree of at least seven and examined them for their function (Arnaud et al., 2010). Ten of them are shown in **Table 2**.

Since there is little information available for *C. albicans*, most of the hub genes we found are still not functionally annotated. We often only found information from ortholog genes in *S. cerevisiae*. The information found indicates that the putative hub genes regulate various cell functions. At least 16 of the 126 hubs are influenced by known antimycotica like *amphotericin B*, *caspofungin*, or the *azole* group as shown in **Table 3**. Thirty-one of the identified hub genes are still not annotated and no functional information is available.

One of the few well studied networks within yeasts is the so-called *GAL*-network. It has been comprehensively studied for *S. cerevisiae* (Johnston, 1987; Lohr et al., 1995). It was also used to

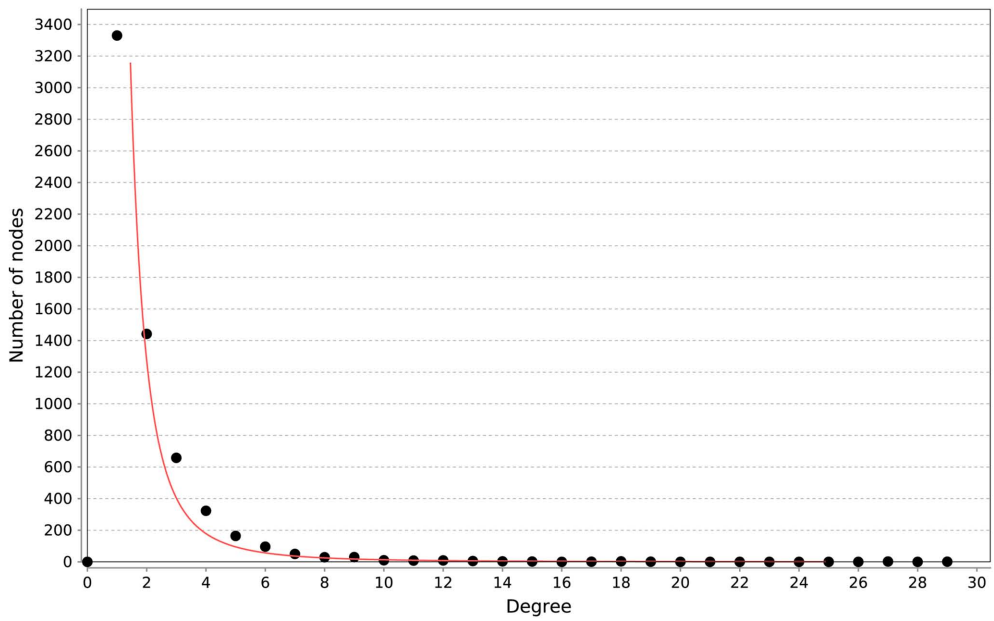


FIGURE 4 | Distribution of degrees for the LASSO-based inference without prior knowledge. The red line represents the fitted power-law. The correlation coefficient of the logarithmical data is 0.95.

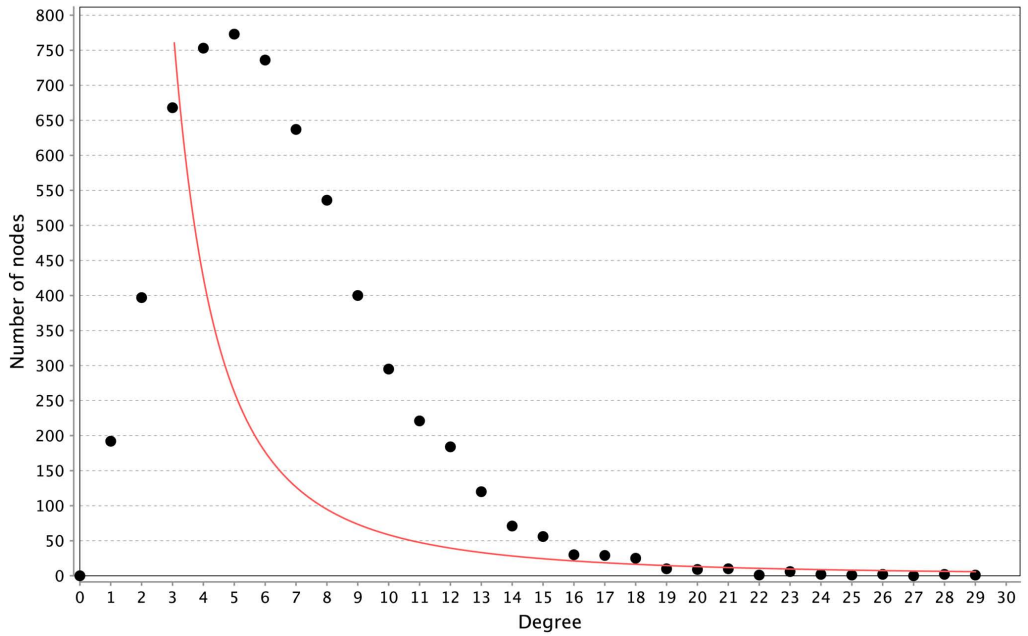


FIGURE 5 | Distribution of degrees for the ARACNE-based inference. The red line represents the fitted power-law. The correlation coefficient of the logarithmical data is 0.08.

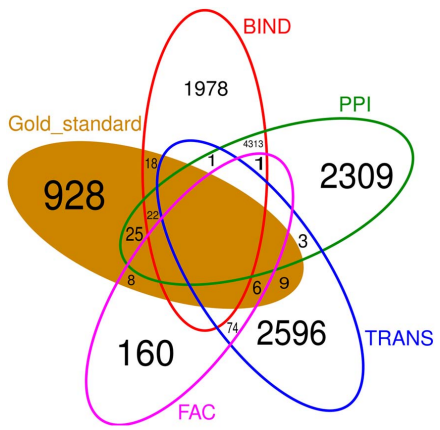


FIGURE 6 | Venn diagram of the four different prior knowledge sources and the *gold standard*. Empty fields contain no common interaction. There is little overlap between the sources of prior knowledge as well as between the prior knowledge and the *gold standard*.

investigate transcriptional rewiring between tit *C. albicans* and *S. cerevisiae* (Rokas and Hittinger, 2007). The GAL-network is responsible for the degradation of galactose. Via *GAL10*, β-D-galactose is transferred to α-D-galactose which is transferred to α-D-galactose 1-phosphate by *GAL1*. *GAL7* then converts α-D-galactose 1-phosphate to α-D-glucose 1-phosphate. The direct regulation *GAL10* → *GAL1* → *GAL7* is predicted by the inferred network models, as can be seen in **Figure 7**, even though it is not part of any prior knowledge. Only the interaction *GAL1* → *GAL7* is part of the *gold standard*.

Table 2 | Ten genes with the highest out degree of the LASSO network inferred with all four sources of prior knowledge (ALL).

Gene name	Out degree
FET31	29
orf19.7450	28
orf19.1300	25
MAL2	20
orf19.4678	19
orf19.1735	18
SGO1	17
orf19.6715	17
Yor353c	15
PSA2	15

The **Figures 8** and **9** illustrate how usage of different sources of prior knowledge affect the connectivity of genes. *PSA2* is involved in nucleotidyltransferase activity and biosynthetic processes (Arnaud et al., 2010). *TKL1* is involved in transketolase activity, part of the cell wall in yeast form, and possibly essential for the viability of the organism.

4. DISCUSSION

We inferred genome-wide scale-free gene regulatory network inference models by exploitation of prior knowledge. The soft integration of prior knowledge can tackle the problem of insufficient data and improves the performance of the inference algorithm. The level of improvement depends on the quality and quantity of the prior knowledge.

Table 3 | Table of 16 hubs which are sensitive to antifungal treatment.

Yor353c	Domain protein of RAM cell wall integrity signaling network; role in cell separation, **azole** sensitivity; required for hyphal growth; lacks orthologs in higher eukaryotes
orf19.5975	Putative zinc finger DNA-binding transcription factor; **fluconazole**-downregulated; expression regulated during planktonic growth
Hmg2	HMG-CoA reductase; enzyme of sterol pathway; inhibited by **lovastatin**; gene not transcriptionally regulated in response to **lovastatin** and **fluconazole**
ASR1	Putative heat shock protein; transcription regulated by cAMP, osmotic stress, **ciclopirox** olamine, **ketoconazole**; stationary phase enriched
YJR073c	Phosphatidylethanolamine *N*-methyltransferase of phosphatidylcholine biosynthesis; downregulation correlates with clinical development of **fluconazole** resistance; **amphotericin B**; and **caspofungin** repressed
Cor1	Putative ubiquinol-cytochrome-*c* reductase; **amphotericin B** induced; repressed by nitric oxide; protein level decreases in stationary phase cultures
Taf19	Putative TFIID subunit; mutation confers hypersensitivity to **amphotericin B**
OPT8	Possible oligopeptide transporter; induced by nitric oxide, **amphotericin B**
AGP2	Amino acid permease; hyphal downregulated; regulated upon white-opaque switching; induced in core **caspofungin** response, during cell wall regeneration, or by **flucytosine**; fungal-specific
FET31	Putative iron transport multicopper oxidase precursor; **flucytosine** induced; **caspofungin** repressed
HIP1	Similar to amino acid permeases; alkaline upregulated; **flucytosine** induced; fungal-specific (no human or murine homolog)
APT1	Adenine phosphoribosyltransferase; **flucytosine** induced; repressed by nitric oxide; protein level decreased in stationary phase yeast cultures
ARX1	Putative ribosomal large subunit biogenesis protein; downregulated during core stress response; decreased expression in response to **prostaglandins**
Ygr090w	Putative U3 snoRNP protein; decreased expression in response to **prostaglandins**; heterozygous null mutant exhibits resistance to **parnafungin**
NOG1	Putative GTPase; mutation confers hypersensitivity to 5-fluorocytosine (5-FC), 5-fluorouracil (5-FU), and tubercidin (7-deazaadenosine); decreased expression in response to **prostaglandins**
Imp4	Putative SSU processome component; decreased expression in response to **prostaglandins**

The data was taken from the Candida Genome Database (Arnaud et al., 2010). Antifungal agents are marked in bold.

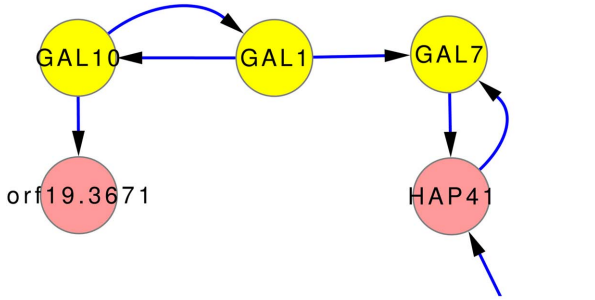


FIGURE 7 | Subnetwork GAL-genes using expression data and the full set of prior knowledge that does not contain these predicted relations. The connection of *GAL10*, *GAL1*, and *GAL7* is well studied in many yeast forms. The network predictions also contain this relationship.

The prior knowledge sources BIND, TRANS, and PPI contain much more interactions than FAC and have also more interactions in common with the *gold standard*. Even though they still have very little overlap with the *gold standard*, the values of the F-measure improve strongly. A possible explanation is that the prior knowledge supports interactions outside the *gold standard*, which afterward supports correct interactions from the *gold standard*.

This emphasizes the importance of extensive data sets to improve the performance of the algorithm. However, the combination of all four sources of prior knowledge does not result in the best performance. With an F-measure of 0.0018, the network is only slightly better than the one created with FAC. This may indicate contradicting information within the prior knowledge sources. Additional refinement concerning weighting prior knowledge with regard to its reliability and the combination of different prior knowledge sources has great potential to further improve the performance of the algorithm.

In general, we can conclude that the higher the influence of the prior knowledge, the better the results concerning the F-measure are, as depicted in **Figure 1**. But there is another conclusion, that can be seen in this Figure: the improvement with prior knowledge is stronger with smaller c-value, i.e., on smaller networks. This seems reasonable since smaller networks have a more strict constraint and the decreased penalty for interactions supported by the prior knowledge has a stronger effect.

However, the inferences correctly predict parts of the GAL-network, as can be seen in **Figure 7**. It shows, that the inference can uncover regulations even without the help of prior knowledge. None of the prior knowledge (ALL) suggest these interactions. The *gold standard* contains only one of them ($GAL1 \rightarrow GAL7$).

It should also be noted that the prior knowledge reflects the knowledge of other species, in particular *S. cerevisiae*, whereas the *gold standard* contains *C. albicans* specific knowledge. We are aware that there are substantial differences between the regulation of *C. albicans* on the one hand and *S. cerevisiae* and other model organisms on the other hand. Therefore, putting to much weight on the prior knowledge from these model

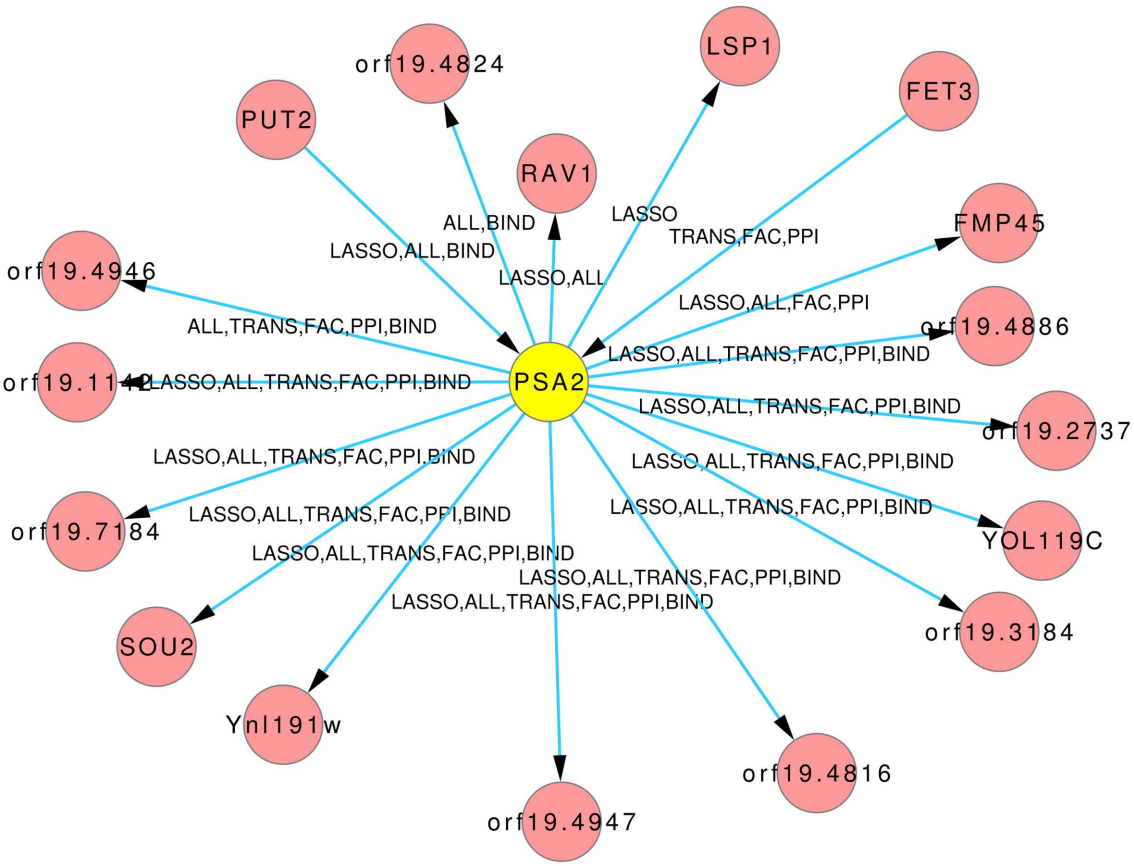


FIGURE 8 | Predicted hub *PSA2*. The labels on the edges tell which inference and prior knowledge predicted this interaction. LASSO for the LASSO inference without prior knowledge. FAC, BIND, TRANS, and PPI for the inferences with the corresponding prior knowledge source. ALL for the inference that exploited all four prior knowledge sources.

organisms can lead to false conclusions. To minimize the probability of such wrong conclusions, we use the *C. albicans* specific *gold standard* to estimate the optimal weighting of the prior knowledge.

All of the presented LASSO-based inferences outperform the models created by mutual information-based methods. CLRNET and MRNET both produced networks with a comparably high number of interactions (15,686,064 with CLRNET and 15,329,450 with MRNET). ARACNE on the other hand produced a network with 39,986. This is by far the smallest of the mutual information-based networks but still about five times larger than the LASSO-based networks. However, with 0.0009 it has a higher F-measure than those of CLRNET and MRNET, which both have a F-measure of 0.00006. It may be correct to assume a correlation between the size of the mutual information-based networks and their F-measure.

The performance evaluation of the network construction algorithm is based on a *gold standard* obtained by automatic scanning of 9,000 full-text research papers. This leads to a *gold standard* of 509 genes and 1,016 interactions. Utilizing this compendium of known interactions, we optimize the parameters of the algorithm in order to increase the performance for best results. A major performance criterion is sparseness, in order to balance comprehensiveness and interpretability of the model. We focused on optimal sparseness, in order to locate the most significant interactions and to increase reliability of the predictions. However, compared to the 6,167 genes of the genome-scale networks, this *gold standard* is still far from adequate. Therefore, the evaluation of the models by comparison to the *gold standard* may favor smaller networks.

The combination of these features with the requirement for scale-freeness is a novel approach. As this is also true for most biological networks and therefore a requirement for a reasonable topological analysis to uncover hubs. Since hubs are of great interest as potential drug targets or biomarker for the development of novel therapies against fungal infections, we concentrated our effort on such a topological analysis and uncovered a list of hubs with many not yet described. Further investigation in this field is still required and continuous improvements in the available data will also enhance the predictive power of our approach.

To further check causality of the predicted gene-to-gene relations, the concept of Granger causality modeling could be applied as proposed by Shojaie and Michailidis (2010) by truncating

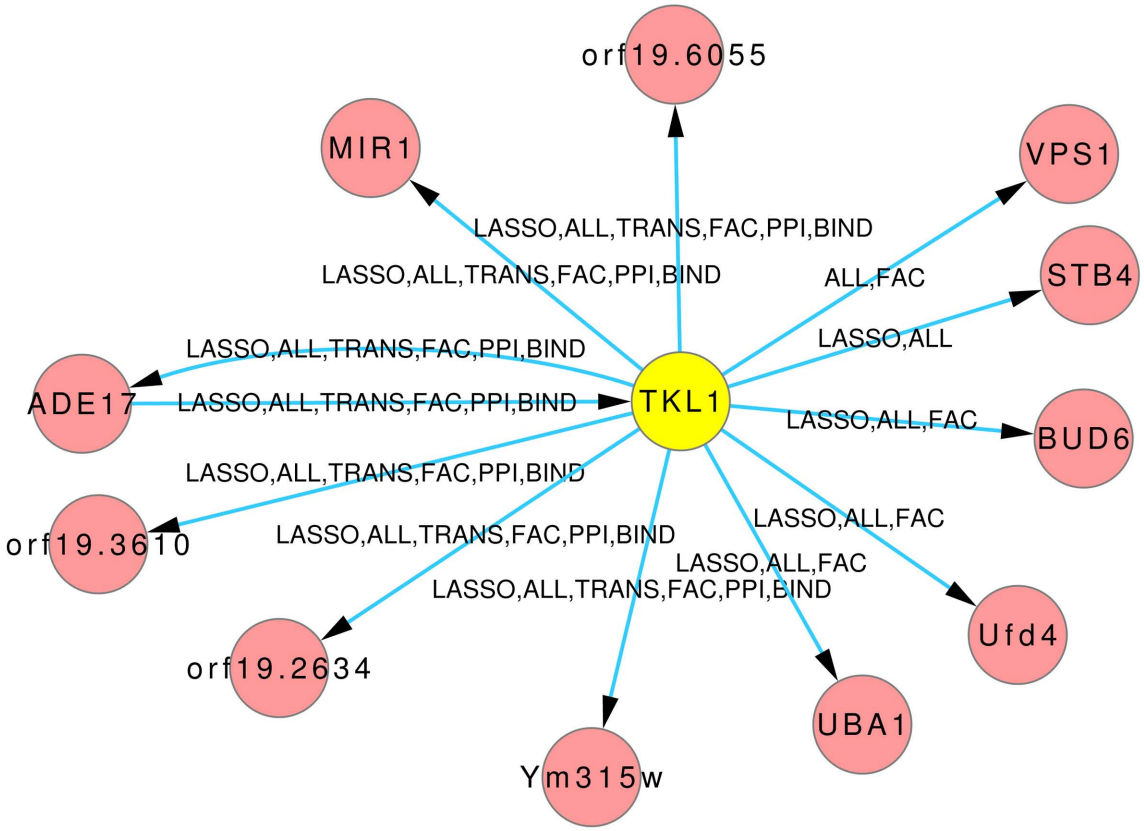


FIGURE 9 | Predicted hub *TKL1*. The labels on the edges tell which inference and prior knowledge predicted this interaction. LASSO for the LASSO inference without prior knowledge. FAC, BIND, TRANS, and PPI for the inferences with the corresponding prior knowledge source. ALL for the inference that exploited all four prior knowledge sources.

LASSO penalty. However, this approach requires time series data whereas the data set analyzed in the present work comprises both, time series and steady state data under different conditions.

We applied our approach to the non-model organism *C. albicans* since there is still little known about this human pathogenic fungus. However, our approach is not limited to *C. albicans* and can be applied to other organisms, where little knowledge is available, as well.

ACKNOWLEDGMENTS

Robert Altwasser and Jörg Linde were supported by the excellence graduate school Jena School for Microbial Communication (JSMC).

REFERENCES

Arnaud, M. B., Costanzo, M. C., Skrzypek, M. S., Shah, P., Binkley, G., Lane, C., Miyasato, S. R., and G, S. (2010). *Candida Genome Database*. Available at: http://www.candidagenome.org/

Barabási, A.-L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113.

Buyko, E., Faessler, E., Wermter, J., and Hahn, U. (2011). Syntactic simplification and semantic enrichment – trimming dependency graphs for event extraction. *Comput. Intell.* 27, 610–644.

D'Enfert, C., and Hube, B. (2007). *Candida: Comparative and Functional Genomics*. Norfolk: Caister Academic Press.

Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene expression omnibus: Ncbi gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210.

Faith, J. J., Hayete, B., Thaden, J. T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J. J., and Gardner, T. S. (2007). Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* 5, e8. doi:10.1371/journal.pbio.0050008

Gustafsson, M., Hörnquist, M., and Lombardi, A. (2004). Large-scale reverse engineering by the lasso. Available at: http://arxiv.org/abs/q-bio/0403012v1

Gustafsson, M., Hörnquist, M., and Lombardi, A. (2005). Constructing and analyzing a large-scale gene-to-gene regulatory network lasso-constrained inference and biological validation. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 2, 254–261.

Guthke, R., Kniemeyer, O., Albrecht, D., Brakhage, A. A., and Möller, U. (2007). Discovery of gene regulatory networks in *Aspergillus fumigatus*. *Lect. Notes Bioinform.* 4366, 22–41.

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

Han, J.-D. J., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J. M., Cusick, M. E., Roth, F. P., and Vidal, M. (2004). Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430, 88–93.

He, X., and Zhang, J. (2006). Why do hubs tend to be essential in protein networks? *PLoS Genet.* 2, e88. doi:10.1371/journal.pgen.0020088

Hecker, M., Goertsches, R. H., Engelmann, R., Thiesen, H.-J., and Guthke, R. (2009a). Integrative

modeling of transcriptional regulation in response to antirheumatic therapy. *BMC Bioinformatics* 10, 262. doi:10.1186/1471-2105-10-262

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

Hube, B. (2004). From commensal to pathogen: stage- and tissue-specific gene expression of *Candida albicans*. *Curr. Opin. Microbiol.* 7, 336–341.

Ihmels, J., Bergmann, S., Berman, J., and Barkai, N. (2005). Comparative gene expression analysis by a differential clustering approach: application to the *Candida albicans* transcription program. *PLoS Genet.* 1, e39. doi:10.1371/journal.pgen.0010039

Johnston, M. (1987). A model fungal gene regulatory mechanism: the gal genes of *Saccharomyces cerevisiae*. *Microbiol. Rev.* 51, 458–476.

Kim, J.-D., Ohta, T., Pyysalo, S., Kano, Y., and Tsujii, J. (2011). Extracting biomolecular events from literature – the bionlp'09 shared task. *Comput. Intell.* 27, 513–540.

Leclerc, R. D. (2008). Survival of the sparsest: robust gene networks are parsimonious. *Mol. Syst. Biol.* 4, 213.

Linde, J., Buyko, E., Altwasser, R., Hahn, U., and Guthke, R. (2011). *Full-Genomic Network Inference for Non-Model Organisms: A Case Study for the Fungal Pathogen Candida albicans*. Paris: WASET.

Linde, J., Wilson, D., Hube, B., and Guthke, R. (2010). Regulatory network modelling of iron acquisition by a fungal pathogen in contact with epithelial cells. *BMC Syst. Biol.* 4, 148. doi:10.1186/1752-0509-4-148

Lohr, D., Venkov, P., and Zlatanova, J. (1995). Transcriptional regulation in the yeast gal gene family: a complex genetic network. *FASEB J.* 9, 777–787.

Margolin, A. A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Favera, R. D., and Califano, A. (2006). ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 7(Suppl. 1), S7. doi:10.1186/1471-2105-7-S1-S7

Meyer, P. E., Kontos, K., Lafitte, F., and Bontempi, G. (2007). Information-theoretic inference of large transcriptional regulatory networks. *EURASIP J. Bioinform. Syst. Biol.* 2007, 8–8.

R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rokas, A., and Hittinger, C. T. (2007). Transcriptional rewiring: the proof is in the eating. *Curr. Biol.* 17, R626–R628.

Shojaie, A., and Michailidis, G. (2010). Discovering graphical granger causality using the truncating lasso penalty. *Bioinformatics* 26, i517–i523.

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432.

Stacklies, W., Redestig, H., Scholz, M., Walther, D., and Selbig, J. (2007). pcaMethods – a bioconductor package providing pca methods for incomplete data. *Bioinformatics* 23, 1164–1167.

Tibshirani, R. (1994). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* 58, 267–288.

Van Rijsbergen, C. (1979). *Information Retrieval*, 2nd edn. London: Butterworths.

Wilson, L. S., Reyes, C. M., Stolpman, M., Speckman, J., Allen, K., and Beney, J. (2002). The direct cost and incidence of systemic fungal infections. *Value Health* 5, 26–34.

Yeung, M. K. S., Tegnr, J., and Collins, J. J. (2002). Reverse engineering gene networks using singular value decomposition and robust regression. *Proc. Natl. Acad. Sci. U.S.A.* 99, 6163–6168.

Zou, H. (2006). The adaptive lasso and its oracle properties. *J. Am. Stat. Assoc.* 101, 1418–1429.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 December 2011; paper pending published: 02 January 2012; accepted: 31 January 2012; published online: 16 February 2012.

Citation: Altwasser R, Linde J, Buyko E, Hahn U and Guthke R (2012) Genome-wide scale-free network inference for Candida albicans. Front. Microbio. ***3****:51. doi: 10.3389/fmicb.2012.00051*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 02 April 2012
doi: 10.3389/fmicb.2012.00114

Multivariate analysis of flow cytometric data using decision trees

Svenja Simon[1]*†, Reinhard Guthke[1], Thomas Kamradt[2] and Oliver Frey[2]†

[1] *Research Group Systems Biology/Bioinformatics, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Jena, Germany*
[2] *Institute of Immunology, Jena University Hospital – Friedrich Schiller University Jena, Jena, Germany*

Edited by:
Franziska Mech, Leibniz-Institute for Natural Product Research and Infection Biology – Hans-Knoell-Institute, Germany

Reviewed by:
Enrico Lugli, National Institutes of Health, USA
Thomas Bley, Technische Universitaet Dresden, Germany

****Correspondence:***
Svenja Simon, Department of Computer and Information Science, University of Konstanz, Universitätsstraße 10, Box 78, 78457 Konstanz, Germany.
e-mail: simon@dbvis.inf.uni-konstanz.de

†Present address:
Svenja Simon, Data Analysis and Visualization Group, Department of Computer and Information Science, University of Konstanz, Konstanz, Germany;
Oliver Frey, Institute of Clinical Chemistry and Laboratory Medicine, Jena University Hospital – Friedrich Schiller University Jena, Jena, Germany.

Characterization of the response of the host immune system is important in understanding the bidirectional interactions between the host and microbial pathogens. For research on the host site, flow cytometry has become one of the major tools in immunology. Advances in technology and reagents allow now the simultaneous assessment of multiple markers on a single cell level generating multidimensional data sets that require multivariate statistical analysis. We explored the explanatory power of the supervised machine learning method called "induction of decision trees" in flow cytometric data. In order to examine whether the production of a certain cytokine is depended on other cytokines, datasets from intracellular staining for six cytokines with complex patterns of co-expression were analyzed by induction of decision trees. After weighting the data according to their class probabilities, we created a total of 13,392 different decision trees for each given cytokine with different parameter settings. For a more realistic estimation of the decision trees' quality, we used stratified fivefold cross validation and chose the "best" tree according to a combination of different quality criteria. While some of the decision trees reflected previously known co-expression patterns, we found that the expression of some cytokines was not only dependent on the co-expression of others *per se*, but was also dependent on the intensity of expression. Thus, for the first time we successfully used induction of decision trees for the analysis of high dimensional flow cytometric data and demonstrated the feasibility of this method to reveal structural patterns in such data sets.

Keywords: flow cytometry, cytokines, machine learning, induction of decision trees, imbalanced data, multidimensionality

1. INTRODUCTION

Flow cytometry is a fundamental technology in immunology. It allows the identification of cell populations as well as functional properties of immune cells with high speed and precision. Because of its ability to analyze thousands of cells per second, this technique is key for the study of immune cell population dynamics in the context of microbial infection or autoimmune disease. Recent advances in flow cytometry instrumentation and reagents provide researchers now with the capability to assess simultaneously multiple phenotypic and functional markers on a single cell level (Perfetto et al., 2004). Assessment of multiple phenotypic and functional markers gives the opportunity for a comprehensive single cell analysis, but the resulting data sets are quite complex. Therefore, the gap between generation of such data and our understanding of it is growing. Conventional analysis approaches are based on filtering of populations (subsets) of interest and subsequent analysis of the expression of certain markers within these populations. This, however, often neglects the multidimensionality of the data. If, for example, the expression of n markers is analyzed in a given subpopulation, the resulting dataset has n dimensions. Using color-coded representation of the third dimension, three different parameters can be displayed in a two-dimensional dot-plot (Roederer and Moody, 2008). However, analysis of data sets with more than three dimensions is heavily impaired by our limited capability to integrate information from more than three dimensions and biased by the experience of the researcher, leaving some information unexploited. By using machine learning methods commonly used in data mining, it should be possible to automate analysis, preclude the operator-introduced bias and reveal structural patterns of the data which would have been unrecognized with conventional approaches (Sachs et al., 2005). Machine learning methods learn decision rules from training data sets to classify new, unknown data sets and thereby can describe the structural patterns contained in the data. Machine learning methods are divided into supervised and unsupervised methods. For supervised learning methods a label (also referred to as a class) is known for the training data set, for instance the outcome of a medical treatment depending on clinical or laboratory parameters. For unsupervised methods a known outcome is not used. In flow cytometry a training data set can

be considered as a collection of analyzed cells with its different expression values for a set of markers (attributes) and an outcome (class label). For instance, a class could be defined by the absence or the presence of a given marker which could be dependent on the expression of the other measured markers. Thus, by supervised machine learning methods the dependency of the expression of one marker on the expression of other markers can be analyzed. A supervised method with an easily understandable graphical representation is the induction of decision trees. Basically, decision trees represent the structural patterns of the data. Beginning at the root, the data set is split at each node according to a marker and a split value is assigned to this node. At each node, the marker and the split value is chosen to maximize a split criterion. In this way the cells are routed down the tree and reach a specific leaf, which gives a classification. The aim of our study was to test the feasibility of this approach for the identification of structural patterns in flow cytometric data. We used data sets from experiments where the expression of six cytokines in antigen-specific T helper (Th) cells from a murine arthritis model were analyzed (Schubert et al., 2004; Frey et al., 2010b). Our results show that the supervised machine learning method induction of decision trees is a versatile tool for identification of structural patterns in multidimensional data obtained by flow cytometry.

2. MATERIALS AND METHODS

2.1. DATA GENERATION AND ACQUISITION

We used a data set from intracellular cytokine staining of activated Th cells (Frey et al., 2011b). The cells were stained and analyzed for the expression of six cytokines as described in the following. DBA/1 mice in the age of 6–12 weeks were subcutaneously immunized at the base of the tail with recombinant glucose-6-phosphate isomerase (G6PI) in an emulsion containing also Freunds complete adjuvant as described (Bruns et al., 2009; Frey et al., 2010a,b, 2011a,b). At day 21 after immunization, the draining lymph nodes (inguinal, axillary, paraaortic) were aseptically removed and prepared to a single cell suspension. In addition, beside the wild type DBA/1 mice (WT) also interferon-gamma (IFN-γ) receptor knock-out DBA/1 mice (KO) were analyzed (Frey et al., 2011b) and we performed the analyses also for other time points (day 9 and day 21 after immunization). Altogether, we studied four conditions: WT-day 21 (standard condition) as well as the additional conditions WT-day 9, KO-day 21, and KO-day 9. The additional conditions have only been applied for the results shown in **Figures 10–12** for a comparative study and to investigate the robustness of the results against experimental variations. For detection of antigen-specific cells by their CD154 expression (Kirchhoff et al., 2007), cells (1×10^7/ml in a 48 well plate) were restimulated with 20 μg/ml G6PI. Control samples were left unstimulated. The total restimulation time was 6 h and Brefeldin A (Sigma) at 5 μg/ml was added to all samples for the last 4 h to block cytokine secretion and to stabilize CD154 expression. These assay conditions have been determined to be optimal for a simultaneous detection of CD154 expression and cytokine production in antigen-specific CD4+ T helper cells. At the end of the restimulation period, cells were washed with ice-cold phosphate-buffered saline (PBS) and incubated with the fixable amine-reactive Aqua viability stain (Invitrogen) for 30 min on ice, fixed with 2% paraformaldehyde in PBS and permeabilized with 0.5% Saponin/0.5% BSA/0.02% NaN3 in PBS. Non-specific binding of antibodies was blocked by preincubation of the cells with anti-CD16/32 (2.4G2) and rat IgG (both at 5 μg/ml) for 8 min, followed by staining with fluorochrome-conjugated mAbs against CD4, CD154, GM-CSF, TNF-α, RANKL, IL-2, IL-17, and IFN-γ (all from BD, eBiosciences, Biolegend, or Miltenyi Biotech). For optimal staining results all antibodies were properly titrated and the binding of the anti-CD4 antibody to fixed and permeabilized cells was verified. After an additional washing step 0.5% Saponin/0.5% BSA/0.02% NaN3 in PBS, cells were resuspended in 0.5% BSA/0.02% NaN3 in PBS and measured within 3 h after staining. Cell analysis was performed on a BD LSR II flow cytometer equipped with 405, 488, and 633 nm laser lines and standard filter sets, except additional detectors for detection of Alexa-700 (red laser, 685 nm long-pass and 710/50 band-pass filters) and Qdot655 (violet laser, 635 nm long-pass and 670/14 band-pass filters, not used for this study). For fluorescence standardization and monitoring of the instrument performance, the cytometer setup, and tracking module of the BD FACSDiVa was used. Compensation for spectral overlap of the fluorochromes was done with the use of singly stained BD CompBeads and a compensation matrix was calculated using the BD FACSDiVa software. At least 1.5 million events were acquired.

2.2. DATA PRE-PROCESSING

For analysis, data were exported as FCS3.0 files. Further preprocessing was done using FlowJo 8.1.1 (Treestar Inc., Ashland, Oregon). For identification of antigen-specific cells the following progressive filtering (also referred to as gating) strategy was used: events were first filtered on a FSC-A vs. FSC-W plot (forward scatter pulse area vs. pulse width) for the exclusion of doublets. Thereafter, a filter was set on lymphocyte in FSC-A vs. SSC-A plot (side scatter pulse area), followed by the exclusion of aqua$^+$ dead cells. Subsequently, cells were filtered for CD4-positive events. The small compartment of antigen-specific T cells was identified by their expression of CD154. The filter for CD154 expression was set using unstimulated control samples. Only the filtered CD154$^+$ events were than exported into a new data file. These resulting data files containing the events from single animals were than concatenated into a single file containing data from four mice and were used for further analyses. This electronic pooling of the data was performed in order to have a sufficient number of cells for further analysis by machine learning methods. The distributions of MFI values for each cytokines were compared between the four biological replicates which were pooled for further analysis. By visual comparison the variance was assessed.

2.3. DATA ANALYSIS

The data analysis was performed using the programming language and statistical software R (R Development Core Team, 2009). To read in the pre-processed FACS data set the R/Bioconductor package *flowCore* was used (Data File Standards Committee of the Society for Analytical Cytology, 1990; Gentleman et al., 2004; Hahne et al., 2009). The intensity of the staining was measured as mean fluorescence intensity (MFI) value. To distinguish between specifically stained cells and background fluorescence

we used appropriate controls, including unstimulated samples and fluorescence-minus-one controls (Hulspas et al., 2009). We defined this MFI value as a cut-off and considered a cell as positive for a given cytokine if their MFI exceeded this cut-off value and cells with a MFI values below this cut-off value are referred to as cytokine-negative cells.

For the induction of decision trees we used the WEKA java implementation J48 of the C4.5 algorithm of Quinlan in revision 8 (Quinlan, 1993; Witten and Frank, 2005) through an interface provided by the R package *RWeka* (Hornik et al., 2009). Since the outcome of the induction of a decision tree is highly dependent on the parameter setting and since it is not known which parameters are the best, our approach is to build decision trees for different parameter settings, to compare their quality and choose the best tree according to some quality criteria. We therefore varied the following parameters:

- pruning (reduced error pruning vs. heuristic or no pruning)
- minimum size of data in each leaf
- exclusion of one or more cytokines for the induction of decision trees (like a brute-force way of feature selection)
- exclusion of cytokine-negative cells

Cytokine-negative cells are cells which are not stained specifically for any cytokine. Their MFI values for each cytokine are under the experimentally determined threshold for cytokine production. Thus they produce none of the measured cytokines. We created decision trees with and without cytokine-negative cells to test if the presence of these cells has an effect on the quality of the decision trees.

Due to the fact that decision trees work best with balanced data (Weiss and Provost, 2003; Sun et al., 2007), the data was also weight according to the class probabilities (*CostSensitiveClassifier* Class of WEKA).

To choose one tree we applied different quality criteria on stratified fivefold cross validation results in a stepwise manner:

1. choose all trees whose *G*-mean (definition see below) is maximal 5% below the best *G*-mean
2. choose all trees of these with ROC AUC greater than 80%, if values greater than 80% exist
3. choose all trees of these whose *F*-mean (definition see below) is maximal 5% below the best *F*-mean of these
4. choose the smallest of these trees

G-mean is the geometric mean of the TP rate and the TN rate. TP (true positive) and TN (true negative) rate is the proportion of the positive cells which were correctly classified as positive (also called sensitivity and recall) or as negative (also called specificity), respectively. We used both values, since we assess both as equally important, and the geometric mean, since this mean gives the smaller value more weight than the arithmetic mean. This helps to filter out trees where only either the TP rate or TN rate is good. The name *G*-mean was introduced by Kubat et al. (1998).
G-mean is the geometric mean of the TP rate and the TN rate. TP (true positive) and TN (true negative) rate is the proportion of the positive cells which were correctly classified as positive (also called sensitivity and recall) or as negative (also called specificity), respectively. We used both values, since we assess both as equally important, and the geometric mean, since this mean gives the smaller value more weight than the arithmetic mean. This helps to filter out trees where only either the TP rate or TN rate is good. The name *G*-mean was introduced by Kubat et al. (1998). For the same reason we used the geometric mean of the F-measures. The F-measure combines the precision and the recall, where precision is the fraction of cells correctly classified as positive. We calculate the F-measure for the class positives and also an F-measure for the class negatives and determined the geometric mean of both, called *F*-mean.

Decision trees route data (cells) down the tree, starting at the top (root), ending at the colored boxes (leaves). Leaves classify the cells either as positive (green) or negative (red). This classification can be correct or incorrect. To decide which route is taken by a cell, the attribute values (MFI values of the other cytokines) are compared to the split values at the branch. The names in the white boxes (inner nodes) state the attribute to which this split value has to be applied to. Additional each node contains the number of cells in the data set, which were routed to this inner node. To ease the analysis of the decision trees and to avoid many look ups at the raw data sets we included some additional information in the raw decision trees. First, we visualized the experimentally determined cut-off values of a cytokine as a further attribute at the corresponding node to allow for a simple assessment between the split value and the cut-off value. Cut-off values are colored in green if the split value is close to the cut-off value, and red or blue if the split value is below or above the cut-off value, respectively. Second, we visualized in each leaf the proportion of positive (or negative) cells in this leaf on all positive (or negative) cells. This allows for an easy assessment of the importance of a particular leave in the overall classification.

3. RESULTS

Figure 1 shows the filtering of the raw data. The resulting data set of antigen-specific activated T helper (Th) cells contains measurements for the six cytokines TNF-α, RANKL, IL-17, IL-2, IFN-γ, and GM-CSF. The intensity of the staining for this cytokines was measured as mean fluorescence intensity (MFI) value. The variance of the MFI values for the same conditions was found to be low as assessed for four biological replicates (**Figure 2**). As shown in **Figure 3A** a huge proportion of antigen-specific activated Th cells (71%) produced TNF-α, followed by expression of RANKL, IL-17, IL-2, IFN-γ, and GM-CSF. Of note, the sum of these cells exceeds 100% because of the co-expression of two or more cytokines which is shown in **Figure 3B**. We found a strong co-expression of TNF-α together with GM-CSF, IFN-γ, and IL-17 (first row in **Figure 3B**), whereas for IL-2 and RANKL this association was lower. There was also a huge proportion of cells (58.65%) which produced only TNF-α and no other cytokines. A strong association was also seen between GM-CSF, IFN-γ, and RANKL (58.25 and 49.51%, respectively, column GM-CSF in **Figure 3B**). In other cases the co-expression was surprisingly low, for instance only 2.58% of the IL-2 positive cells also co-produced GM-CSF. However, such pairwise comparisons are limited because they neglect the possibility that the expression of a certain cytokine could be dependent of

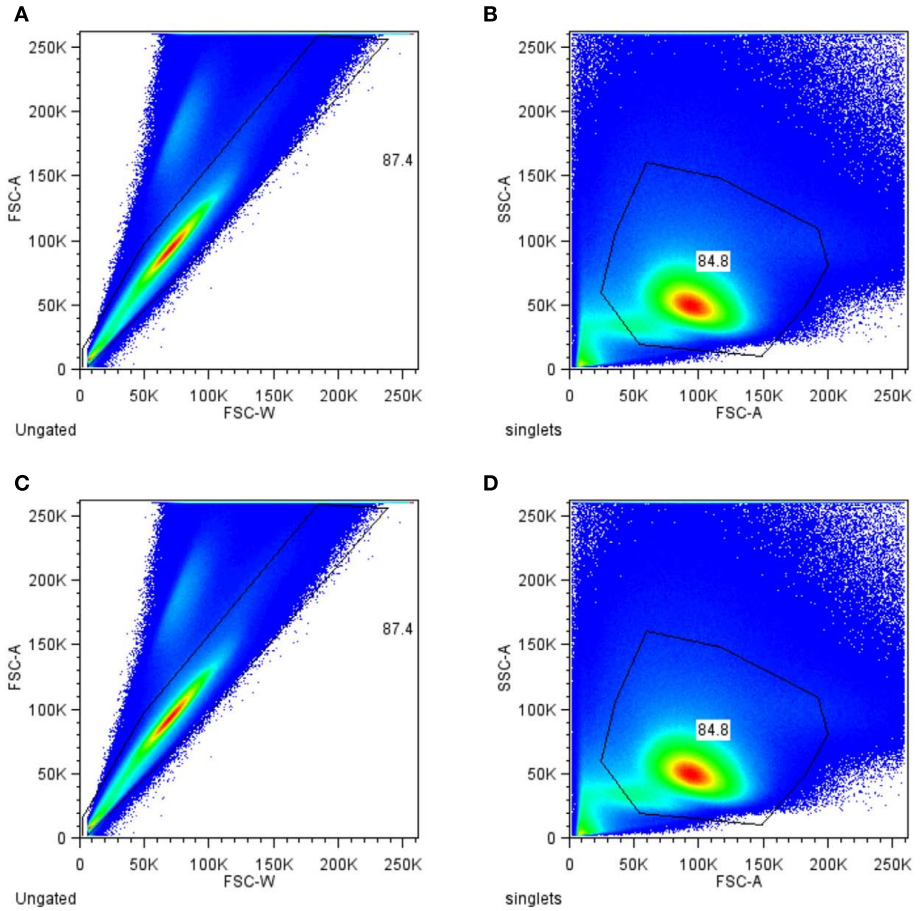


FIGURE 1 | Progressive filtering (gating) strategy for identification of antigen-specific T helper cells. Total events measured were filtered (gated) for single cell **(A)** and for lymphocytes **(B)**. To exclude dead cells aqua dye positive cells were excluded from further analysis **(C)**. The remaining cells were further selected for CD4 **(D)**.

the co-expression of two or more cytokines simultaneously. To study the dependencies of the co-expression patterns of cytokines we formed a data set for each cytokine, with the expression of this cytokine as the class and with the measurements of the five remaining cytokines as attributes. We considered cells as positive for a given cytokine according to the cut-off value defined by our biological and staining control samples. Since the data sets were highly imbalanced (for instance only 1.4% of all cells produce GM-CSF, see **Figure 3A**) we weighted the data according to the class probabilities (extitCostSensitiveClassifier Class of WEKA). The resulting data sets were used to build decision trees to test if the graphical representation of the decision trees can reveal structural patterns in the data. This approach yielded 13,392 different decision trees for each given cytokine. We assumed that if a decision tree gives a good classification, the structure of the decision tree reflect reasonable patterns of co-expression. We therefore aimed at the identification of the best decision tree out of the 13,392 generated for each cytokine.

With the approach explained in the section 3 we chose the "best" tree for every cytokine out of the many trees which were build with different parameter settings.

For instance, the tree of IL-2 (**Figure 4**) is to be read as followed:

For this decision only cells with expression of at least one cytokine were used. The tree thus begins with 2590 cells at the root. Based on the MFI value of the root attribute TNF-α the 2590 cells are routed down the tree. If the TNF-α MFI value of a cell is equal or below the split value 2390, the cells are routed to the inner left node. Otherwise the cells are routed down to the right leaf which classifies 1859 cells as IL-2 negative. This classification is correct for 1611 cells, they are true negative (TN). For 248 cells this classification is wrong, they are false negative (FN). Further information in this leave show that the leaf captures 78.66% of the IL-2 negative cells and 45.76% of the IL-2 positive cells. Due to the imbalance of the data set there is a high percentage of the IL-2 positive cells in this leaf, but only a small number of false negative cells. 20% of all measured cells produce IL-2 (see **Figure 3A**). The inner node splits the cells based on the MFI value of RANKL and routes them to the right leaf if the RANKL MFI value is above the split value of 1817. This leaf classifies the cells again as IL-2 negative. If the RANKL MFI value is equal or below the split value, they are routed to the left leaf, which classifies these cells as

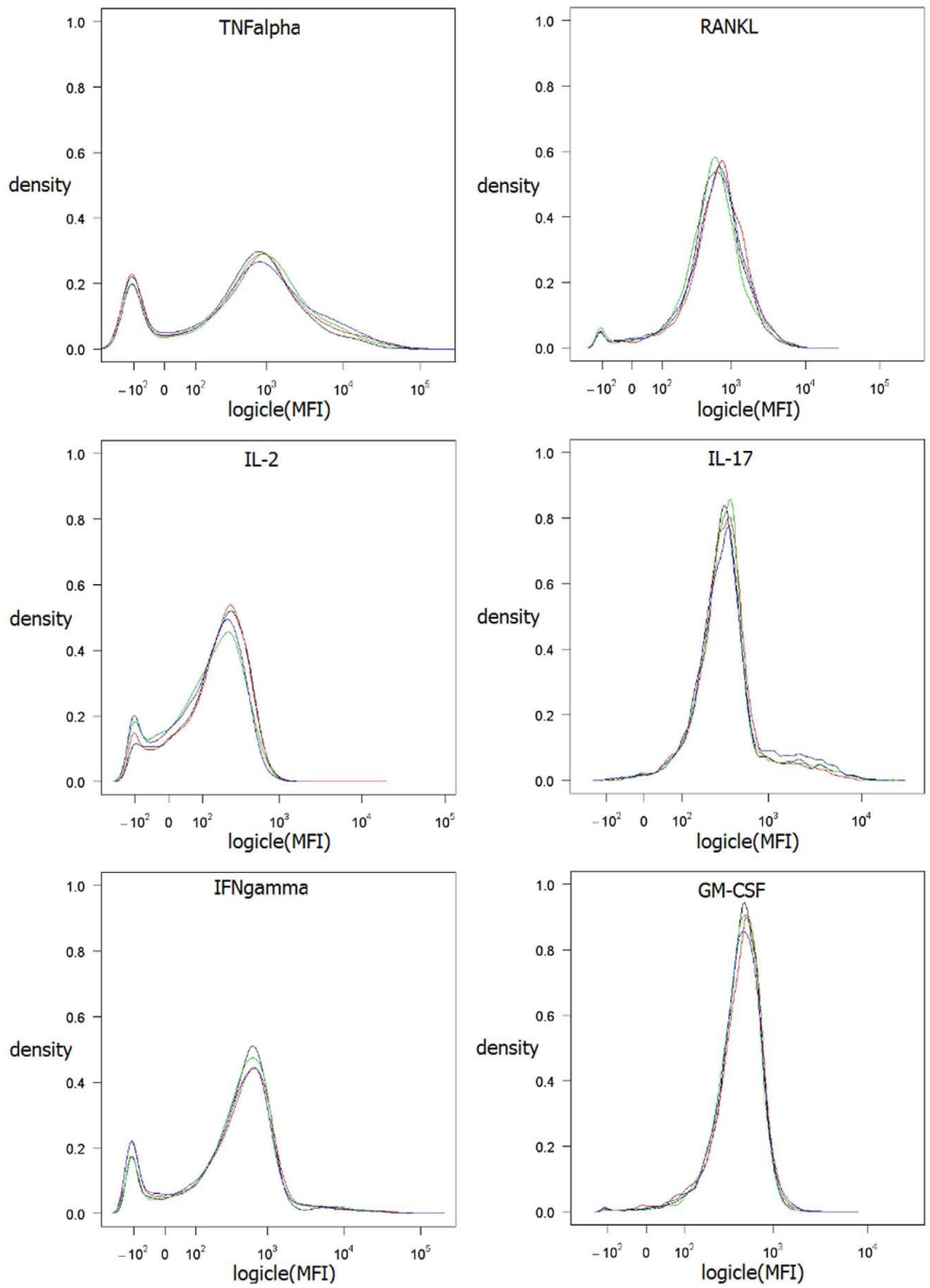


FIGURE 2 | Kernel density estimation of MFI values of the four biological replicates. Method *density* from R (R Development Core Team, 2009) was used with default parameter settings. Data values are *logicle* transformed according to Parks et al. (2006) using *flowCore* (Hahne et al., 2009).

IL-2 positive. This leaf contains 356 cells, of which 251 are classified correctly as positive (true positive – TP) and 105 are wrongly classified as positive (false positive – FN). This leaf than contains 46.31% of the IL-2 positive cells and 5.13% of the IL-2 negative cells. Furthermore, the split values of TNF-α and RANKL capture all cells which are negative for TNF-α and RANKL but positive for IL-2. Therefore, the split values have to be so low that they do not route cells down to this leaf which express RANKL or TNF-α. This is indeed true since the split values are very close to the experimentally determined cut-off values of these cytokines. The experimentally determined cut-off values are shown below the root and inner nodes, the green color indicates that the split values are close to the cut-off value. Finally, the proportion of RANKL and TNF-α negative cells in this leaf exceeds 100%. Thus, it is clear

A

	% of all
TNFalpha	71
RANKL	24
IL-2	20
IL-17	14
IFNgamma	5.1
GM-CSF	1.4

B

	GM-CSF	IFNgamma	IL-17	IL-2	RANKL	TNFalpha
TNFalpha	75.73	77.09	68.63	45.94	40.92	**58.65**
RANKL	49.51	25.61	29.74	19.37	**43.13**	13.91
IL-2	13.59	11.32	14.05	**44.1**	16.59	13.37
IL-17	31.07	11.32	**19.28**	7.93	14.38	11.28
IFNgamma	58.25	**11.05**	13.73	7.75	15.01	15.36
GM-CSF	**1.94**	16.17	10.46	2.58	8.06	4.19

FIGURE 3 | Cytokine expression by antigen-activated T helper cells. Expression of the cytokines TNF-α, RANKL, IL-2, IL-17 IFN-γ, and GM-CSF was measured in lymph node cells as described in section 2. Data depict the proportion of activated cells (identified by expression of the marker CD154) that produce the indicated cytokine **(A)**. As the cells can express multiple cytokines simultaneously, pair-wise co-expression patterns of cytokines are shown in **(B)**. The primary cytokines are listed from left to right as the column names. For each primary cytokine the percentage of cells co-expressing one of the other cytokines (listed as row names) are shown. As an example GM-CSF expressing cells often co-express TNF-α (75.73% of all GM-CSF expressing cells) but only rarely IL-2 (13.59%). (Green background = low co-expression, red background = high co-expression).

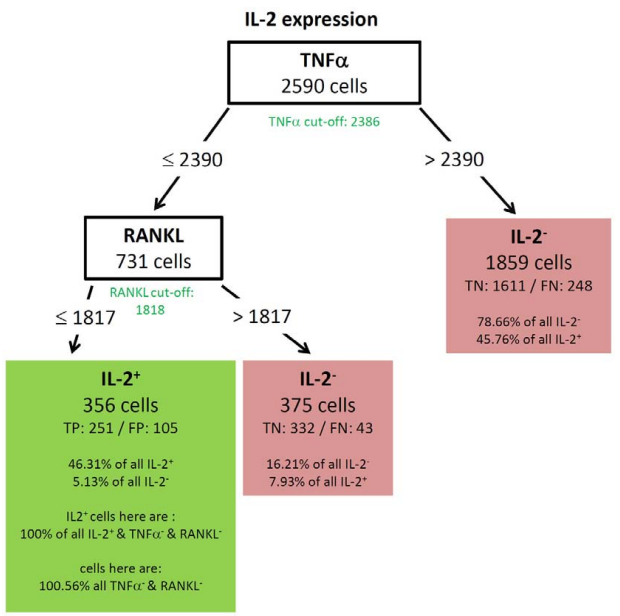


FIGURE 4 | Best decision tree for the classification of cells as positive or negative for IL-2 expression. Cells are classified based on the MFI values of first TNF-α and second RANKL. The green colored cut-off values below the inner nodes, indicate that the split values are near the cut-off values. Therefore, these nodes dived the cells in TNF-α (respectively RANKL) negative and positive cells. The cells are routed to the right leaf if they do not express TNF-α. This leaf captures 78.66% of all IL-2 negative cells, but also 45.76% of the IL-2 positive cells. The leaf in the middle classifies cells as IL-2 negative, which do not express TNF-α but RANKL. The left node classifies cells as IL-2 positive and captures 46.31% of the IL-2 positive cells. It is stated that this node contains 100.56% of the TNF-α and RANKL negative cells. Thus, also cells which are positive for TNF-α and/or RANKL are in this leaf. This is due to the split values of TNF-α. The split value is slightly above the real cut-off, therefore a few TNF-α positive cells have been routed to the left and are therefore contained in the leftmost leaf. Since the split value of RANKL is below the cut-off value, no RANKL positive cells can be contained in the leftmost leaf.

that there are few cells in this leave which are positive for TNF-α, RANKL, or both. A closer comparison of the split values and cut-off values reveal that these cells can only be TNF-α and not RANKL positive, since the split value of TNF-α is slightly above the cut-off value, but the split value of RANKL is slightly below the cut-off value.

All resulting decision trees (besides the tree of IL-17) were of sufficient quality to reveal meaningful structural patterns. This implies that there are associations between the expressions of different cytokines. An interesting common finding for the decision trees for TNF-α (**Figure 5**) and IL-2 (**Figure 4**) was the fact that the chosen split thresholds of all used cytokines (RANKL, IL-2 respectively TNF-α, RANKL) were close to the experimentally determined cut-off value of these cytokines. These finding suggests that the expression (or non-expression) of TNF-α and IL-2 depends on if the other cytokines are expressed or not. Interestingly, there was an inverse relationship between the cytokines: no expression of RANKL and IL-2 classified cells as positive for TNF-α (**Figure 5**). Similarly, no expression of TNF-α and RANKL classified cells as positive for IL-2 (**Figure 4**). One obvious reason for this classification is that TNF-α and IL-2-expressing cells have a high proportion of cells producing only a single cytokine (see **Figure 3B**). While the IL-2-expressing Th cells contain 44.1% single producers (**Figure 3B**) by bivariate analysis, our multidimensional analysis classified 46.31% of all IL-2 positive cells into the left leave of the decision tree (**Figure 4**). These cells do neither produce TNF-α nor RANKL and can therefore be considered as IL-2 single producers. We therefore can conclude that only the IL-2 single producers are classified correctly. However the decision tree can not reveal patterns in the IL-2 positive cells which are co-expressed with other cytokines. The TNF-α tree (**Figure 5**) has a similar structure as the IL-2 tree (**Figure 4**). Cells are classified as TNF-α positive if they neither produce RANKL nor IL-2. Unlike in the IL-2 tree, the TNF-α positive leave does not contain only TNF-α single producers (74.87% TNF-α positive cells, **Figure 5** vs. 58.65% TNF-α single producers in **Figure 3B**). We therefore conclude from the two trees for cytokine expression with a high percentage of single producers that the decision trees could reveal this pattern. Furthermore, other subsets with a high percentage of single producers were used to filter out cells negative for the cytokine of interest. Therefore, the decision trees detect nearly exactly the experimentally determined cut-off values of these

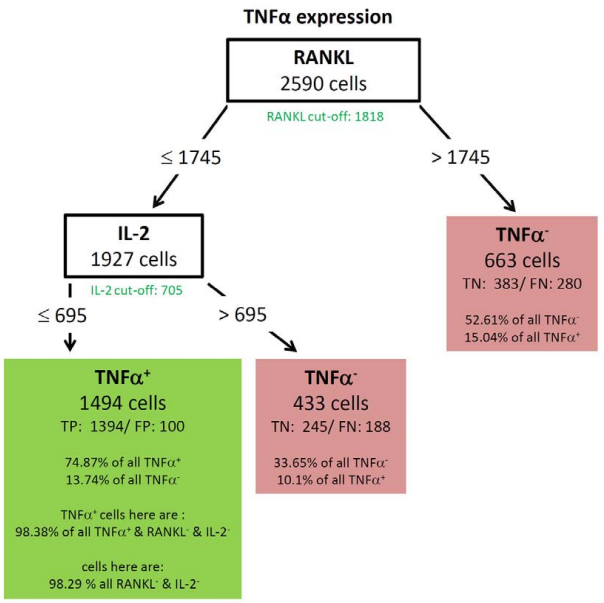


FIGURE 5 | Best decision tree for the classification of cells as positive or negative for TNF-α expression. Cells are classified based on the MFI values of first RANKL and second IL-2. The green colored cut-off values below the inner nodes, indicate that the split values are close to the cut-off values. Therefore, these nodes dived the cells in RANKL (respectively IL-2) negative and positive cells. The cells are routed to the right leaf if they do not express RANKL. Cells of this leaf are classified as negative and this leaf capture 52.61% of all TNF-α negative cells, and only classifies 15.04% of the TNF-α positive cells incorrectly. The leaf in the middle classifies cells as TNF-α negative which do not express RANKL, but do express IL-2. The left node classifies cells as TNF-α positive and captures 74.87% of the TNF-α positive cells. This node contains not all RANKL and IL-2 negative cells, but 98.38% of them. Since the split values of RANKL and IL-2 are slightly below the real cut-off a few RANKL and IL-2 positive cells have been routed to the right and are therefore not contained in the leftmost leaf.

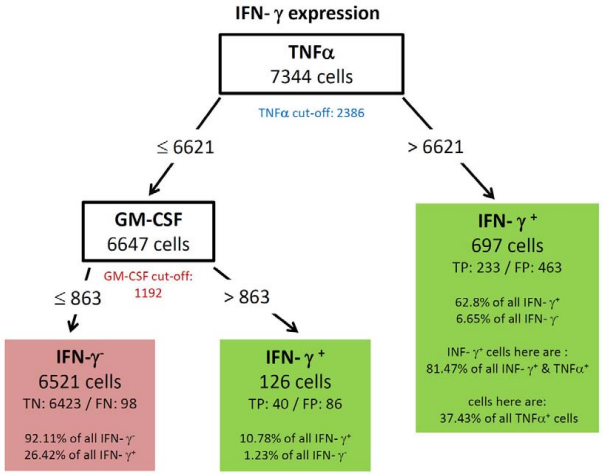


FIGURE 6 | Best decision tree for the classification of cells as positive or negative for IFN-γ expression. Cells are classified based on the MFI values of first TNF-α and second GM-CSF. The blue colored TNF-α (red colored GM-CSF) cut-off value indicates that the split value is high above (below) the cut-off value. Therefore, these nodes dived the cells not just in TNF-α (respectively GM-CSF) negative and positive cells. Only a proportion of the cells which express TNF-α are routed to the right leaf. Contained in this leafs are only 37.43% of all TNF-α positive cells due to this high split value compared to the cut-off value. This leaf captures 62.8% of the IFN-γ positive cells, these are 233 cells (TP – true positive). The leaf also contains the information that this leave captures 81.47% of all IFN-γ and TNF-α positive cells. The leaf in the middle classifies cells as IFN-γ positive and captures 10.78% of all IFN-γ positive cells. The left node classifies cells as IFN-γ negative, thus this leaf wrongly classifies 26.42% of the IFN-γ positive, but captures 92.11% of all IFN-γ negative cells.

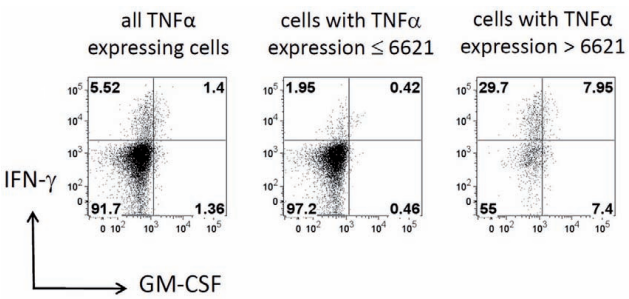


FIGURE 7 | Cells with high TNF-α expression are enriched for IFN-γ expressing cells. Dot plots show the expression of IFN-γ vs. GM-CSF in the same data set for all cells with expression of TNF-α (left), for cells with expression of TNF-α below the split value 6621 given by the decision tree shown in **Figure 6** (middle) and for cells with high TNF-α expression (right). Horizontal and vertical lines represent the experimentally determined cut-off values that define the expression for these cytokines as estimated by controls. The cells in the lower left quadrant produce neither IFN-γ nor GM-CSF and the cells in the upper right quadrant produce both. Numbers in quadrants give the percentage of the cells in the respective quadrant.

cytokines. RANKL (tree not shown) also had a high percentage of single producers. We thus expected a tree with the same structure like for TNF-α and IL-2. Compared to these easy and compact trees, the RANKL decision tree was quite complex, however it could be pruned to the same structure like the IL-2 and TNF-α tree (not shown). This pruning only slightly impaired the classification and resulted in a tree with TNF-α as root and IL-2 as next split attribute. As for RANKL and TNF-α the split values were very close to the experimentally determined cut-off values. Cells were classified as RANKL positive if TNF-α and IL-2 were not expressed and classified as RANKL negative if one of them was expressed. Other decision trees (**Figures 6** and **8**) had split values highly above the experimentally determined cut-off values. These high split values also revealed some biologically relevant information. As an example, the tree for IFN-*ma* (**Figure 6**) was splitted into IFN-γ positive and negative cells by the expression of TNF-α with an MFI of about 6621. Due to this high split value, the node to the right (MFI for TNF-α > 6621) only contained 37.43% of all TNF-α positive cells. However, this node contained 81.47% of all TNF-α and IFN-γ positive cells. Given that the expression of TNF-α started above an MFI of 2368 (as measured by controls), it can be concluded that especially a high expression of TNF-α is associated with the expression of IFN-γ. Routing down the tree of IFN-γ further, the next node contained GM-CSF expression

as split attribute for cells with a TNF-α expression below 6621 (**Figure 6**). However, the split value of 863 for GM-CSF expression was far below the threshold for GM-CSF positive cells as estimated by biological controls (MFI > 1192). This lead to the classification IFN-γ negative for cells below this threshold (TN rate is 92.11%, FN 26.42%) and a classification as IFN-γ positive for cells above this threshold (TP rate 10.78%). Since the TP rate was only around 11% and the split value did not correspond with the true cut-off of GM-CSF, it can be concluded that IFN-γ expression is probably only loosely associated with the expression of GM-CSF. Most of the IFN-γ negative cells do not express GM-CSF since the true negative (TN) rate is high.

To further confirm the relationship between the production of IFN-γ and TNF-α, we filtered the data on TNF-α high (MFI > 6621) and TNF-α low (MFI < 6621) cells. As shown in **Figure 7**, TNF-α-high cells were highly enriched for IFN-γ producing cells (5.52 vs. 29.7%; left vs. right plot in **Figure 7**), while TNF-α-low cells are depleted of IFN-γ producing cells (5.52 vs. 1.95%; left vs. middle plot in **Figure 7**). Strikingly, although segregation into TNF-α high and low populations did also enrich GM-CSF producers, the proportion between GM-CSF single positive cells and GM-CSF, IFN-γ double positive cells was similar for unfiltered TNF-α low and TNF-α high cells (**Figure 7**). Thus, the co-production of GM-CSF and TNF-α seems to be independent of IFN-γ production.

GM-CSF showed a complex decision tree with many leaves (**Figure 8**). Comparable to IFN-γ (**Figure 6**), the first split was at a high level of TNF-α production, but captured most of the GM-CSF and TNF-α positive cells (80.77%). Further splits were at IFN-γ and RANKL expression again with split values above the cut-off value of these cytokines. Cells with expression of TNF-α below 7213, IFN-γ below 2458, and RANKL below 2729 were classified as GM-CSF negative cells (TN = 88.84% in **Figure 8**). The decision tree of IL-17 (**Figure 9**) did not provide useful patterns, because the split values were always below the real cut-off values of these cytokines. Nevertheless the classification is quite good.

To validate the robustness of identified patterns we generated decision trees not only for the standard condition (WT-day 21) as presented up to this point, but also for three additional experimental conditions (WT-day 9, KO-day 21, and KO-day 9). The trees for IFN-γ are almost identical as shown in **Figures 10A–C** and **6**. The four trees demonstrate that high TNF-α production is required for IFN-γ expression. Quite similar, the expression of GM-CSF is the most important split criterion for the expression of IL-17 as shown in **Figure 9** for the standard condition and the three trees shown in **Figures 11A–C** for the additional conditions. In addition, also the pattern found for the standard condition WT-day 21 (**Figure 9**) is very similar to that found for KO-day 21 shown in **Figure 11C**. These two trees show that IL-17 is expressed if both GM-CSF and TNF-α are expressed. The importance of TNF-α as the second important criterion for IL-17 production is only evident at the day 21 after immunization (**Figures 9** and **11C**), not at the day 9 (**Figures 11A,B**). Further comparisons of the induced trees demonstrate that the expression of TNF-α is the most

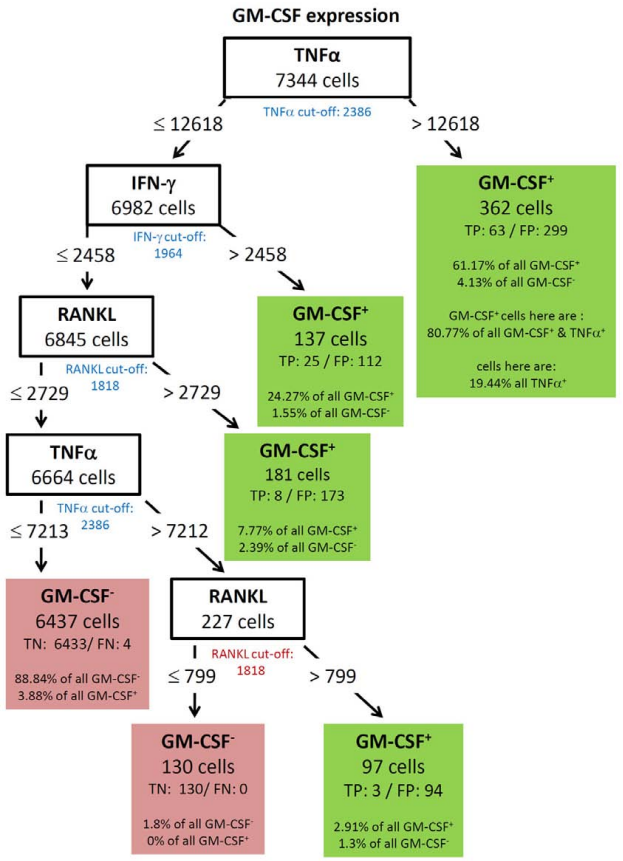


FIGURE 8 | Best decision tree for the classification of cells as positive or negative for GM-CSF expression. Cells are first classified based on the MFI values of TNF-α. The blue colored TNF-α cut-off value indicates that the split value is high above the cut-off value. Therefore, only a proportion of the cells which express TNF-α are routed to the right leaf. A closer look at this leafs shows that due to the split value high above the cut-off value only 19.44% of all TNF-α positive cells are contained in this leaf. But this leaf captures 61.17% of the GM-CSF positive cells. The leaf also contains the information that this leave captures 80.77% of all GM-CSF and TNF-α positive cells. Form the root (top) cells with a TNF-α MFI value equal or below 12618 are routed down to the next inner node. This node splits the cells on the MFI value of IFN-γ. Again, the blue colored cut-off values indicates that not all IFN-γ positive cells in this node are routed to the leaf at the right. This leaf also classifies cells as GM-CSF positive. This leaf captures less of the GM-CSF positive cells – around one out of four. The other two leafs which give a positive classification captures even less of the GM-CSF positive cells, but summed up the four leaves which give a positive classification captures 96.12% of all GM-CSF positive cells.

important split criterion also for the GM-CSF as well as IL-2 production not only for the standard condition (**Figures 4** and **8** for IL-2 and GM-CSF, respectively) but also for the knock-out mutant when observed at 21 days after immunization (KO-day 21, **Figures 12A,B** for IL-2 and GM-CSF, respectively). These rules have not been confirmed when measured 9 days after immunization.

As stated in the section 3 we use different parameter settings and chose the best tree according to specific quality criteria. The

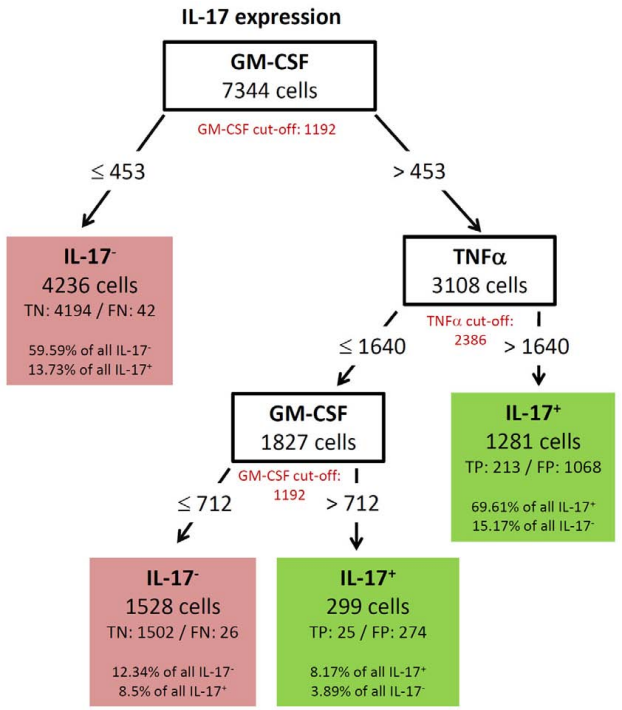


FIGURE 9 | Best decision tree for the classification of cells as positive or negative for IL-17 expression. Cells are first classified based on the MFI values of GM-CSF. The red colored GM-CSF cut-off value indicates that the split value is high below the cut-off value. The cells routed to the left node do not express GM-CSF and this value split captures 59.59% of all IL-17 negative cells and only wrongly classifies 13.73% of the IL-17 positive cells. The next split on the right is on TNF-α, but again the split value is below the cut-off value. The split values of GM-CSF and TNF-α are below the cut-off value and thus not all cells, which are routed down the path to the rightmost leaf express this two cytokines. Nevertheless this leaf correctly classifies 69.61% of the IL-17 positive cells as positive. Only 15.17% of the IL-17 negative cells are wrongly classified in this leaf. Also the last split value again on GM-CSF is below the cut-off value.

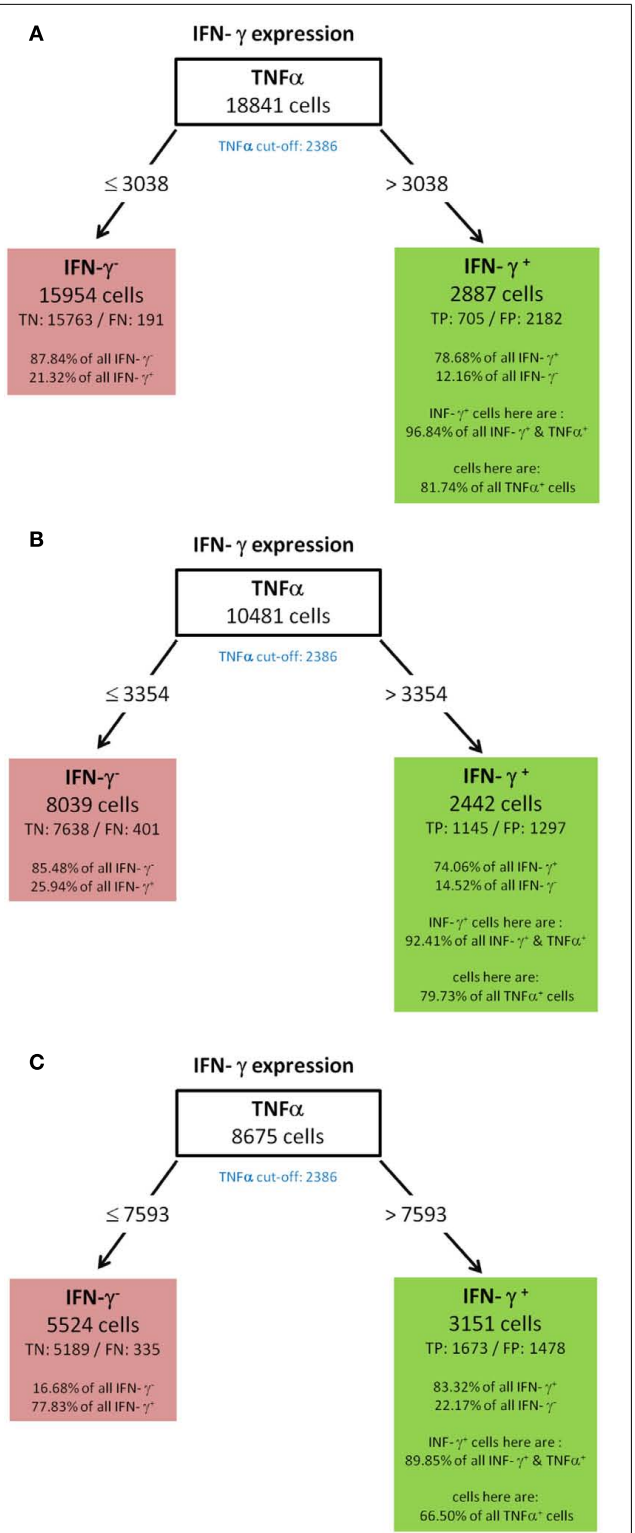


FIGURE 10 | IFN-γ (A) WT-day 9, (B) KO-day 9, (C) KO-day 21.

question is whether our quality criteria always choose a tree created with the same parameter settings. The answer is no. The parameter settings for the induction of the decision trees clearly vary for each cytokine to meet our quality criteria. For instance, the induction of the "best" decision trees for GM-CSF, IFN-γ, and IL-17 required the inclusion of all cells (cytokine-positive and negative) while for IL-2, RANKL, and TNF-α, a better classification could be reached when cytokine-negative cells were omitted. This is due to the fact that IL-2, RANKL, and TNF-α have a high percentage of cells which only produce this cytokine (see **Figure 3**). The corresponding decision trees have the characteristic that positive cells are routed to the left most node (see **Figures 4** and **5**), thus these cells are correctly classified as positive if they do not produce the cytokines used for this trees. If also the cytokine-negative cells have been used to induce decision trees for IL-2, RANKL, and TNF-α, then also all cytokine-negative cells would have been routed to this leaf and that would have worsen the classification. Consequently, although the induction of decision trees seems to be a promising approach for the analysis of multidimensional data, standard parameter settings that are suitable for all data sets cannot be proposed. Our approach to choose a decision tree will be discussed in the next section.

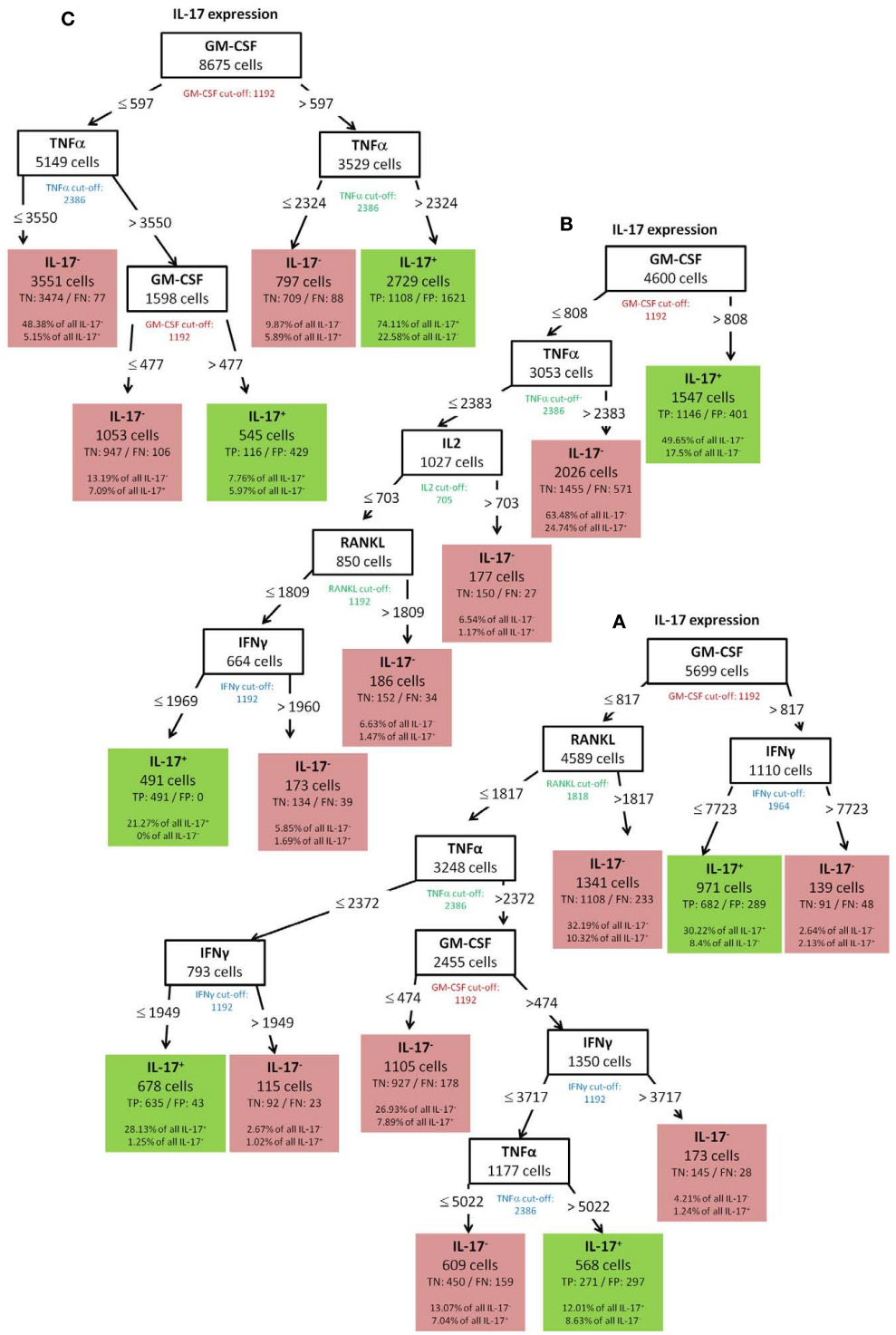


FIGURE 11 | IL-17 (A) WT-day 9, (B) KO-day 9, (C) KO-day 21.

4. DISCUSSION

Commonly used sequential filtering strategies have several disadvantages in the analysis of multi-parametric flow cytometry data. They are time-consuming and subjective and therefore information which is contained in the dataset might be lost. Perhaps more importantly, this approach is mainly descriptive and not quantitative. The establishment of an improved workflow for the analysis of flow cytometric data is a demanding need since

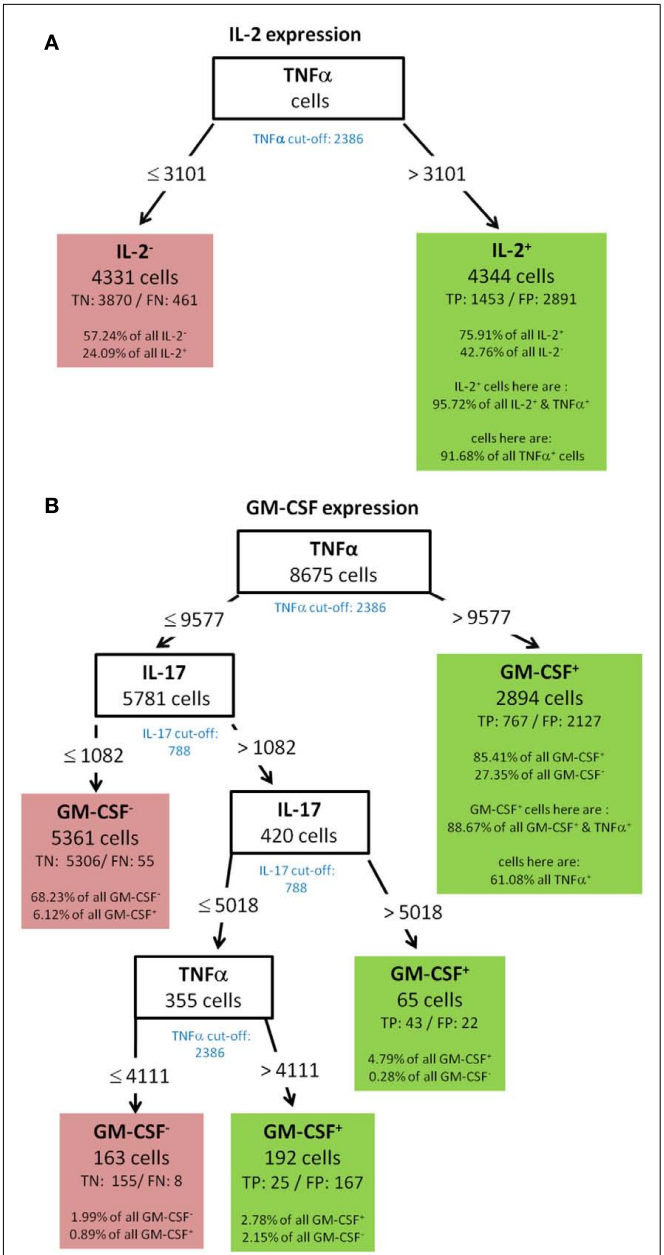


FIGURE 12 | (A) IL-2 KO-day 21, **(B)** GM-CSF KO-day 21.

further increases in data complexity can be expected by the current technical advance in flow cytometry technology (Perfetto et al., 2004; Lugli et al., 2010; Bendall et al., 2011; Mittag and Tarnok, 2011). We tested the feasibility of the induction of decision trees to identify structural patterns in multidimensional flow cytometric data. Data sets from intracellular cytokine staining of antigen-specific T helper cells were analyzed in our proof-of-concept study. T helper cell cytokine production is critical for their capacity to regulate different aspects of the immune response. Cytokine secretion must occur in a coordinated way for maximum efficiency of an immune response. In several independent studies, it has been demonstrated that the protective or pathogenic potential of a T cell response is not determined by the production of a single cytokine. It is rather correlated to their capacity for a coordinated expression of cytokines. For instance, protection against infection with the intracellular parasite *Leishmania major* is related to a high number of Th cells producing TNF-α, IFN-γ, and IL-2 simultaneously (Darrah et al., 2007; Seder et al., 2008). The identification of patterns of cytokine expression by machine learning might be a useful tool for a better understanding of both T cell immunology and system biology of microbial infection, which critically depends on bidirectional interactions between the pathogen and the host. This prompted us to test the feasibility of the induction of decision trees for the analysis of highly complex flow cytometric data. We hypothesized that good retrieval of information requires good classification by the decision tree. Unfortunately, universally applicable criteria that assess the quality of a decision tree do not exist; these criteria depend on both the data and on the intention of the researcher. Furthermore, there are no general applicable parameter settings in machine learning. Thus, we used different parameter settings and ended up with a large set of decision trees from which we had to choose. Our approach was to select the one tree we considered to be the "best" tree. For the selection of the best, we chose a combination of different criteria, including the geometric mean of the TP rate and the TN rate, the area under the ROC curve (AUC) and geometric mean of the F-measures (see section 3). Since decision trees works best with balanced data, we also weighted our highly imbalanced data sets regarding positive and negative cells (see **Figure 3A**). Although, we were able to identify a tree of sufficient for each cytokine, the parameter settings for the induction of the decision trees clearly varied to meet our quality criteria. For instance, the induction of the "best" decision trees for GM-CSF, IFN-γ, and IL-17 required the inclusion of all cells (cytokine-positive and negative). In contrast for IL-2, RANKL, and TNF-α, a better classification could be reached when cytokine-negative cells were omitted. This shows that although the induction of decision trees seems to be a promising approach for the analysis of multidimensional data, standard parameter settings, which are suitable for all data sets, cannot be proposed. The choice and combination of the quality measure was adjusted for our task. We consequently used primarily the geometric mean of the TP and TN rate since we considered both as equally important and chose the smallest tree, since we aimed at a visual expert inspection of the trees. Other tasks may require other criteria. An extension to our approach to chose only the "best" tree could be to provide the expert user with a set of good trees. A comparison of these set could reveal more insight and will be the scope of future work. Another discussable point is the weighting of the data sets according to their imbalance. Due to the weighting of the data, a node is also classified as positive if the number of cells negative for this cytokine is much higher than the number of cells positive for this cytokine. Thus, the precision of classification becomes lower. Such low precision heavily impairs the ability to predict the expression of a given marker depending on the known expression of other markers, in order to avoid its direct measurement. Since we used decision trees to find structural patterns in the data and to generate hypothesis from these patterns, precision does not play a critical role.

The work was focused on standard condition (WT-day 21). However, for comparative reasons and to check the robustness of identified rules, additional conditions were investigated using knock-out mice (KO) and measurement already 9 days after immunization. The importance of high TNF-α expression for IFN-γ was confirmed for all conditions while for GM-CSF and IL-2 production only when measured 21 days after immunization. Furthermore, TNF-α was identified as the second important criterion also for IL-17 production but again only when measured 21 days after immunization. Summarizing, some rules of co-expression have been confirmed with different experimental conditions. Thus, the method of induction of decision trees is able to extract robust rules. Interestingly, the strong dependency of the expression of one cytokine on the expression of others which we found reproducibly between different time points and despite differing genotypes implies some biological significance of these findings. We have previously shown in a kinetic study that TNF-α is one of the earliest cytokines produced after activation of antigen-specific T cells (Frey et al., 2010b). Given that expression of the other cytokines starts later, the strong relationship between TNF-α and the expression of the other cytokines could argue for a hard-wired connection between the expression of these mediators. High TNF-α expression has been described as a marker of polyfunctional T cells in another study (Darrah et al., 2007; Seder et al., 2008), supporting our hypothesis that TNF-α expression is highly correlated with the expression of other effector cytokines.

In conclusion, the presented results show that data analysis with decision trees can easily reveal structural patterns in flow cytometric data that would have been missed by conventional analysis. Such patterns can be used for the generation of hypothesis on the complex biology of certain subsets of cells.

ACKNOWLEDGMENTS

Funding: Svenja Simon and Reinhard Guthke were supported by the Deutsche Forschungsgemeinschaft (DFG) Priority Program 1160 "Colonisation and infection by human-pathogenic fungi." Thomas Kamradt and Oliver Frey were supported by the Deutsche Forschungsgemeinschaft (DFG) Priority Program 1468 "Immunobone" (KA 755-6/1).

REFERENCES

Bendall, S. C., Simonds, E. F., Qiu, P., Amir, E.-A. D., Krutzik, P. O., Finck, R., Bruggner, R. V., Melamed, R., Trejo, A., Ornatsky, O. I., Balderas, R. S., Plevritis, S. K., Sachs, K., Pe'er, D., Tanner, S. D., and Nolan, G. P. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* 332, 687–696.

Bruns, L., Frey, O., Morawietz, L., Landgraf, C., Volkmer, R., and Kamradt, T. (2009). Immunization with an immunodominant self-peptide derived from glucose-6-phosphate isomerase induces arthritis in DBA/1 mice. *Arthritis Res. Ther.* 11, R117.

Darrah, P. A., Patel, D. T., De Luca, P. M., Lindsay, R. W. B., Davey, D. F., Flynn, B. J., Hoff, S. T., Andersen, P., Reed, S. G., Morris, S. L., Roederer, M., and Seder, R. A. (2007). Multifunctional T(H)1 cells define a correlate of vaccine-mediated protection against *Leishmania major*. *Nat. Med.* 13, 843–850.

Data File Standards Committee of the Society for Analytical Cytology. (1990). Data file standard for flow cytometry. *Cytometry* 11, 323–332.

Frey, O., Bruns, L., Morawietz, L., Dunussi-Joannopoulos, K., and Kamradt, T. (2011a). B cell depletion reduces the number of autoreactive T helper cells and prevents glucose-6-phosphate isomerase-induced arthritis. *PLoS ONE* 6, e24718. doi:10.1371/journal.pone.0024718

Frey, O., Tania, M., Kelchtermans, H., Schurgers, E., Kamradt, T., and Matthys, P. (2011b). Ameliorated course of glucose-6-phosphate isomerase (G6PI)-induced arthritis in IFN-γ receptor knockout mice exposes an arthritis-promoting role of IFN-γ. *J. Autoimmun.* 36, 161–169.

Frey, O., Meisel, J., Hutloff, A., Bonhagen, K., Bruns, L., Kroczek, R. A., Morawietz, L., and Kamradt, T. (2010a). Inducible costimulator (ICOS) blockade inhibits accumulation of polyfunctional T helper 1/T helper 17 cells and mitigates autoimmune arthritis. *Ann. Rheum. Dis.* 69, 1495–1501.

Frey, O., Reichel, A., Bonhagen, K., Morawietz, L., Rauchhaus, U., and Kamradt, T. (2010b). Regulatory T cells control the transition from acute into chronic inflammation in glucose-6-phosphate isomerase-induced arthritis. *Ann. Rheum. Dis.* 69, 1511–1518.

Gentleman, R., Carey, V., Bates, D., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J., and Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5, 1–16.

Hahne, F., LeMeur, N., Brinkman, R. R., Ellis, B., Haaland, P., Sarkar, D., Spidlen, J., Strain, E., and Gentleman, R. (2009). flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinformatics* 10, 106.

Hornik, K., Buchta, C., and Zeileis, A. (2009). Machine learning: R meets Weka. *Comput. Stat.* 24, 225–232.

Hulspas, R., O'Gorman, M. R. G., Wood, B. L., Gratama, J. W., and Sutherland, D. R. (2009). Considerations for the control of background fluorescence in clinical flow cytometry. *Cytometry B Clin. Cytom.* 76B, 355–364.

Kirchhoff, D., Frentsch, M., Leclerk, P., Bumann, D., Rausch, S., Hartmann, S., Thiel, A., and Scheffold, A. (2007). Identification and isolation of murine antigen-reactive T cells according to CD154 expression. *Eur. J. Immunol.* 37, 2370–2377.

Kubat, M., Holte, R., and Matwin, S. (1998). Machine learning for the detection of oil spills in satellite radar images. *Mach. Learn.* 30, 195–215.

Lugli, E., Roederer, M., and Cossarizza, A. (2010). Data analysis in flow cytometry: the future just started. *Cytometry A* 77A, 705–713.

Mittag, A., and Tarnok, A. (2011). "Recent advances in cytometry applications: preclinical, clinical, and cell biology," in *Recent Advances in Cytometry, Part B: Advances in Applications, Vol. 103 of Methods in Cell Biology*, 5th Edn, eds Z. Darzynkiewicz, E. Holden, A. Orfao, W. Telford, and D. Wlodkowic, (Waltham: Elsevier Academic Press Inc.), 3–20.

Parks, D., Roederer, M., and Moore, W. (2006). A new "logicle" display method avoids deceptive effects of logarithmic scaling for low signals and compensated data. *Cytometry A* 69A, 541–551.

Perfetto, S., Chattopadhyay, P., and Roederer, M. (2004). Innovation – seventeen-colour flow cytometry: unravelling the immune system. *Nat. Rev. Immunol.* 4, 648–655.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. San Francisco: Morgan Kaufmann Publishers.

R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Roederer, M., and Moody, M. A. (2008). PolyChromatic plots: graphical display of multidimensional data. *Cytometry A* 73A, 868–874.

Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D., and Nolan, G. (2005). Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 308, 523–529.

Schubert, D., Maier, B., Morawietz, L., Krenn, V., and Kamradt, T. (2004). Immunization with glucose-6-phosphate isomerase induces T cell-dependent peripheral polyarthritis in genetically unaltered mice. *J. Immunol.* 172, 4503–4509.

Seder, R. A., Darrah, P. A., and Roederer, M. (2008). T-cell quality in memory and protection: implications for vaccine design. *Nat. Rev. Immunol.* 8, 247–258.

Sun, Y., Kamel, M., Wong, A., and Wang, Y. (2007). Cost-sensitive boosting for classification of imbalanced data. *Pattern Recognit.* 40, 3358–3378.

Weiss, G., and Provost, F. (2003). Learning when training data are costly: the effect of class distribution on tree induction. *J. Artif. Intell. Res.* 19, 315–354.

Witten, I. H., and Frank, E. (2005). *Data Mining – Practical Machine Learning Tools and Techniques*, 2nd Edn. San Francisco: Morgan Kaufmann Publishers.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 December 2011; accepted: 12 March 2012; published online: 02 April 2012.

*Citation: Simon S, Guthke R, Kamradt T and Frey O (2012) Multivariate analysis of flow cytometric data using decision trees. Front. Microbio. **3**:114. doi: 10.3389/fmicb.2012.00114*
This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.

frontiers in
MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 14 September 2012
doi: 10.3389/fmicb.2012.00284

Comparison of sepsis-induced transcriptomic changes in a murine model to clinical blood samples identifies common response patterns

Sandro Lambeck[1]*, Martina Weber[2], Falk A. Gonnert[1,2], Ralf Mrowka[3] and Michael Bauer[1,2]

[1] Integrated Research and Treatment Center – Center for Sepsis Control and Care, Jena University Hospital, Jena, Germany
[2] Department of Anaesthesiology and Intensive Care Therapy, Jena University Hospital, Jena, Germany
[3] Department of Internal Medicine III – Experimental Nephrology, Jena University Hospital, Jena, Germany

Edited by:
Marc Thilo Figge, Hans-Knoell-Institute, Germany

Reviewed by:
Oliver Kurzai, Friedrich-Schiller-University Jena, Germany
Hortense Slevogt, University Hospital Jena, Germany
Daniela Albrecht-Eckardt, BioControl Jena GmbH, Germany

***Correspondence:**
Sandro Lambeck, Center for Sepsis Control and Care, Jena University Hospital, Erlanger Allee 101, 07747 Jena, Germany.
e-mail: sandro.lambeck@med.uni-jena.de

Experimental models, mimicking physiology, and molecular dynamics of diseases in human, harbor the possibility to study the effect of interventions and transfer results from bench to bedside. Recent advances in high-throughput technologies, standardized protocols, and integration of knowledge from databases yielded rising consistency and usability of results for inter-species comparisons. Here, we explored similarities and dissimilarities in gene expression from blood samples of a murine sepsis model (peritoneal contamination and infection, PCI) and patients from the pediatric intensive care unit (PICU) measured by microarrays. Applying a consistent pre-processing and analysis workflow, differentially expressed genes (DEG) from PCI and PICU data significantly overlapped. A major fraction of DEG was commonly expressed and mapped to adaptive and innate immune response related pathways, whereas the minor fraction, including the chemokine (C–C motif) ligand 4, exhibited constant inter-species disparities. Reproducibility of transcriptomic observations was validated experimentally in PCI. These data underline, that inter-species comparison can obtain commonly expressed transcriptomic features despite missing homologs and different protocols. Our findings point toward a high suitability of an animal sepsis model and further experimental efforts in order to transfer results from animal experiments to the bedside.

Keywords: innate and adaptive immune response, major histocompatibility complex class II, sepsis, systems biology, transcriptomics

INTRODUCTION

Application of high-throughput technologies like microarrays extended the understanding of the complex molecular interactions in infectious disease processes (Jenner and Young, 2005; Polpitiya et al., 2009). Contrary to a single gene study, unbiased quantitative insights are collected by simultaneously monitored gene expression levels. These can be subject to data- and knowledge-driven analyses by bioinformatic and systems biology approaches in order to extract hypotheses about molecular interactions and predict interventional targets. Validation of these hypotheses may lead to translational approaches and ultimately optimize clinical practice (Cavaillon and Adib-Conquy, 2006). A major precondition for this comprises the inter-species comparison of molecular features in response to disease including their interactions as well as the underlying dynamics, which are covered by translational systems biology (Vodovotz et al., 2008). Experiments in animal models like *M. musculus* are often used to study human gene expression based on the assumption that the expression and function are similar for the majority of orthologous genes between both species (Lu et al., 2009). Although >12,000 orthologous gene pairs, i.e., homologous genes in the same syntenic location, were found in human and mouse, genomic differences with transcriptional and regulatory relevance exist (Mestas and Hughes, 2004). Comparison of expression differences in meta-studies, focusing on comparing lists of differentially expressed genes (DEG) derived from literature, e.g., studies about dietary restriction spanning six organisms for different species (Han and Hickey, 2005) or of sepsis markers in humans (Tang et al., 2010), yielded no agreement or only small intersections, respectively. This may be related to biological variability as well as different protocols for experimental set-up and data analysis. In contrast, promising transcriptomic comparisons from murine and human samples with respect to conserved pathways (Lim et al., 2010) as well as similar expression of homologs (Dowell, 2011) have been reported. Using a consistent workflow and filtering according to microarray quality control (MAQC, Shi et al., 2006) criteria, DEG with approximately 50% overlap in murine and human monocytes were obtained (Ingersoll et al., 2010), which is driving attempts to find conserved markers for diseases like sepsis across species.

Sepsis is a syndrome characterized by a systemic inflammatory response to infection (Bone et al., 1992) with high incidence in the intensive care unit (ICU) and especially critical for the Pediatric Intensive Care Unit (PICU) patients (Kovarik and Siegrist, 1998). The most severe form of sepsis, septic shock, is characterized by hypotension or hypoperfusion despite adequate fluid resuscitation. Therapies aimed at reducing the inflammatory burden of

acute sepsis patients have been studied in clinical trials with little success, underlining the need to revise current approaches including the basic research in animals (Marshall, 2008). Experimental approaches in mimicking the sepsis syndrome include the use of peritonitis models like the polymicrobial challenges by peritoneal contamination and infection (PCI; Gonnert et al., 2011) and cecal ligation and puncture (CLP; Hubbard et al., 2005) or pneumonia models (Weber et al., 2012) aiming to closely reflect the pathophysiology observed in humans in the rodent models (Rittirsch et al., 2008).

In the present study, we investigated inter-species comparability based on gene expression patterns from experimental and clinical samples from septic blood. We compared time-resolved transcriptomic data obtained from (I) a murine polymicrobial sepsis model induced by PCI (Gonnert et al., 2011) and (II) sepsis and septic shock patients from the PICU and identified genes and pathways which could be translationally (across species) promising for diagnostics and interventions.

MATERIALS AND METHODS

DATA CHARACTERISTICS

Table 1 is summarizing the animal (I, PCI) and human (II, PICU) sepsis data used in this comparison further described in detail below.

(I) Animal experiments were carried out after approval of the local ethics committee. Using a recently published rodent sepsis model (Gonnert et al., 2011), 8-week-old female C57/BL6 mice were subject to the PCI challenge. Eight mice, equally divided into 6 and 24 h observational groups, received an intraperitoneal injection of 200 µl of a human feces suspension (0.21×10^6 CFU/g, predicted to be lethal). Control mice ($n = 4$) were used for adjustments. Blood was drawn with a 20G Sterican needle (Braun, Melsungen, Germany) from the *Vena cava caudalis* in a syringe prepared with 200 µl PAXgene reagent (PreAnalytix®, Qiagen, Hilden, Germany). Clearing from globin RNA from whole blood samples was done by the *GLOBINclearTM Mouse/Rat Kit* by Ambion® (Applied Biosystems, Darmstadt, Germany). RNA integrity and concentration analyses were performed with the Bioanalyzer 2100 (Agilent Technologies®, Palo Alto, CA, USA) according to the manufacturers specifications (*RNA 6000 Nano Assay*) using *RNA 6000 Nano LabChip® Kits* by Agilent Technologies®. Amplification and biotinylation of RNA was in agreement to the protocol of the *Illumina® TotalPrep RNA Amplifikation Kit* from Ambion (Applied Biosystems, Darmstadt, Germany). Hybridization was in accordance to the *Whole-Genome Gene Expression with IntelliHybTM Seal* protocol (Revision B) from Illumina® (San Diego, CA, USA). Scanning of arrays was done with the BeadArray Reader 500 X for Illumina MouseWG-6 v2.0 expression beadchips. Validation of expression changes for candidate features PCI (6 h) was performed using a lower dose (0.14×10^6 CFU/g, predicted to be lethal) for $n = 2$ gender- and age-matched mice monitored on Illumina MouseRef-8 V2 chips and compared to a control in a follow-up run.

(II) A publically available data set containing gene expression values from blood samples of PICU sepsis and septic shock patients and controls, along protocols as described elsewhere (Wong et al., 2009). Briefly, pediatric patients with septic shock were characterized by a significantly higher illness severity, higher mortality ($n = 27$ died within 28 days), and an increased degree of organ failure, compared to sepsis patients. Total RNA was isolated from whole blood samples using the PaxGene Blood RNA System according to the manufacturer's specifications. Raw data for gene expression profiling was obtained from *Gene Expression Omnibus* (accession ID GSE13904).

DATA ANALYSIS

Computations were performed using R software (http://www.r-project.org/) version 2.15 and Bioconductor (Gentleman et al., 2004) packages. To assure data comparability (Ramasamy et al., 2008) a consistent workflow was applied.

(I) PCI data was subjected to quality control and the robust spline procedure from the R package *lumi* (Du et al., 2008), combining the features of quantile and loess normalization and $\log_2$ transformed signals were obtained. Bead types having detection values $p < 0.01$ in at least four of the RNA samples were taken. Subsequently, bead types were mapped to most recent annotation and gene-centered agglomeration by Entrez IDs was performed. Normalized detected bead types from PCI samples and controls are provided in Supplementary Tables 1 and 2, for the initial run and the validation data set, respectively.

(II) PICU data was pre-processed using chip definition file from *Brainarray* (v. 15.1, 2012) which aggregates probes into updated gene-centered probe set definitions mapped to Entrez IDs (Sandberg and Larsson, 2007). Further preprocessing was performed using quantile and loess normalization.

Table 1 | Characteristics of transcriptomic data used in this analysis.

Organisms (age)	PCI (experimental sepsis)	PICU (clinical sepsis and septic shock)	
	Female C57BL/6 mice (8-week old)	Pediatric intensive care unit (PICU) patients (2–5-year old)	
Syndrome	Sepsis	Sepsis	Septic shock
Case number [time (h)]	4, 4 (6, 24)	32, 20 (24, 72)	67, 39 (24, 72)
Controls	4	18	
Chip	Illumina Mouse WG 6 v2	Affymetrix HGU 133plus2	

Entrez IDs were mapped using the Array Information Library Universal Navigator (AILUN)-service (Chen et al., 2007), facilitating cross-species and cross-platform comparisons by the use of homologous gene definitions. Merged data sets were re-normalized to avoid potential bias due to different log2 signal ranges. DEG were separately identified in PCI and PICU data according to micro array quality control (MAQC; Shi et al., 2006) – criteria and standard thresholds by: (1) average twofold difference for sample groups (opposed to controls) and (2) FDR (Benjamini and Hochberg, 1995)-adjusted p-values from moderated t-statistics <0.05 using *limma* (Smyth, 2004). Because different hybridization properties of the probes on species–specific arrays can profoundly bias cross-platform comparison of gene expression levels the use of relative expression changes is preferred to absolute expression signals (Liao and Zhang, 2006). Parametric tests (t-statistics) were applied to account for the small sample-sized PCI data. Pearson correlation (r) was used to assess similarity.

Hypergeometric tests were performed for the evaluation if intersections of separately identified DEG from PICU and PCI data could occur by chance (Fury et al., 2006). In addition, Monte–Carlo simulations (MCS, 4,000 iterations) were employed to non-parametrically assess the significance of intersections for two and more groups by re-sampling procedures (Ramasamy et al., 2008). One-sided p-values were defined by the fraction of counts derived from the empirical distributions of randomly expected intersections being larger than the number of observed intersections from the original gene lists.

To include potential differences in PICU (sepsis) and PICU (septic shock), commonly expressed DEG, i.e., found in human and murine data at any time point, were obtained by: DEGC = {DEG(PICU sepsis) ∪ DEG(PICU septic shock)} ∩ DEG(PCI). Differences in signs of expression changes were assessed by sign (DEGC(PICU)) ≠ sign (DEGC(PCI)). Hierarchical clustering (Eisen et al., 1998) of average log2 fold change data was applied for features drawn in a heatmap. Gene set enrichment analysis (GSEA) was performed using the *Database for Annotation, Visualization, and Integrated Discovery* (DAVID, v6.7; Huang et al., 2009) and its internal variant of Fisher's exact test for enrichment within the pathways defined by the *Kyoto Encyclopedia of Genes and Genomes* (KEGG; Kanehisa et al., 2008) for murine and human data and backgrounds.

RESULTS

DIFFERENTIALLY EXPRESSED GENES SIGNIFICANTLY OVERLAP IN BOTH SPECIES

Pre-processing yielded $k = 18{,}988$ and $l = 9{,}325$ gene-centered features from human and murine samples, respectively. Merging by Entrez IDs narrowed down the lists to $m = 7{,}461$ homologous features. Converted log2FC values (opposed to controls) of common detected features (m) indicated overall positive and significant correlations ($p < 0.05$) for PCI and PICU, but higher correlation within PICU data ($r > 0.85$). Combining DEG from PCI and PICU data as shown in the Venn diagrams for the early time point (**Figure 1A**) and both time points (**Figure 1B**) indicated time-persistent expression changes. Observed intersections of DEG from PCI and PICU were significantly higher than would be expected by chance (**Figure 1A**; e.g., $n_1 = 89$ common DEG in PCI and PICU, $p < 0.05$ hypergeometric test; MCS sampling out of m features). Notably, septic shock induced more DEG than sepsis in PICU data (196 DEG were exclusively assigned to septic shock for both time points). DEGC, i.e., DEG found in PCI and PICU at any time point (**Figure 1B** highlighted in gray, $n_{12} = 135$, $p < 0.05$), as depicted in the heatmap (**Figure 1C**), indicated hierarchically grouped clusters for down- and upregulated genes. A small subset of DEGC (10 out of 135) was expressed with different signs, comprising, e.g., the chemokine (C-X-C motif) ligand 4 (*CCL4*) and *MARCKSL1* genes with high average disparity.

Assessing the dynamic changes for each group by the MAQC-criteria, e.g., PICU sepsis 72 h (late) *versus* PICU sepsis 24 h (early), yielded no qualitative changes in PICU sepsis data and only one gene in PICU septic shock data, but 141 features for PCI data (PCI 24 h *versus* PCI 6 h).

GENE SET ENRICHMENT ANALYSIS IDENTIFIES PATHWAYS RELATED TO THE INNATE AND ADAPTIVE IMMUNE RESPONSE

Next, we thought to determine the pathways exhibiting highest enrichment for the DEGC. Equally signed DEGC mapped to pathways from the immune response using DAVID knowledge base on KEGG (*M. musculus* definitions) as shown in **Table 2**.

Top enrichment of commonly downregulated genes from the DEGC was found in the *CAMs*-category, containing *CD8B1*, *ICAM2*, and histocompatibility 2 class II locus genes (*H2-OB*, *H2-EB1*, and *H2-DMa*). Histocompatibility 2 class II locus genes were also included in the *Antigen processing and presentation* pathway and indicated an overlap for further categories, e.g., *Asthma*.

Top enriched category termed *Cytokine–Cytokine-Receptor Interaction* for commonly upregulated genes comprised chemokine (C-X-C motif) ligand 16, and related genes from interleukin receptors 1, 10, 17, and 18. Further enrichment was found for the *Insulin signaling pathway* (e.g., flotilin 1 and 2) and the Toll-like receptor (TLR) signaling pathway (*CD14*, *TLR2*, *TLR4*, and *TLR6*). Mapping of DEGC by human definitions/background to KEGG did not alter the ranking of enriched pathways.

VALIDATION OF THE TRANSCRIPTOMIC RESPONSE IN MURINE SEPSIS

Validation by a second run of wet lab experiments for the murine PCI challenge to a slightly lower dose yielded similar expression changes, i.e., significant positive correlation ($r > 0.7$, $p < 0.05$) by log2FC values of all detected genes on both Illumina platforms. Higher correlation ($r > 0.9$) was obtained for detectable DEGC ($n = 115$) in comparison of log2FC values of PCI data (first versus second run) as depicted in **Figure 2A** (including two highlighted outliers). Detailed log2 signals are exemplarily shown in **Figure 2B** for a selection of the aforementioned genes.

DISCUSSION

In the present work, we compared the transcriptomic response in blood samples from a murine sepsis model and pediatric septic patients. Batch-effects, due to different protocols, in this inter-species and cross-platform comparison, were approached by standard bioinformatic and microarray data analysis procedures including probe set re-annotation, group-wise control adjustments and using stringent MAQC-criteria for filtering of

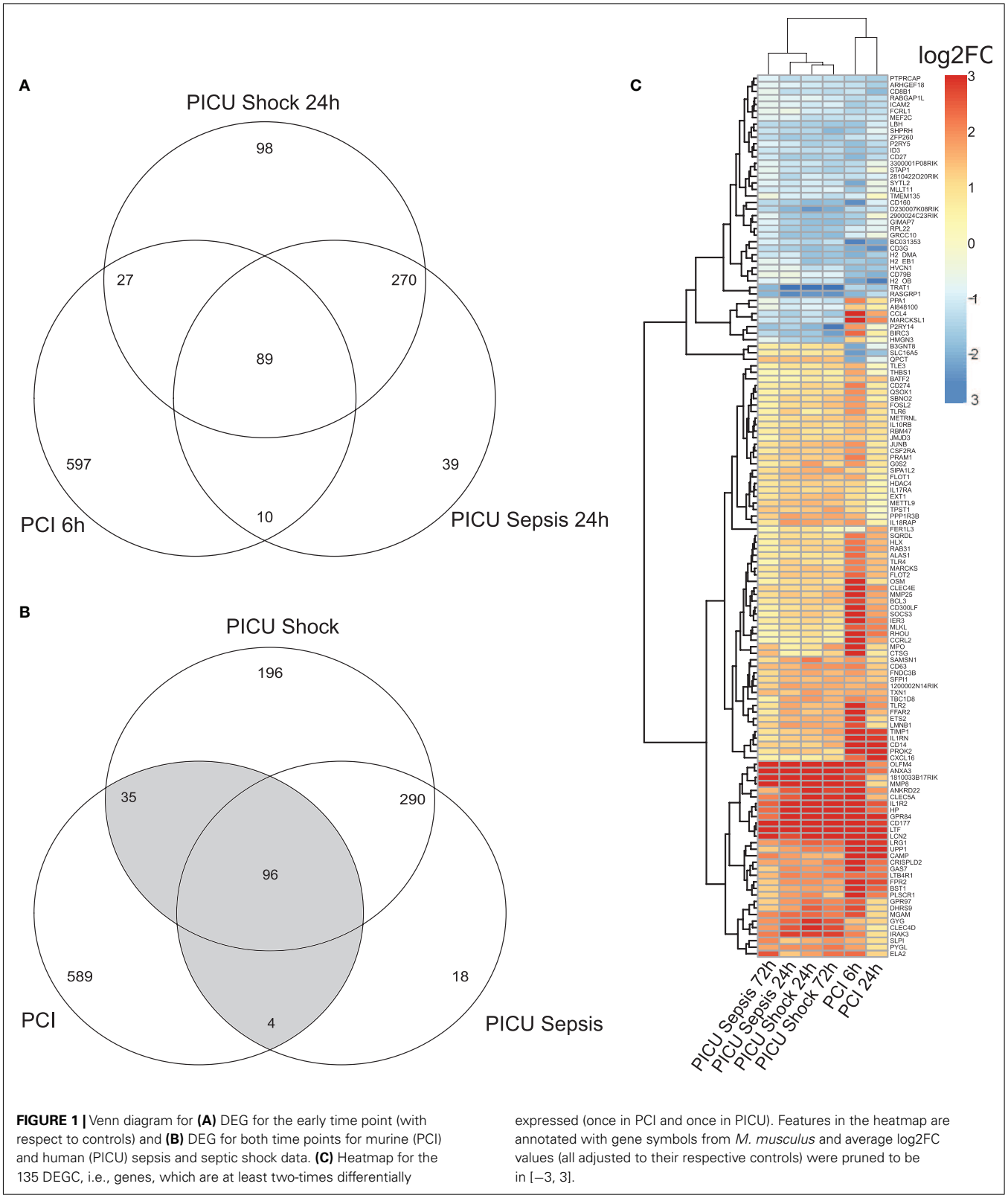


FIGURE 1 | Venn diagram for **(A)** DEG for the early time point (with respect to controls) and **(B)** DEG for both time points for murine (PCI) and human (PICU) sepsis and septic shock data. **(C)** Heatmap for the 135 DEGC, i.e., genes, which are at least two-times differentially expressed (once in PCI and once in PICU). Features in the heatmap are annotated with gene symbols from *M. musculus* and average log2FC values (all adjusted to their respective controls) were pruned to be in [−3, 3].

DEG. The analyzed spectrum of time-resolved gene expression data demonstrated positive correlations for PCI and PICU data with significantly overlapping DEG. Furthermore, commonly expressed genes were found throughout the observation period for both species, whereas the early time points (6 h PCI, 24 h PICU) already comprised the majority of identified DEG.

Table 2 | Top five enriched KEGG pathways obtained from DAVID (*p*-values Fisher's exact test $p < 0.05$) supported by FDR-adjusted *p*-values within the common down- and upregulated DEG from PCI and PICU.

	Pathway	Hits	%	*p*-Value	FDR-adjusted *p*-value
Commonly downregulated genes (32 DAVID IDs)	Cell adhesion molecules (CAMs)	5	15.6	3.0E−04	7.7E−03
	Antigen processing and presentation	4	12.5	9.8E−04	1.3E−03
	Asthma	3	9.4	2.4E−03	2.1E−02
	Intestinal immune network for IgA production	3	9.4	6.3E−03	4.1E−02
	Allograft rejection	3	9.7	7.3E−03	3.7E−02
Commonly upregulated genes (87 DAVID IDs)	Cytokine–cytokine receptor interaction	7	8	2.0E−03	7.2E−02
	Insulin signaling pathway	5	5.7	6.8E−03	1.2E−01
	Toll-like receptor signaling pathway	4	4.6	1.7E−02	2.0E−01
	Jak-STAT signaling pathway	4	4.6	5.2E−02	4.0E−01
	Hematopoietic cell lineage	3	3.4	7.90E−02	4.7E−01

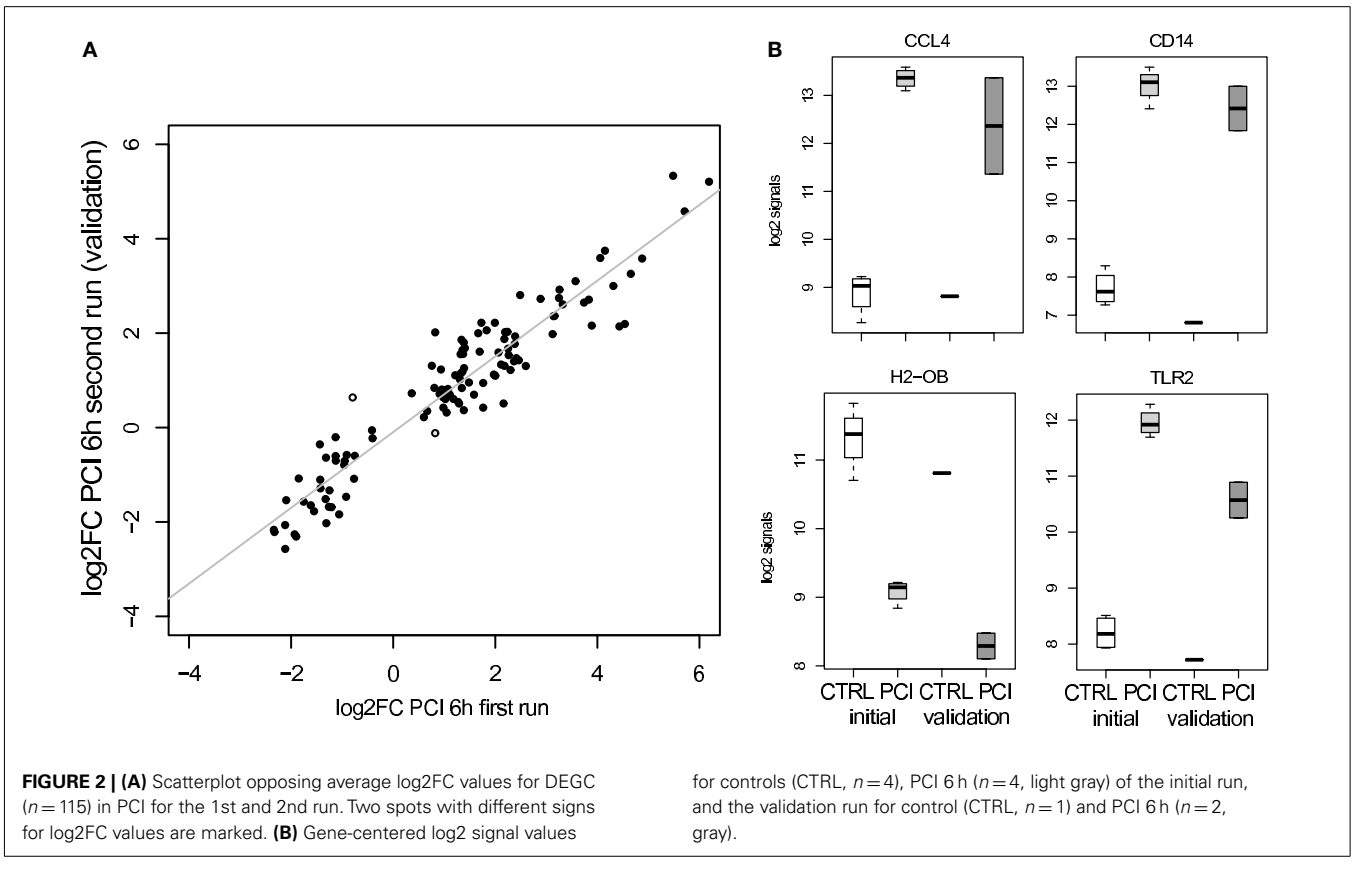


FIGURE 2 | (A) Scatterplot opposing average log2FC values for DEGC ($n = 115$) in PCI for the 1st and 2nd run. Two spots with different signs for log2FC values are marked. **(B)** Gene-centered log2 signal values for controls (CTRL, $n = 4$), PCI 6 h ($n = 4$, light gray) of the initial run, and the validation run for control (CTRL, $n = 1$) and PCI 6 h ($n = 2$, gray).

LITERATURE SUMMARY FOR GENES AND PATHWAYS RELATED TO THE INNATE AND ADAPTIVE IMMUNE RESPONSE IN SEPSIS

Although not statistically significant after multiple test correction commonly upregulated features mapped to *Cytokine–Cytokine-Receptor Interaction*, *Insulin signaling pathway*, and the *TLR signaling pathway*. The latter comprised *TLR2*, *TLR4*, and *TLR6* genes and can be grouped to the innate immune response (Tsujimoto et al., 2008). The differential regulation of these pathogen receptors is supported by a recent transcriptomic meta-study covering over 10 publications from human septic blood samples indicating a common activation of pathogen recognition receptors and corresponding signal transduction cascades (Tang et al., 2010).

Further features, which were identified commonly expressed in PCI and PICU, included the repressed MHC II genes, e.g., related to *Antigen processing and presentation*. Among many transcriptomic studies, MHC II related genes were found mostly stable downregulated and may characterize an immune paralysis (Prucha et al., 2004; Tang et al., 2009) related to reduced lymphocyte populations in sepsis due to apoptosis (Hotchkiss and Nicholson, 2006).

Furthermore a lot of unmapped features (not within KEGG pathway definitions for both species) were found and could be subject to enrichment tests in other knowledge bases. Comparison of DEGC to another published experimental sepsis study yielded, e.g., the significantly upregulated *LCN2* gene in the murine CLP

model (Chung et al., 2006), whose protein mediating an innate immune response by sequestering iron (Flo et al., 2004).

Our results regarding the involved genes and pathways are supported by adult data from a recently published time-resolved transcriptomic study (Xiao et al., 2011), comparing a large cohort of severe trauma and burn injury patients to healthy subjects, where a sustained transcriptionally downregulated adaptive immune response (e.g., MHC II genes) as well as an upregulated innate immune response (including, e.g., *TLR*, *LTF*, *LCN2*, and *HP* transcripts) were found. As the degree of deregulation of these genes was able to discriminate between complicated and uncomplicated clinical courses, they might be of interest for novel diagnostic approaches in sepsis and other clinical entities.

LIMITATIONS AND PERSPECTIVES

Limitations of our study include the differences in the wet lab protocols, because variation in the relative proportions of distinct cell types may contain valuable molecular information (Palmer et al., 2006). Furthermore, the number of homologous genes mapped by Entrez IDs across species may influence the results, as observed for the limiting PCI data set (*k*, *l*, *m*). Observed disparities in expression of the DEG, e.g., high average disparity in murine and pediatric sepsis samples for the *CCL4* and *MARCKSL1* genes, may reflect a bias within the mapping or the underlying genomic features. Because their expression can also be affected by missing homologs through complex interaction networks, experimental workarounds may include the use of transgenic or humanized mice (Shultz et al., 2007).

These limitations notwithstanding, we have done a first step, in comparing functional genomic-based experimental sepsis to pediatric patients data for potential modeling applications in translational systems biology. Future studies could also address the impact of different therapeutic approaches in sepsis across species, e.g., the response to antibiotics in combination with immunomodulators in PCI (Bauhofer et al., 2008) on the transcriptomic level.

CONCLUSION

We found highly comparable gene expression patterns in blood samples from a murine sepsis model and pediatric septic patients characterized by commonly expressed genes from the adaptive and innate immune response. Findings point toward a high suitability of an animal sepsis model to study the complex molecular mechanisms and to establish diagnostic as well as treatment options for pediatric sepsis patients.

AUTHOR CONTRIBUTIONS

Sandro Lambeck performed the data analysis and provided access to computations to the reviewers, and drafted the manuscript. Martina Weber carried out the initial wet lab experiments. Martina Weber and Falk A. Gonnert participated in data interpretation. Ralf Mrowka and Michael Bauer supervised the work.

ACKNOWLEDGMENTS

Sandro Lambeck, Falk A. Gonnert, and Michael Bauer were supported by the Federal Ministry of Education and Research (BMBF), Germany, FKZ: 01EO1002 for the Center for Sepsis Control and Care (CSCC). Furthermore, grant support FKZ: 031A113B (quickTLR) is appreciated by Michael Bauer and Ralf Mrowka.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Microbial_Immunology/10.3389/fmicb.2012.00284/abstract

REFERENCES

Bauhofer, A., Huttel, M., Lorenz, W., Sessler, D. I., and Torossian, A. (2008). Differential effects of antibiotics in combination with G-CSF on survival and polymorphonuclear granulocyte cell functions in septic rats. *BMC Infect. Dis.* 8, 55. doi:10.1186/1471-2334-8-55

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

Bone, R., Balk, R., Cerra, F., Dellinger, R., Fein, A., Knaus, W., Schein, R., and Sibbald, W. (1992). Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. *Chest* 101, 1644–1655.

Cavaillon, J. M., and Adib-Conquy, M. (2006). Bench-to-bedside review: endotoxin tolerance as a model of leukocyte reprogramming in sepsis. *Crit. Care* 10, 233.

Chen, R., Li, L., and Butte, A. J. (2007). AILUN: reannotating gene expression data automatically. *Nat. Methods* 4, 879.

Chung, T. P., Laramie, J. M., Meyer, D. J., Downey, T., Tam, L. H., Ding, H., Buchman, T. G., Karl, I., Stormo, G. D., Hotchkiss, R. S., and Cobb, J. P. (2006). Molecular diagnostics in sepsis: from bedside to bench. *J. Am. Coll. Surg.* 203, 585–598.

Dowell, R. D. (2011). The similarity of gene expression between human and mouse tissues. *Genome Biol.* 12, 101.

Du, P., Kibbe, W. A., and Lin, S. M. (2008). Lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24, 1547–1548.

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14863–14868.

Flo, T. H., Smith, K. D., Sato, S., Rodriguez, D. J., Holmes, M. A., Strong, R. K., Akira, S., and Aderem, A. (2004). Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. *Nature* 432, 917–921.

Fury, W., Batliwalla, F., Gregersen, P. K., and Li, W. (2006). Overlapping probabilities of top ranking gene lists, hypergeometric distribution, and stringency of gene selection criterion. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 1, 5531–5534.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y., and Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5, R80.

Gonnert, F. A., Recknagel, P., Seidel, M., Jbeily, N., Dahlke, K., Bockmeyer, C. L., Winning, J., Lösche, W., Claus, R. A., and Bauer, M. (2011). Characteristics of clinical sepsis reflected in a reliable and reproducible rodent sepsis model. *J. Surg. Res.* 170, 123–134.

Han, E. S., and Hickey, M. (2005). Microarray evaluation of dietary restriction. *J. Nutr.* 135, 1343–1346.

Hotchkiss, R. S., and Nicholson, D. W. (2006). Apoptosis and caspases regulate death and inflammation in sepsis. *Nat. Rev. Immunol.* 6, 813–822.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.

Hubbard, W. J., Choudhry, M., Schwacha, M. G., Kerby, J. D., Rue, L. W. III, Bland, K. I., and Chaudry, I. H. (2005). Cecal ligation and puncture. *Shock* 24(Suppl. 1), 52–57.

Ingersoll, M. A., Spanbroek, R., Lottaz, C., Gautier, E. L., Frankenberger, M., Hoffmann, R., Lang, R., Haniffa, M., Collin, M., Tacke, F., Habenicht, A. J., Ziegler-Heitbrock, L., and Randolph, G. J. (2010). Comparison of gene expression profiles between human, and mouse monocyte subsets. *Blood* 115, e10–e19.

Jenner, R. G., and Young, R. A. (2005). Insights into host responses against pathogens from transcriptional profiling. *Nat. Rev. Microbiol.* 3, 281–294.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., and Yamanishi, Y. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, D480–D484.

Kovarik, J., and Siegrist, C. A. (1998). Immunity in early life. *Immunol. Today* 19, 150–152.

Liao, B. Y., and Zhang, J. (2006). Evolutionary conservation of expression profiles between human and mouse orthologous genes. *Mol. Biol. Evol.* 23, 530–540.

Lim, E., Wu, D., Pal, B., Bouras, T., Asselin-Labat, M. L., Vaillant, F., Yagita, H., Lindeman, G. J., Smyth, G. K., and Visvader, J. E. (2010). Transcriptome analyses of mouse, and human mammary cell subpopulations reveal multiple conserved genes, and pathways. *Breast Cancer Res.* 12, R21.

Lu, Y., Huggins, P., and Bar-Joseph, Z. (2009). Cross species analysis of microarray expression data. *Bioinformatics* 25, 1476–1483.

Marshall, J. C. (2008). Sepsis: rethinking the approach to clinical research. *J. Leukoc. Biol.* 83, 471–482.

Mestas, J., and Hughes, C. C. (2004). Of mice and not men: differences between mouse and human immunology. *J. Immunol.* 172, 2731–2738.

Palmer, C., Diehn, M., Alizadeh, A. A., and Brown, P. O. (2006). Cell-type specific gene expression profiles of leukocytes in human peripheral blood. *BMC Genomics* 7, 115. doi:10.1186/1471-2164-7-115

Polpitiya, A. D., McDunn, J. E., Burykin, A., Ghosh, B. K., and Cobb, J. P. (2009). Using systems biology to simplify complex disease: immune cartography. *Crit. Care Med.* 37, 16–21.

Prucha, M., Ruryk, A., Boriss, H., Möller, E., Zazula, R., and Russwurm, S. (2004). Expression profiling: toward an application in sepsis diagnostics. *Shock* 22, 29–33.

Ramasamy, A., Mondry, A., Holmes, C. C., Altman, D. G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.* 5, e184. doi:10.1371/journal.pmed.0050184

Rittirsch, D., Flierl, M. A., and Ward, P. A. (2008). Harmful molecular mechanisms in sepsis. *Nat. Rev. Immunol.* 8, 776–787.

Sandberg, R., and Larsson, O. (2007). Improved precision and accuracy for microarrays using updated probe set definitions. *BMC Bioinformatics* 8, 48. doi:10.1186/1471-2105-8-48

Shi, L., Reid, L. H., Jones, W. D., Shippy, R., Warrington, J. A., Baker, S. C., Collins, P. J., de Longueville, F., Kawasaki, E. S., Lee, K. Y., Luo, Y., Sun, Y. A., Willey, J. M., Setterquist, R. A., Fischer, G. M., Tong, W., Dragan, Y. P., Dix, D. J., Frueh, F. W., Goodsaid, F. M., Herman, D., Jensen, R. V., Johnson, C. D., Lobenhofer, E. K., Puri, R. K., Schrf, U., Thierry-Mieg, J., Wang, C., Wilson, M., Wolber, P. K., Zhang, L., Slikker, W., Jr., Shi, L., and Reid, L. H. (2006). The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotechnol.* 24, 1151–1161.

Shultz, L. D., Ishikawa, F., and Greiner, D. L. (2007). Humanized mice in translational biomedical research. *Nat. Rev. Immunol.* 7, 118–130.

Smyth, G. K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3, 3.

Tang, B., McLean, A., Dawes, I., Huang, S., and Lin, R. (2009). Gene-expression profiling of peripheral blood mononuclear cells in sepsis. *Crit. Care Med.* 37, 882–888.

Tang, B. M., Huang, S. J., and McLean, A. S. (2010). Genome-wide transcription profiling of human sepsis: a systematic review. *Crit. Care* 14, R237.

Tsujimoto, H., Ono, S., Efron, P. A., Scumpia, P. O., Moldawer, L. L., and Mochizuki, H. (2008). Role of Toll-like receptors in the development of sepsis. *Shock* 29, 315–321.

Vodovotz, Y., Csete, M., Bartels, J., Chang, S., and An, G. (2008). Translational systems biology of inflammation. *PLoS Comput. Biol.* 4, e1000014. doi:10.1371/journal.pcbi.1000014

Weber, M., Lambeck, S., Ding, N., Henken, S., Kohl, M., Deigner, H. P., Enot, D. P., Igwe, E. I., Frappart, L., Kiehntopf, M., Claus, R. A., Kamradt, T., Weih, D., Vodovotz, Y., Briles, D. E., Ogunniyi, A. D., Paton, J. C., Maus, U. A., and Bauer, M. (2012). Hepatic induction of cholesterol biosynthesis reflects a remote adaptive response to pneumococcal pneumonia. *FASEB J.* 26, 2424–2436.

Wong, H., Cvijanovich, N., Allen, G. L., Lin, R., Anas, N., Meyer, K., Freishtat, R. J., and Shanley, T. P. (2009). Genomic expression profiling across the pediatric systemic inflammatory response syndrome, sepsis and septic shock spectrum. *Crit. Care Med.* 37, 1558–1566.

Xiao, W., Mindrinos, M. N., Seok, J., Cuschieri, J., Cuenca, A. G., Gao, H., Hayden, D. L., Hennessy, L., Moore, E. E., Minei, J. P., Bankey, P. E., Johnson, J. L., Sperry, J., Nathens, A. B., Billiar, T. R., West, M. A., Brownstein, B. H., Mason, P. H., Baker, H. V., Finnerty, C. C., Jeschke, M. G., López, M. C., Klein, M. B., Gamelli, R. L., Gibran, N. S., Arnoldo, B., Xu, W., Zhang, Y., Calvano, S. E., McDonald-Smith, G. P., Schoenfeld, D. A., Storey, J. D., Cobb, J. P., Warren, H. S., Moldawer, L. L., Herndon, D. N., Lowry, S. F., Maier, R. V., Davis, R. W., Tompkins, R. G., and the Inflammation, and Host Response to Injury Large-Scale Collaborative Research Program. (2011). A genomic storm in critically injured humans. *J. Exp. Med.* 208, 2581–2590.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 December 2011; accepted: 18 July 2012; published online: 14 September 2012.

*Citation: Lambeck S, Weber M, Gonnert FA, Mrowka R and Bauer M (2012) Comparison of sepsis-induced transcriptomic changes in a murine model to clinical blood samples identifies common response patterns. Front. Microbio. **3**:284. doi: 10.3389/fmicb.2012.00284*

This article was submitted to Frontiers in Microbial Immunology, a specialty of Frontiers in Microbiology.