Aspergilli: Models for systems biology in filamentous fungi
Julian Brandl and Mikael R. Andersen

Abstract
Aspergillus is a diverse genus of filamentous fungi including common house hold mold as well as human pathogens. More than 350 species are currently part of this genus and all their genomes are soon to be sequenced. The availability of this vast amount of data will allow for more in-depth understanding of genetic traits governing desirable properties like enzyme production as well as the pathogenic potency of the organisms. In this review we give an overview of the systems biology research conducted in Aspergilli. This research has covered omics technologies like genomics, transcriptomics and proteomics where outstanding contributions are highlighted. From past developments it becomes apparent that CRISPR technology will speed up genetic research in the Aspergillus field. This speed up will allow for an increase in systems biology targeted research by accelerating data generation. The increase in throughput of data generation both per experiment and per time will lead to future challenges in the data handling, integration and interpretation.

Addresses
Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark

Corresponding author: Andersen, Mikael R (mr@bio.dtu.dk)

URL: http://www.staff.dtu.dk/miran

Current Opinion in Systems Biology 2017, 6:67–73
This review comes from a themed issue on Systems biology of model organisms (2017)
Edited by Jens Nielsen and Kiran Raosaheb Patil
For a complete overview see the Issue and the Editorial
Available online 30 September 2017
https://doi.org/10.1016/j.coisb.2017.09.005
2452-3100 © 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords
Aspergillus, Systems biology, Genomics, Proteomics, Transcriptomics, Genome-scale modeling.

Introduction
The genus Aspergillus comprises more than 350 known species of filamentous fungi [1] exhibiting a great variation in lifestyle (e.g. habitat, pathogenicity) and metabolic properties. Probably the best known representative of this genus is Aspergillus nidulans, a genetic model organism that has been extensively studied with respect to metabolism, cellular development, and regulation. The ability to genetically manipulate A. nidulans and combine genetic traits by sexual crossing has been crucial for its pronounced role as genetic model organism.

Besides being used as genetic model organisms for filamentous fungi and eukaryotes in general, other species in the genus are known for their occurrence as mold in buildings as well as on food products. Aspergillus niger is a very common food spoilage fungus which can be explained by the saprophytic lifestyle endowing A. niger with a large repertoire of plant biomass degrading enzymes [2]. While this arsenal enables the undesired growth of A. niger on food, those enzymes are at the same time interesting compounds for biotechnological production [3]. Some of the most prominent enzymes produced in A. niger are glucose oxidase — a common enzyme for the quantification of blood sugar — and α-amylase that is widely used in the processing of starch in the food industry. A. niger is furthermore well known for its ability to accumulate organic acids, most prominently citric acid in very high amounts [4]. Another member of this genus — Aspergillus oryzae — has also been used for centuries in the production of soy sauce and has consequently resulted in A. niger and A. oryzae as Generally Recognized as Safe (GRAS) an important label for the biotechnological production of food components [5]. This label combined with the high productivity in both organisms has established those fungi as industrial work horses for enzyme production of fungal origin.

In addition to these biotech-relevant representatives of the genus, there are other species that are mostly known for being pathogens to humans (Aspergillus flavus and Aspergillus fumigatus) causing aspergillosis in immuno-compromised patients resulting in difficult and potentially life-threatening infections [6] or plants (A. flavus and A. niger), where the fungus spoils crops on the field or post-harvest. Having this wide variety of organisms that are of relevance for humans with regard to health and biotechnological applications sparked an early interest in conducting research on a systems level in many of those fungi. For these reasons — and due to the diverse and active Aspergillus research community — the genus has become a model genus for many types of studies. In particular, the advent of genome sequencing...
has paved the way for the application of several -omics technologies for the elucidation of metabolic phenotypes, responses to perturbations as well as for the identification of differences and commonalities between individual organisms.

In this review, we will outline the -omics and systems biology contributions in the Aspergillus genus, and how they have made them one of the leading genera in omics-based research, if not the leading genus.

Systems biology in Aspergilli
An important development for the understanding of biology on a systems level is the development in the field of the “omics” technologies i.e. genomics, transcriptomics, proteomics and metabolomics. These technologies allow the acquisition of a snapshot of state of the system allowing the unbiased observation of system-wide effects induced by environmental and temporal changes.

Genomics
Genomics research in the genus of Aspergillus has received most attention since the start of the genomics era in filamentous fungi. Even though the Neurospora community published the fungal genome for Neurospora crassa in 2003 [7], the Aspergillus field was quick to follow with the triple publications of A. nidulans [8], A. oryzae [9], and A. fumigatus [10] in a 2005 issue of Nature. Since the publication of the first three Aspergillus genomes, genomes sequenced from the Aspergillus genus have been diverse and plentiful. First, the Broad Institute released the genome of Aspergillus terreus in 2006, followed by the sequencing of A. niger CBS 513.88 [11], an important industrial work horse. A large number of single genomes have been published since then, such as Aspergillus parasiticus [12], Aspergillus udagawae [13], A. flavus [14], Aspergillus caldusatus [15], Aspergillus luchuensis [16], Aspergillus bombayensis [17], or even isolates from the International Space Station [18] to name a few.

A part of the success of genomics research in Aspergilli is the community support as seen in the manual curation of the A. nidulans genome [19]. The activity of the community is also evident in the establishment of the active and curated genome repository, the Aspergillus genome database (AspGD) [20,21], which has now been continued in FungiDB [22].

A large part of the genomics research in Aspergilli has been published in the form of comparative genomics studies, where new genomes are compared to previously published ones. One notable example is the first comparative study of the genome sequence of three strains of the same species, the wild type citric acid producer A. niger ATCC 1015, the industrial enzyme-producer A. niger CBS 513.88, and the NRRL3 laboratory strain [23]. In a similar study performed in 2008 [24] the authors compared two isolates of A. fumigatus. From these studies it became particularly evident that there is a large amount of species diversity at the genomic level in Aspergilli. Comparative genomics has also been employed for analysis of diversity in published genomes for specific traits, e.g. intragenic tandem repeats [25] or carbohydrate-active enzymes [26], showing how re-analysis of genomes across specific sets can generate new insights.

Recently, Aspergillus comparative genomics has increased in scale. The most prominent example is the community-driven analysis of 19 Aspergillus genomes, 10 of them from new de novo genomes [27]. This type of analysis will only expand in the near future, as a part of the community is sequencing and annotating a member of all 350 species currently defined in the Aspergillus genus. This project has so far released genomes for more than 25 Aspergillus species, and have completed an additional 150 which are about to be published. The prospects of having access to the diversity of Aspergilli at the genome level are fascinating and will allow many future studies of fungal genomics at so far unseen levels. This has also led to proposals of similar studies for the Penicillium and Trichoderma genera. Thus, the extent and diversity of Aspergillus genomics has become leading in the area of fungal genomics, and is expected to remain so for the foreseeable future.

Transcriptomics
The availability of selected genomes of the Aspergillus genus allowed it to enter the transcriptomics era relatively early [28,29]. With microarrays being available, the first comparisons of system-wide effects of experimental perturbations, in some cases called comparative transcriptomics [29–31] was possible. Another application has been the early (2011) single-cell study of hyphal differentiation [32] in A. niger.

The Aspergillus area has also been relatively early adopters of RNA-seq, with the first study based on A. oryzae published in 2010 [33], and currently about 30 studies published.

One specific and efficient application of transcriptomics data is the elucidation and study of metabolic pathways. Using comparative transcriptomics between two nutrient conditions allowed the putative identification of genes involved in the catabolism of galactaric acid by Kuivanen et al. [34] by screening for upregulated genes, or in another case to identify genes involved in itaconic acid production [35].

Another powerful area of application for transcriptomics based experiments has been the observation of global
changes to environmental stresses like hypoxia [36–38], pH [39], and heat [10], shedding light on mechanisms allowing filamentous fungi to cope with these changes that might lead to insights for understanding fungal durability and towards targets for combating them [40]. In the case of hypoxia, A. fumigatus seems to possess a robust cellular response involving the upregulation of e.g. glycolysis, iron uptake and ergosterol biosynthesis and also led to the identification of a potential antifungal target in the Q10 biosynthesis pathway [41].

In summary, Aspergillus transcriptomics studies are abundant and have given insight on key traits in the fungal lifestyle and differentiation, to name just a few areas.

Proteomics
As transcriptomics analysis is frequently used as an approximation for proteomic changes it is not surprising that proteomics data has been used in similar ways as described for transcriptomics data above. An area of special interest in the proteomics of Aspergilli has been the set of secreted biomass degrading enzymes.

One example for the analysis of proteomics data for differences in the secretome between conditions is the study by Lu et al. [42]. The authors collected data for both the intra- and extracellular proteome under different conditions thereby proving and excellent insight conditional changes in the proteome of A. niger. A similar study comparing the secretome of A. fumigatus on three different carbon sources [43] additionally found many enzymes to be post-translationally modified by deamination. Comparing substrate dependent differences in membrane associated proteins has led to the identification of transporters for glucose [44], as well as to the first eukaryotic L-rhamnose transporter [45].

The combination of proteomics and transcriptomics has also been used as an important tool in understanding the development of several Aspergilli. As filamentous fungi propagate by forming conidia spores, processes governing the germination of spores are of vital importance. Using a combination of transcriptomic and proteomic data, Novodvorska et al. [46] have been able to show, that fungal spores are still metabolically active and switch from fermentative to respiratory growth upon triggering outgrowth of the spores. Similarly Anjo et al. [47] found that spores of A. fumigatus proceed from a metabolic active state via dormancy to an autophagic state under nutrient limitation.

Integration of multi-omics data sets has proven useful for the identification of genes contributing to complex biological traits like protein secretion. A notable example for the integration and transcriptomics data is the identification of leads for strain improvement by Jacobs et al. [48]. Using the combination of these technologies allowed the authors to identify metabolic engineering targets based on the correlation of changes in transcript with changes in protein level. Those genes represent good candidates for genetic engineering as there is no obvious regulation on another level like post-translational regulation.

Metabolomics
Metabolomics on a systems level has been applied to analyze the underlying mechanism of productivity by comparing strains with varying levels of production. Metabolic flux analysis (MFA) is used for the investigation of intracellular carbon fluxes and has been applied to A. niger and A. oryzae. Comparing the flux distribution of A. niger overexpressing fructofuranosidase to the parent strain Driouch et al. [49] found a redistribution the metabolic flux to the Pentose Phosphate Pathway (PPP) leading to an increase in NADPH production. The same observation has been made by Lu et al. [50] in high glucoamylase producer of A. niger as well as by Pedersen et al. [51] in A. oryzae producing alpha-amylase. The observed modulation of the flux through the PPP therefore appears to be a general feature of Aspergilli expressing recombinant proteins.

Another area of research that has gained a lot of interest in Aspergilli is the biosynthesis of secondary metabolites. The diversity of secondary metabolites produced by different Aspergilli has been reviewed elsewhere [1] and the information about individual metabolites has been collected in the A2MDB database [52]. One of the challenges in this area is that most of the biosynthetic gene clusters are inactive under standard laboratory conditions. Targeted methods employing promoter exchange of individual genes [53] or heterologous expression [54] as well as untargeted approaches like epigenetic modifiers [55] have proven successful in identifying new compounds and the corresponding gene clusters.

Modeling
Mathematical modeling is at the core of systems biology in order to integrate information and quantitatively analyze phenotypes thereby allowing predictions about the behavior of the organism or elucidating mechanisms underlying experimental observations. The organism which sparked the most interest in developing a model is A. niger and the unique citric acid accumulation phenotype. The metabolic network of A. niger has therefore been subject to several modeling efforts to elucidate key aspects of the high citric acid production. Stoichiometric modeling has been pioneered in filamentous fungi with the reconstruction of the central metabolism of A. niger [56] and A. nidulans [57] which paved the way for the development of genome-scale models. The first genome-scale model of A. niger has
been developed by our group in 2008 [58] and has recently been updated by Liu et al. [59]. Similarly genome-scale models are available for *A. terreus* [60], *A. oryzae* [61] and *A. nidulans* [62].

Besides these genome-scale stoichiometric models, kinetic models have been developed and analyzed for different aspects of *Aspergillus* physiology. Separate models have been developed for the catabolism of arabinose [63] and xylose [64] as well as for the citric acid production [65] in *A. niger* in order to identify key enzymatic steps as metabolic engineering targets.

**Discussion**

As reviewed above, the early adaptation of omics-methods *Aspergillus* research has brought significant contributions within a broad range of omics, showing how the Aspergilli are suitable models for study of fungal-specific, microbe-specific, and general eukaryotic traits. Seeing how rapid the development has been since the publication of the first genomes in 2005, it will be interesting to follow how the field develops. Based on the recent developments in the area, we will hazard some guesses.

One trend that we believe to see coming up, is how future research in Aspergilli will shift more towards the development and utilization of high-throughput technologies. One of the major hurdles for performing genetic research in Aspergilli in a systematic manner has been the low efficiency for targeted integration and gene knockout. Since the introduction of the CRISPR/Cas9 technology to Aspergilli [66] the efficiency of targeted genetic manipulation has been boosted significantly [34,67,68] enabling the assessment of a large number of those modifications in a high-throughput manner. Leveraging this gain in strain development will speed up the elucidation of pathways both in primary and secondary metabolism as well as enabling the development of genome-wide screens e.g. knockout libraries as has been performed in other model systems [69]. Similarly a catalytically inactive variant of Cas9 (dCas9) might be used for high-throughput elucidation of transcriptional regulation by performing knock-down or activation experiments of individual transcription factors.

Another development could be faster identification of genes contributing to a specific phenotype. One example of this trend is the elucidation of secondary metabolite biosynthesis. Today the identification of new secondary metabolites and the mapping to the corresponding gene-cluster is cumbersome. Several efforts have been made in the past in order to speed up the process and developing methods for scaling-up this process leading to a more rapid screening of whole libraries of gene clusters [54]. In combination with large-scale comparative genomics, as seen in *Penicillium* [70], this could accelerate the field.

One of the existing challenges for a true understanding of the system is the lack of information about the function of several genes [71]. At the same time, information about the subcellular localization of individual proteins is even scarcer than for functional annotation, but has been shown to even be important for understanding the cellular context of known functions [72].

A final trend concerns how the masses of new data will be treated and stored. As the field is moving towards larger scale experimental work as well as the acquisition of increasing amounts of data per experiment, new challenges with respect to storing, handling and integration of this data arise in order to enable a systems level understanding. Some of the challenges for future research will be the management and the infrastructure for handling and integrating omics-data and experimental data generated by high-throughput experiments. At the moment, since data storage and metadata is not standardized, a lot of omics-data becomes “single-use”, meaning that it is only analyzed in the original publication, due to poor accessibility for other researchers. This is both a waste and slows down the progress of the field. One of the initiatives aiming at developing solutions for this problem is the EU project called ELIXIR which aims at developing a joint infrastructure for life science research.

While storage and sharing of this data is a challenge in itself, the integration and joint interpretation tops the challenges of data handling in complexity. As biological systems are inherently complex, a systems level understanding will only be achieved by integrating all available data in a coherent framework. Fungal research is already moving in this direction with the integration of large datasets in FungiDB and its sophisticated methods for data queries, but sustained support and use of these functions will be required for reaching the full potential of genome-wide studies and engineering.

**Acknowledgement**

The authors acknowledge funding from the Novo Nordisk Foundation, Grant NNF13OC0004831, Biotechnology-based Synthesis and Production.

**References**

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest


organisms. Straining conservation of non-coding regulatory regions between those organisms.


Early comparative genomics study of 3 Aspergillus species demonstrating conservation of non-coding regulatory regions between those organisms.


Publication and comparative analysis of 10 new Aspergillus genomes and comprehensive review of research considering metabolism and genetics in this field.


Aspergillus systems biology Brandl and Andersen 71

Current Opinion in Systems Biology 2017, 6:67–73


53. Alvarez-Vasquez F, Gonzalez-Atalco C, Torres NV: Metabolism of citric acid production by Aspergillus niger: model


