Linking secondary metabolites to gene clusters through genome sequencing of six diverse Aspergillus species

The fungal genus of Aspergillus is highly interesting, containing everything from industrial cell factories, model organisms, and human pathogens. In particular, this group has a prolific production of bioactive secondary metabolites (SMs). In this work, four diverse Aspergillus species (A. campestris, A. novofumigatus, A. ochraceoroseus, and A. steynii) have been whole-genome PacBio sequenced to provide genetic references in three Aspergillus sections. A. taichungensis and A. candidus also were sequenced for SM elucidation. Thirteen Aspergillus genomes were analyzed with comparative genomics to determine phylogeny and genetic diversity, showing that each presented genome contains 15–27% genes not found in other sequenced Aspergilli. In particular, A. novofumigatus was compared with the pathogenic species A. fumigatus. This suggests that A. novofumigatus can produce most of the same allergens, virulence, and pathogenicity factors as A. fumigatus, suggesting that A. novofumigatus could be as pathogenic as A. fumigatus. Furthermore, SMs were linked to gene clusters based on biological and chemical knowledge and analysis, genome sequences, and predictive algorithms. We thus identify putative SM clusters for aflatoxin, chlorflavonin, and ochrindol in A. ochraceoroseus, A. campestris, and A. steynii, respectively, and novofumigatonin, ent-cycloechinulin, and epi-aszonalenins in A. novofumigatus. Our study delivers six fungal genomes, showing the large diversity found in the Aspergillus genus; highlights the potential for discovery of beneficial or harmful SMs; and supports reports of A. novofumigatus pathogenicity. It also shows how biological, biochemical, and genomic information can be combined to identify genes involved in the biosynthesis of specific SMs.

General information
Publication status: Published
Organisations: Department of Biotechnology and Biomedicine, Network Engineering of Eukaryotic Cell factories, Fungal Chemodiversity, Natural Product Discovery, Department of Biotechnology, Eukaryotic Molecular Cell Biology, United States Department of Energy, University of Manchester
Corresponding author: Andersen, M. R.
Pages: E753-E761
Publication date: 23 Jan 2018
Peer-reviewed: Yes

Publication information
Journal: Proceedings of the National Academy of Sciences of the United States of America
Volume: 115
Issue number: 4
Efficient Oligo nucleotide mediated CRISPR-Cas9 Gene Editing in Aspergilli

CRISPR-Cas9 technologies are revolutionizing fungal gene editing. Here we show that survival of specific Cas9/sgRNA mediated DNA double strand breaks (DSBs) depends on the non-homologous end-joining, NHEJ, DNA repair pathway and we use this observation to develop a tool to assess protospacer efficiency in Aspergillus nidulans. Moreover, we show that in NHEJ deficient strains, highly efficient marker-free gene targeting can be performed. Indeed, we show that single-stranded oligo nucleotides efficiently works as repair templates of specific Cas9/sgRNA induced DNA DSBs in A. nidulans, A. niger, and in A. oryzae indicating that this type of repair may be wide spread in filamentous fungi. Importantly, we demonstrate that by using single-stranded oligo nucleotides for CRISPR-Cas9 mediated gene editing it is possible to introduce specific point mutations as well gene deletions at efficiencies approaching 100%. The efficiency of the system invites for multiplexing and we have designed a vector system with the capacity of delivering Cas9 and multiple sgRNAs based on polymerase III promoters and tRNA spacers. We show that it is possible to introduce two point mutations and one gene insertion in one transformation experiment with a very high efficiency. Our system is compatible with future high-throughput gene-editing experiments.
**Friends and Foes - Comparative Genomics of 23 Aspergillus Flavi Species**

Flavi is a highly diverse section in genus Aspergillus encompassing species used in food fermentation and enzyme production (A. oryzae and A. sojae) as well as toxigenic and foodpoilers (A. parasiticus and A. flavus) in addition to many less studied species. Here we have whole genome sequenced 19 Flavi species and used comparative genomic tools to investigate the section. We have examined similarities and differences of this section with a special focus on the carbohydrate active enzymes (CAZy) and secondary metabolites to get an understanding of what is unique features for section Flavi. In addition we have studied selected secondary metabolite gene clusters (SMGC) found widely across the section to get an understanding of cluster evolution and development.

**General information**

Publication status: Published
Organisations: Section for Synthetic Biology, Network Engineering of Eukaryotic Cell factories, Department of Biotechnology and Biomedicine, New Bioactive Compounds, Novo Nordisk Foundation Center for Biosustainability, Section for Microbial and Chemical Ecology, Fungal Chemodiversity, Eukaryotic Molecular Cell Biology, Natural Product Discovery, Kikkoman Corporation, Aix-Marseille University, Joint Genome Institute, Aix-Marseille Université, Pacific Northwest National Laboratory
Number of pages: 1
Publication date: 2018
Peer-reviewed: Yes
Event: Poster session presented at 14th European Conference on Fungal Genetics, Haifa, Israel.

**Investigation of inter- and intraspecies variation through genome sequencing of Aspergillus section Nigri**

Aspergillus section Nigri comprises filamentous fungi relevant to biomedicine, bioenergy, health, and biotechnology. To learn more about what genetically sets these species apart, as well as about potential applications in biotechnology and biomedicine, we sequenced 23 genomes de novo, forming a full genome compendium for the section (26 species), as well as 6 Aspergillus niger isolates. This allowed us to quantify both inter- and intraspecies genomic variation. We further predicted 17,903 carbohydrate-active enzymes and 2,717 secondary metabolite gene clusters, which we condensed into 455 distinct families corresponding to compound classes, 49% of which are only found in single species. We performed metabolomics and genetic engineering to correlate genotypes to phenotypes, as demonstrated for the metabolite aurasperone, and by heterologous transfer of citrate production to Aspergillus nidulans. Experimental and computational analyses showed that both secondary metabolism and regulation are key factors that are significant in the delineation of Aspergillus species.

**General information**

Publication status: Published
Organisations: Department of Biotechnology and Biomedicine, Section for Synthetic Biology, Network Engineering of Eukaryotic Cell factories, Novo Nordisk Foundation Center for Biosustainability, New Bioactive Compounds, Section for Microbial and Chemical Ecology, Fungal Chemodiversity, Natural Product Discovery, Fungal Biomedicine and Biology, Eukaryotic Molecular Cell Biology
Corresponding author: Andersen, M. R.
Pages: 1688–1695
Publication date: 2018
Peer-reviewed: Yes
The Aspmine - Comparative genomics analysis of 6 new species of Aspergillus section Sparsi, Ochraceorosei, Tanneri and Robusti

Background: Sequencing genomes of filamentous fungi will reveal the genetic mechanisms that lead to a high diversity and frequency of bioactive compounds.

Conclusions: Sequencing of new species identifies thousands of new proteins and gene clusters involved in metabolite production.

Genetic diversity of 100+ Aspergillus species: The aspMine analysis resource

The Aspergillus Mine - publishing bioinformatics.

Genome analysis is no longer a field reserved for specialists and experimental laboratories are doing groundbreaking research using genome sequencing and analysis. In this new era, it is essential that data, analysis and results are shared.
between scientists. But this can be a challenge, even more so with no computational specialist. Here we present a setup for analysis and publication of genome data of 70 species of Aspergillus fungi. The platform is based on R, Python and uses the RShiny framework to create interactive web-applications. It allows all participants to create interactive analysis which can be shared with the team and in connection with publications. We present analysis for investigation of genetic diversity, secondary and primary metabolism and general data overview. The platform, the Aspergillus Mine, is a collection of analysis tools based on data from collaboration with the Joint Genome Institute. The Aspergillus Mine is not intended as a genomic data sharing service but instead focuses on creating an environment where the results of bioinformatic analysis is made available for inspection. The data and code is public upon request and figures can be obtained directly from the web-app. This resource will be of great benefit to the Aspergillus community which is in a rapid development in regards to genome sequencing and analysis. At the moment, the service includes analysis of more than 70 genomes, and is expected to double in the next 6 months, with the final goal of the project is the analysis of 300 Aspergillus species.

**General information**
Publication status: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Metabolomics Platform, Eucaryotic Molecular Cell Biology, CBS-KNAW Fungal Biodiversity Centre, Joint Genome Institute, Joint Bioenergy Institute
Number of pages: 1
Publication date: 2016
Peer-reviewed: Yes
Keywords: Comparative genomics, Bioinformatic, Fungi, Aspergillus, Data publishing

A CRISPR/Cas9 system for genetic engineering of filamentous fungi

The number of fully sequenced fungal genomes is rapidly increasing. Since genetic tools are poorly developed for most filamentous fungi, it is currently difficult to employ genetic engineering for understanding the biology of these fungi and to fully exploit them industrially. For that reason there is a demand for developing versatile methods that can be used to genetically manipulate non-model filamentous fungi. To facilitate this, we have developed a CRISPR-Cas9 based system adapted for use in filamentous fungi. The system is simple and versatile, as RNA guided mutagenesis can be achieved by
transforming a target fungus with a single plasmid. The system currently contains four CRISPR-Cas9 vectors, which are equipped with commonly used fungal markers allowing for selection in a broad range of fungi. Moreover, we have developed a script that allows identification of protospacers that target gene homologs in multiple species to facilitate introduction of common mutations in different filamentous fungi. With these tools we have performed RNA-guided mutagenesis in six species of which one has not previously been genetically engineered. Moreover, for a wild-type *Aspergillus aculeatus* strain, we have used our CRISPR Cas9 system to generate a strain that contains an AACU_pyrG marker and demonstrated that the resulting strain can be used for iterative gene targeting.

**General information**
- Publication status: Published
- Organisations: Department of Systems Biology, Eucaryotic Molecular Cell Biology
- Contributors: Nødvig, C. S., Nielsen, J. B., Kogle, M. E., Mortensen, U. H.
- Number of pages: 18
- Publication date: 2015
- Peer-reviewed: Yes

**Publication information**
- Journal: *P L o S One*
- Volume: 10
- Issue number: 7
- Article number: e0133085
- ISSN (Print): 1932-6203
- Ratings:
  - BFI (2015): BFI-level 1
  - Scopus rating (2015): CiteScore 3.32 SJR 1.427 SNIP 1.149
- Web of Science (2015): Indexed yes
- Original language: English
- Electronic versions:
  - A_Crispr_Cas9_System_for_Genetic.pdf
- DOIs:
  - 10.1371/journal.pone.0133085

**Bibliographical note**
This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Research output:** Contribution to journal › Journal article – Annual report year: 2015 › Research › peer-review

**Diversity of carbohydrate metabolism in species of *Aspergillus***
The filamentous fungus *Aspergillus niger* and its close relatives in *Aspergillus* section *Nigri* are of broad interest to the scientific community including applied, medical and basic research. The fungi are prolific producers of native and heterologous proteins, organic acids (in particular citrate), and secondary metabolites (including bioactives and toxins such as ochratoxin A). Because of these abilities they represent a substantial economic interests in bioenergy applications. In a project collaboration with the US Joint Genome Institute and JBEI we are sequencing 300 different species of *Aspergillus* and establishing an online analysis platform for the scientific community, *aspMine300*.

**General information**
- Publication status: Published
- Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Molecular Biology, Metabolomics Platform, Eucaryotic Molecular Cell Biology, Aix-Marseille University, Joint Bioenergy Institute, Joint Genome Institute, CBS-KNAW Fungal Biodiversity Centre
- Number of pages: 1
- Publication date: 2015
- Peer-reviewed: Yes
- Event: Abstract from Exploring the genomic complexity and diversity of eukaryotes, Sant Feliu de Guixols, Spain.
- Source: PublicationPreSubmission
- Source ID: 124072321
- Research output: Contribution to conference › Conference abstract for conference – Annual report year: 2015 › Research › peer-review

**Gene cluster dynamics throughout the *Aspergillus* genus***
In the era of high-throughput sequencing, comparative genomics can be applied for evaluating species diversity. In this project, we aim to compare the genomes of 300 species of filamentous fungi from the Aspergillus genus, a complex task. To be able to define species, clade, and core features, this project uses BLAST on the amino acid level to discover orthologs. With a potential of 300 Aspergillus species each having ~12,000 annotated genes, traditional clustering would demand supercomputing. Instead, our approach reduces the research space by identifying isoenzymes within each genome creating intragenomic protein families (iPFs), and then connecting iPFs across all genomes. The initial findings in a set of 31 species show that ~48% of the annotated genes are core genes (genes shared between all species) and 2-24% of the genes are defining the individual species.

The methods presented here will allow for detailed investigation into mapping of genotypetophenotype across a very large set of genomes without losing information.