Research outputs:

**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 novofumigatontin, ent-cycloechinulin, and epiaszonalenins 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.
Ratings:
BFI (2019): BFI-level 2
Web of Science (2019): Indexed yes
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 8.59 SJR 6.092 SNIP 2.626
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 8.56 SJR 6.576 SNIP 2.642
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 8.84 SJR 6.814 SNIP 2.691
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): CiteScore 8.86 SJR 6.898 SNIP 2.734
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): CiteScore 9.5 SJR 7.073 SNIP 2.738
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): CiteScore 9.49 SJR 6.868 SNIP 2.697
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): CiteScore 9.31 SJR 6.864 SNIP 2.646
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 6.898 SNIP 2.545
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 7.025 SNIP 2.556
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 7.034 SNIP 2.449
Web of Science (2008): Indexed yes
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 6.849 SNIP 2.45
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 6.94 SNIP 2.555
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 7.197 SNIP 2.629
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 7.129 SNIP 2.515
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 6.913 SNIP 2.503
Web of Science (2002): Indexed yes
Scopus rating (2001): SJR 7.189 SNIP 2.47
Web of Science (2001): Indexed yes
Scopus rating (2000): SJR 8.751 SNIP 2.458
**Friends and foes - comparative genomics of 23 Aspergillus Flavi species**

**General information**

State: 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, Natural Product Discovery, Eukaryotic Molecular Cell Biology, Joint Genome Institute, Pacific Northwest National Laboratory
Number of pages: 1
Publication date: 2018
Peer-reviewed: Yes
Event: Abstract from 14th European Conference on Fungal Genetics, Haifa, Israel.

Electronic versions:
Abstract_ECFG2018_final_pdf.pdf

Research output: Research - peer-review › Journal article – Annual report year: 2018

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 foodspoilers (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.

**Immense diversity found in secondary metabolite gene clusters in filamentous fungi and bacteria using comparative genomics**

Secondary metabolism in microorganisms is defined as non-life essential metabolism such as the production of mycotoxins and antibiotic compounds. This definition has to some extent been obscured as it is hard to define what is essential for the life of a microorganism in natural environments. Who is to say that the secondary metabolite penicillin produced by species of Penicillium is not essential for the life of the fungi. The compounds produced in these reactions are possible antimicrobial, biofuels, food spoilers and other important or valuable compounds. Secondary metabolism is, therefore, the subject of great curiosity for both biological and financial reasons. At the Center for Microbial Secondary Metabolites (CeMiSt) we are interested in the natural role of secondary metabolites and their response to other organisms. Here we present the analysis of four different sets of organisms that are of interest in the study of natural environments like...
soil and food products. The number of known secondary metabolite compounds far exceed the number of characterized genes involved in these pathways and the effort required to elucidate the connection between metabolites and genes require time-consuming and expensive chemical characterization. Because of the benefits and challenges in this elucidation much effort has been given to the computational prediction of genes involved in secondary metabolism, with the most prominent software being AntiSMASH (bacteria and fungi) and SMURF (fungi). These methods provide valuable knowledge in the selection of gene targets for further analysis and can increase the speed of gene to compound association tremendously. Although taxonomy groups living organisms into specific groups at different levels, genera, species and so on, the diversity within each group can vary. We investigated the general conservation of protein function, using protein sequences, by an all against all BLASTp comparison. The fraction, as a function of the total number of proteins in a species, of shared proteins was then calculated and presented in a matrix in Figure 2. The cutoff for significant hits is defined as a reciprocal BLAST [Altschul (1990)] hit with percent identity of over 50% and alignment coverage percentage sum (query and hit coverage) over 130%. The diagonal line is a self comparison, and is therefore always 100% (purple). We see that the overall diversity between species is much larger for Bacilli (20-40%) than for Pseudomonas (50-70%). The Penicillium are as diverse as the Bacillus while the Aspergilli are the most diverse of all the sets.

General information
State: Published
Organisations: Section for Synthetic Biology, Network Engineering of Eukaryotic Cell factories, Department of Biotechnology and Biomedicine, Section for Microbial and Chemical Ecology, Fungal Chemodiversity, Bacterial Interactions and Evolution, Infection Microbiology, New Bioactive Compounds, Novo Nordisk Foundation Center for Biosustainability, Joint Bioenergy Institute
Number of pages: 1
Publication date: 2018
Peer-reviewed: Yes
Event: Poster session presented at 17th European Conference on Computational Biology, Athens, Greece.
Source: PublicationPreSubmission
Source-ID: 158210001
Research output: Research - peer-review » Poster – Annual report year: 2018

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
State: 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
Pages: 1688–1695
Publication date: 2018
Peer-reviewed: Yes

Publication information
Journal: Nature Genetics
Volume: 50
ISSN (Print): 1061-4036
Ratings:
BFI (2019): BFI-level 3
Web of Science (2019): Indexed yes
BFI (2018): BFI-level 3
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 21.12 SJR 22.243 SNIP 5.867
Web of Science (2017): Impact factor 27.125
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 20.83 SJR 21.979 SNIP 6.709
Web of Science (2016): Impact factor 27.959
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Web of Science (2015): Impact factor 31.616
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): CiteScore 22.76 SJR 23.986 SNIP 6.332
Web of Science (2014): Impact factor 29.352
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): CiteScore 24.17 SJR 24.193 SNIP 6.287
Web of Science (2013): Impact factor 29.648
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): CiteScore 27.17 SJR 25.621 SNIP 7.137
Web of Science (2012): Impact factor 35.209
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): CiteScore 25.75 SJR 25.298 SNIP 7.206
Web of Science (2011): Impact factor 35.532
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Web of Science (2010): Impact factor 36.377
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 20.87 SNIP 5.222
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 17.931 SNIP 4.809
Web of Science (2007): Indexed yes
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 14.345 SNIP 5.272
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 13.814 SNIP 5.329
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 13.523 SNIP 5.059
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 13.631 SNIP 5.2
Novofumigatonin biosynthesis involves a non-heme iron-dependent endoperoxide isomerase for orthoester formation

Novofumigatonin (1), isolated from the fungus Aspergillus novofumigatus, is a heavily oxygenated meroterpenoid containing a unique orthoester moiety. Despite the wide distribution of orthoesters in nature and their biological importance, little is known about the biogenesis of orthoesters. Here we show the elucidation of the biosynthetic pathway of 1 and the identification of key enzymes for the orthoester formation by a series of CRISPR-Cas9-based gene-deletion experiments and in vivo and in vitro reconstitutions of the biosynthesis. The novofumigatonin pathway involves endoperoxy compounds as key precursors for the orthoester synthesis, in which the Fe(II)/α-ketoglutarate-dependent enzyme NvfI performs the endoperoxidation. NvE, the enzyme catalyzing the orthoester synthesis, is an Fe(II)-dependent, but cosubstrate-free, endoperoxide isomerase, despite the fact that NvE shares sequence homology with the known Fe(II)/α-ketoglutarate-dependent dioxygenases. NvfE thus belongs to a class of enzymes that gained an isomerase activity by losing the α-ketoglutarate-binding ability.
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.

General information
State: 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, Joint Genome Institute, Westerdijk Fungal Biodiversity Institute, Joint Bioenergy Institute, Aix-Marseille University
Number of pages: 1
Publication date: 2018
Peer-reviewed: Yes
Event: Poster session presented at 14th European Conference on Fungal Genetics, Haifa, Israel.
Source: PublicationPreSubmission
Source-ID: 158209954
Research output: Research - peer-review > Poster – Annual report year: 2018
although this did not always correlate to the ability of the species to use the corresponding sugar as a carbon source.
Uncovering secondary metabolite evolution and biosynthesis using gene cluster networks and genetic dereplication

The increased interest in secondary metabolites (SMs) has driven a number of genome sequencing projects to elucidate their biosynthetic pathways. As a result, studies revealed that the number of secondary metabolite gene clusters (SMGCs) greatly outnumbers detected compounds, challenging current methods to dereplicate and categorize this amount of gene clusters on a larger scale. Here, we present an automated workflow for the genetic dereplication and analysis of secondary metabolism genes in fungi. Focusing on the secondary metabolite rich genus *Aspergillus*, we categorize SMGCs across genomes into SMGC families using network analysis. Our method elucidates the diversity and dynamics of secondary metabolism in section *Nigri*, showing that SMGC diversity within the section has the same magnitude as within the genus. Using our genome analysis we were able to predict the gene cluster responsible for biosynthesis of malformin, a potentiator of anti-cancer drugs, in 18 strains. To proof the general validity of our predictions, we developed genetic engineering tools in *Aspergillus brasiliensis* and subsequently verified the genes for biosynthesis of malformin.

General information

State: Published
Organisations: New Bioactive Compounds, Network Engineering of Eukaryotic Cell factories, Department of Biotechnology and Biomedicine, Section for Synthetic Biology, Eukaryotic Molecular Cell Biology, Section for Microbial and Chemical Ecology, Fungal Chemodiversity, Natural Product Discovery, Fungal Biomedicine and Biology, Joint Genome Institute, Federal University of Viçosa
Number of pages: 12
Publication date: 2018
Peer-reviewed: Yes

Publication Information

Journal: Scientific Reports
Volume: 8
Article number: 17957
ISSN (Print): 2045-2322
Ratings:
BFI (2019): BFI-level 1
Web of Science (2019): Indexed yes
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 4.36 SJR 1.533 SNIP 1.245
Web of Science (2017): Impact factor 4.122
Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus Aspergillus

Background:
The fungal genus Aspergillus is of critical importance to humankind. Species include those with industrial applications, important pathogens of humans, animals and crops, a source of potent carcinogenic contaminants of food, and an important genetic model. The genome sequences of eight aspergilli have already been explored to investigate aspects of fungal biology, raising questions about evolution and specialization within this genus.

Results:
We have generated genome sequences for ten novel, highly diverse Aspergillus species and compared these in detail to sister and more distant genera. Comparative studies of key aspects of fungal biology, including primary and secondary metabolism, stress response, biomass degradation, and signal transduction, revealed both conservation and diversity among the species. Observed genomic differences were validated with experimental studies. This revealed several highlights, such as the potential for sex in asexual species, organic acid production genes being a key feature of black aspergilli, alternative approaches for degrading plant biomass, and indications for the genetic basis of stress response. A genome-wide phylogenetic analysis demonstrated in detail the relationship of the newly genome sequenced species with other aspergilli.

Conclusions:
Many aspects of biological differences between fungal species cannot be explained by current knowledge obtained from genome sequences. The comparative genomics and experimental study, presented here, allows for the first time a genus-wide view of the biological diversity of the aspergilli and in many, but not all, cases linked genome differences to phenotype. Insights gained could be exploited for biotechnological and medical applications of fungi.

General information
State: Published
Organisations: Department of Biotechnology and Biomedicine, Network Engineering of Eukaryotic Cell factories, Fungal Chemodiversity, Westerdijk Fungal Biodiversity Institute, Utrecht University, United States Department of Energy,
Expansions and reductions in fungal primary metabolism studied across 100 fungal species

General information
State: Published
Organisations: Department of Biotechnology and Biomedicine, Network Engineering of Eukaryotic Cell factories
Contributors: Brandl, J., Rasmussen, J. L. N., Vesth, T. C., Andersen, M. R.
Number of pages: 1
Publication date: 2017
Peer-reviewed: Yes
Event: Abstract from 29th Fungal Genetics Conference, Pacific Grove, United States.
Electronic versions:
Expansions and reductions in fungal primary metabolism studied across 100 fungal species
Research output: Research - peer-review › Conference abstract for conference – Annual report year: 2017

Genetic diversity of 100+ Aspergillus species - the aspMine analysis resource

General information
State: Published
Organisations: Department of Biotechnology and Biomedicine, Network Engineering of Eukaryotic Cell factories, Novo Nordisk Foundation Center for Biosustainability, New Bioactive Compounds, CBS-KNAW Fungal Biodiversity Centre,
Genetic diversity of 100+ Aspergillus species: The aspMine analysis resource

General information
State: 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, Joint Genome Institute, Utrecht University
Publication date: 2017
Peer-reviewed: Yes
Event: Poster session presented at Comparative genomics of eukaryotic microbes: Dissecting sources of evolutionary diversity.
Electronic versions: poster_asilomar.pdf
Research output: Research - peer-review > Poster – Annual report year: 2018

Genus level analysis of secondary metabolism reveals the origin of Aspergillus hybrid NRPS-PKS gene clusters

General information
State: Published
Organisations: Department of Biotechnology and Biomedicine, Network Engineering of Eukaryotic Cell factories, New Bioactive Compounds, Fungal Chemodiversity, Natural Product Discovery, Eukaryotic Molecular Cell Biology, Joint Genome Institute
Number of pages: 1
Publication date: 2017
Peer-reviewed: Yes
Event: Abstract from 14th International Aspergillus Meeting, Pacific Grove, United States.
Electronic versions: abstract_asperfest_2017.pdf
Research output: Research - peer-review > Conference abstract for conference – Annual report year: 2017

Identifying more than 300 Biosynthetic Gene Clusters with Potential Resistance Genes in over 75 Aspergillus species using Resistance Gene-Guided Genome Mining

General information
State: Published
Organisations: Department of Biotechnology and Biomedicine, Network Engineering of Eukaryotic Cell factories, Fungal Chemodiversity, New Bioactive Compounds, Natural Product Discovery, Eukaryotic Molecular Cell Biology, Joint Genome Institute
Number of pages: 1
Publication date: 2017
Peer-reviewed: Yes
Event: Poster session presented at 29th Fungal Genetics Conference, Pacific Grove, United States.
Electronic versions: 29FGC_abstract3.pdf
What drives speciation? Examination into the evolutionary events of more than 100 Aspergillus species

General information
State: Published
Organisations: Department of Biotechnology and Biomedicine, Network Engineering of Eukaryotic Cell factories, Novo Nordisk Foundation Center for Biosustainability, New Bioactive Compounds, Fungal Chemodiversity, Natural Product Discovery, Joint Genome Institute, Joint Bioenergy Institute
Number of pages: 1
Publication date: 2017
Peer-reviewed: Yes
Event: Abstract from 29th Fungal Genetics Conference, Pacific Grove, United States.
Electronic versions:
What_drives_speciation_Examination_into_the_evolutionary_events_of_more_than_100_Aspergillus_species.pdf

Approaches for Comparative Genomics in Aspergillus and Penicillium
The number of available genomes in the closely related fungal genera Aspergillus and Penicillium is rapidly increasing. At the time of writing, the genomes of 62 species are available, and an even higher number is being prepared. Fungal comparative genomics is thus becoming steadily more powerful and applicable for many types of studies. In this chapter, we provide an overview of the state-of-the-art of comparative genomics in these fungi, along with recommended methods. The chapter describes databases for fungal comparative genomics. Based on experience, we suggest strategies for multiple types of comparative genomics, ranging from analysis of single genes, over gene clusters and CaZymes to genome-scale comparative genomics. Furthermore, we have examined published comparative genomics papers to summarize the preferred bioinformatic methods and parameters for a given type of analysis, highly useful for new fungal geneticists. Moreover, the chapter contains a detailed overview of comparative genomics studies of key fungal traits such as primary metabolism, secondary metabolism, and secretome analysis. Finally, we gaze into a possible future of the field by comparing the current state of fungal comparative genomics to the development in bacterial genomics, where the comparison of hundreds of genomes has been performed for a while.

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories
Contributors: Rasmussen, J. L. N., Theobald, S., Brandl, J., Vesth, T. C., Andersen, M. R.
Publication date: 2016

Host publication information
Title of host publication: Aspergillus and Penicillium in the Post-genomic Era
Publisher: Caister Academic Press
Editors: de Vries, R. P., Benoît Gelber, I., Rørdam Andersen, M.
ISBN (Electronic): 978-1-910190-40-1
Source: PublicationPreSubmission
Source-ID: 123997713
Research output: Research - peer-review › Conference abstract for conference – Annual report year: 2017

aspMine - online comparative analysis of species from the Aspergillus genus

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Joint Genome Institute, Utrecht University
Contributors: Vesth, T. C., Rasmussen, J. L. N., Theobald, S., de Vries, R. P., Grigoriev, I. V., Baker, S. E., Andersen, M. R.
Number of pages: 1
Pages: 71-71
Publication date: 2016

Host publication information
Co-evolution of secondary metabolite gene clusters and their host

Secondary metabolite gene cluster evolution is mainly driven by two events: gene duplication and annexation and horizontal gene transfer. Here we use comparative genomics of Aspergillus species to investigate the evolution of secondary metabolite (SM) gene clusters across a wide spectrum of species. We investigate the dynamic evolutionary relationship between the cluster and the host by examining the genes within the cluster and the number of homologous genes found within the host and in closely related species.

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Eucaryotic Molecular Cell Biology
Number of pages: 1
Publication date: 2016
Peer-reviewed: Yes
Electronic versions:
Comparative_genomics_of_four_Aspergillus_species_with_focus_on_identification_of_specific_secondary_metabolite_gene_clusters.pdf
Research output: Research - peer-review » Conference abstract for conference – Annual report year: 2016

Comparative genomics of four Aspergillus species with focus on identification of specific secondary metabolite gene clusters

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Joint Genome Institute, Joint Bioenergy Institute
Number of pages: 1
Pages: 418-418
Publication date: 2016

Host publication information
Title of host publication: Book of abstracts from the 13th European Conference on Fungal Genetics
Article number: CSST31
Electronic versions:
Comparative_genomics_of_four_Aspergillus_species_with_focus_on_identification_of_specific_secondary_metabolite_gene_clusters.pdf
Research output: Research - peer-review » Conference abstract in proceedings – Annual report year: 2016

FunGeneClusterS: Predicting fungal gene clusters from genome and transcriptome data

Secondary metabolites of fungi are receiving an increasing amount of interest due to their prolific bioactivities and the fact that fungal biosynthesis of secondary metabolites often occurs from co-regulated and co-located gene clusters. This makes the gene clusters attractive for synthetic biology and industrial biotechnology applications. We have previously published a method for accurate prediction of clusters from genome and transcriptome data, which could also suggest cross-chemistry, however, this method was limited both in the number of parameters which could be adjusted as well as in user-friendliness. Furthermore, sensitivity to the transcriptome data required manual curation of the predictions. In the present work, we have aimed at improving these features.

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories
Contributors: Vesth, T. C., Brandl, J., Andersen, M. R.
Number of pages: 8
Pages: 122-129
Ortholog identification in genera of high genetic diversity and evolution

In the era of high-throughput sequencing, comparative genomics is vastly used in the discovery of genetic diversity between species, but also in defining the core and pan genome of single species to whole genera. Current comparative approaches are implementing ortholog identification to establish genome annotations, gene or protein evolutions or defining functional features in individual species and groups.

Ortholog prediction of the Aspergillus genus applicable for synthetic biology

The Aspergillus genus contains leading industrial microorganisms, excelling in producing bioactive compounds and enzymes. Using synthetic biology and bioinformatics, we aim to re-engineer these organisms for applications within
human health, pharmaceuticals, environmental engineering, and food production. In this project, we compare the genomes of ~300 species from the Aspergillus genus to generate a high-resolution pan-genomic map, representing genetic diversity spanning ~200 million years. We are identifying genes specific to species and clades to allow for guilt-by-association-based mapping of genotype-to-phenotype. To achieve this, we have developed orthologous protein prediction software that utilizes genus-wide genetic diversity. The approach is optimized for large data sets, based on BLASTp considering protein identity and alignment coverage, and clustering using single linkage of bi-directional hits. The result is orthologous protein families describing the genomic and functional features of individual species, clades and the core/pan genome of Aspergillus; and applicable to genotype-to-phenotype analyses in other microbial genera.

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Joint Bioenergy Institute, Joint Genome Institute
Contributors: Rasmussen, J. L. N., Vesth, T. C., Theobald, S., Frisvad, J. C., Grigoriev, I. V., Baker, S. E., Andersen, M. R.
Number of pages: 1
Publication date: 2016
Peer-reviewed: Yes
Source: PublicationPreSubmission
Source-ID: 127371130
Research output: Research - peer-review › Conference abstract for conference – Annual report year: 2016

Phylogenomic analysis of secondary metabolism genes sheds light on their evolution in Aspergilli
The World Health Organization is reporting a rising number of multiple drug resistant pathogens every year, increasing the need for new drug development. However, current methods for natural product discovery rely on time consuming experimental work, making them unable to keep up with this demand. In the aspMine project, we are sequencing and analyzing over 300 species of Aspergilli, a group of filamentous fungi rich in natural compounds. The vast amount of data obtained from these species challenges the way we were mining for products and requires new pipelines for secondary metabolite analysis. Natural products are encoded by genes located in close proximity, called secondary metabolic gene clusters, which makes them interesting targets for genomic analysis. We use a modified version of the Secondary Metabolite Unique Regions Finder (SMURF) algorithm, combined with InterPro annotations to create approximate maximum likelihood trees of conserved domains from secondary metabolic genes across 56 species, giving insights into the secondary metabolism gene diversity and evolution. In this study we can describe the evolution of nonribosomal peptide synthetases (NRPS), polyketide synthases (PKS) and hybrids thereof, find possible common ancestors and detect horizontal gene transfer events. Finally, we have performed large scale analysis of gene cluster dynamics and evolution, which provides us with better understanding of speciation in Aspergilli. With this new insights into the evolution of natural products, an application in synthetic natural product assembly lies within our grasp.

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Metabolomics Platform, Eukaryotic Molecular Cell Biology, Pacific Northwest National Laboratory
Number of pages: 1
Publication date: 2016
Peer-reviewed: Yes

Bibliographical note
Poster presentation
Source: PublicationPreSubmission
Source-ID: 127424624
Research output: Research - peer-review › Poster – Annual report year: 2016

Speciation over 200 million years – What makes an Aspergillus species

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Joint Genome Institute
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.

Diversity of carbohydrate metabolism in species of A spergillus

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.
Gene cluster dynamics throughout the Aspergillus genus

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Metabolomics Platform, Eucaryotic Molecular Cell Biology, Joint Genome Institute, Joint Bioenergy 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: Research - peer-review › Conference abstract for conference – Annual report year: 2015

Genome mining of the genetic diversity in the Aspergillus genus - from a collection of more than 30 Aspergillus species

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 show that ~48% of the annotated genes are core genes (genes shared between all species) and 24% of the genes are defining the individual species. The methods presented here will allow for a detailed investigation into mapping of genotypetophenotype across a very large set of genomic sequences.

General information
State: Published
Organisations: Department of Systems Biology, Network Engineering of Eukaryotic Cell Factories, Fungal Chemodiversity, Metabolomics Platform, Eucaryotic Molecular Cell Biology, Joint Genome Institute, Joint Bioenergy Institute, CBS-KNAW Fungal Biodiversity Centre, Joint Bioenergy Institute
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: 12755462
Research output: Research - peer-review › Conference abstract for conference – Annual report year: 2016

Determining and comparing protein function in Bacterial genome sequences

In November 2013, there was around 21,000 different prokaryotic genomes sequenced and publicly available, and the number is growing daily with another 20,000 or more genomes expected to be sequenced and deposited by the end of
2014. An important part of the analysis of this data is the functional annotation of genes – the descriptions assigned to genes that describe the likely function of the encoded proteins. This process is limited by several factors, including the definition of a function which can be more or less specific as well as how many genes can actually be assigned a function based on known functions.

This thesis describes the development of new tools for comparative functional annotation and a system for comparative genomics in general. As novel sequenced genomes are becoming more readily available, there is a need for standard analysis tools. The system CMG-biotools is presented here as an example of such a system and was used to analyze a set of genomes from the Negativicutes class, a group of bacteria closely related to Gram positives but which has a different cell wall structure and stains Gram negative, as the name indicates. The results of this work show that genomes of this class have very little homology to other known genomes making functional annotation based on sequence similarity very difficult.

Inspired in part by this analysis, an approach for comparative functional annotation was created based public sequenced genomes, CMGfunc. Functionally related groups of proteins were clustered based on sequence domains so that each group represented a protein function. Each function was then modeled using Artificial Neural Networks (ANN) and the model was evaluated based on its ability to identify true positives and negatives, that is proteins with or without the function of the model. The models were used to annotate a number of proteins without functional annotations and predicted functions for 98% of the genes. Evaluation of the precision of the method was performed, using data from the Critical Assessment of Functional Annotation (CAFA) project, and correct predictions were made in about 60% of the cases.

This project has highlighted the difficulties and challenges in functional annotation and computational analysis of sequence data. It has provided possible solutions for creating reproducible pipelines for comparative genomics as well as constructed a number of functional models not based on sequence similarity. Although much work is still left to be done, resources are flowing into the area of sequence analysis and progress is being made every day. As such, many different approach are being tried out and tested which will, in time, improve the knowledge gained from sequencing genomes.

**General information**
- **State**: Published
- **Organisations**: Department of Systems Biology
- **Contributors**: Vesth, T. C., Ussery, D., Lagesen, K.
- **Number of pages**: 101
- **Publication date**: 2014

**Publication information**
- **Publisher**: Technical University of Denmark (DTU)
- **Original language**: English
- **Electronic versions**: PhDThesis.PDF

**Amino Acid Usage Is Asymmetrically Biased in AT- and GC-Rich Microbial Genomes.**

Introduction: Genomic base composition ranges from less than 25% AT to more than 85% AT in prokaryotes. Since only a small fraction of prokaryotic genomes is not protein coding even a minor change in genomic base composition will induce profound protein changes. We examined how amino acid and codon frequencies were distributed in over 2000 microbial genomes and how these distributions were affected by base compositional changes. In addition, we wanted to know how genome-wide amino acid usage was biased in the different genomes and how changes to base composition and mutations affected this bias. To carry this out, we used a Generalized Additive Mixed-effects Model (GAMM) to explore non-linear associations and strong data dependences in closely related microbes; principal component analysis (PCA) was used to examine genomic amino acid and codon frequencies, while the concept of relative entropy was used to analyze genomic mutation rates. Results: We found that genomic amino acid frequencies carried a stronger phylogenetic signal than codon frequencies, but that this signal was weak compared to that of genomic %AT. Further, in contrast to codon usage bias (CUB), amino acid usage bias (AAUB) was differently distributed in AT- and GC-rich genomes in the sense that AT-rich genomes did not prefer specific amino acids over others to the same extent as GC-rich genomes. AAUB was also associated with relative entropy: genomes with low AAUB contained more random mutations as a consequence of relaxed purifying selection than genomes with higher AAUB. Conclusion: Genomic base composition has a substantial effect on both amino acid- and codon frequencies in bacterial genomes. While phylogeny influenced amino acid usage more in GC-rich genomes, AT-content was driving amino acid usage in AT-rich genomes. We found the GAMM model to be an excellent tool to analyze the genomic data used in this study.

**General information**
- **State**: Published
- **Organisations**: Department of Systems Biology, Center for Biological Sequence Analysis, Norwegian School of Veterinary Science
- **Contributors**: Bohlin, J., Brynildsrud, O. B., Vesth, T. C., Skjerve, E., Ussery, D.
- **Publication date**: 2013
- **Peer-reviewed**: Yes
CMG-Biotools, a Free Workbench for Basic Comparative Microbial Genomics.
This paper shows the strength and diverse use of the CMG-biotools system. The system can be installed on a wide range of host operating systems and utilizes as much of the host computer as desired. It allows the user to compare multiple genomes, from various sources using standardized data formats and intuitive visualizations of results. The examples presented here clearly shows that users with limited computational experience can perform complicated analysis without much training.

General information
State: Published
Organisations: Department of Systems Biology, Center for Biological Sequence Analysis, Department of Electrical Engineering, Electromagnetic Systems
Contributors: Vesth, T. C., Lagesen, K., Acar, Ö., Ussery, D.
Pages: e60120
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: P L o S One
Volume: 8
Issue number: 4
ISSN (Print): 1932-6203
Ratings:
BFI (2019): BFI-level 1
Web of Science (2019): Indexed yes
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 3.01 SJR 1.164 SNIP 1.111
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 3.11 SJR 1.236 SNIP 1.101
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 3.32 SJR 1.427 SNIP 1.136
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 3.54 SJR 1.559 SNIP 1.148
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 3.94 SJR 1.772 SNIP 1.153
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): CiteScore 4.15 SJR 1.982 SNIP 1.156
Web of Science (2012): Impact factor 3.73
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): CiteScore 4.58 SJR 2.425 SNIP 1.233
Web of Science (2011): Impact factor 4.092
ISI indexed (2011): ISI indexed no
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 1
Veillonella, Firmicutes: Microbes disguised as Gram negatives

The Firmicutes represent a major component of the intestinal microflora. The intestinal Firmicutes are a large, diverse group of organisms, many of which are poorly characterized due to their anaerobic growth requirements. Although most Firmicutes are Gram positive, members of the class Negativicutes, including the genus Veillonella, stain Gram negative. Veillonella are among the most abundant organisms of the oral and intestinal microflora of animals and humans, in spite of being strict anaerobes. In this work, the genomes of 24 Negativicutes, including eight Veillonella spp., are compared to 20 other Firmicutes genomes; a further 101 prokaryotic genomes were included, covering 26 phyla. Thus a total of 145 prokaryotic genomes were analyzed by various methods to investigate the apparent conflict of the Veillonella Gram stain and their taxonomic position within the Firmicutes. Comparison of the genome sequences confirms that the Negativicutes are distantly related to Clostridium spp., based on 16S rRNA, complete genomic DNA sequences, and a consensus tree based on conserved proteins. The genus Veillonella is relatively homogeneous: inter-genus pairwise comparison identifies at least 1,350 shared proteins, although less than half of these are found in any given Clostridium genome. Only 27 proteins are found conserved in all analyzed prokaryote genomes. Veillonella has distinct metabolic properties, and significant similarities to genomes of Proteobacteria are not detected, with the exception of a shared LPS biosynthesis pathway. The clade within the class Negativicutes to which the genus Veillonella belongs exhibits unique properties, most of which are in common with Gram-positives and some with Gram negatives. They are only distantly related to Clostridia, but are even less closely related to Gram-negative species. Though the Negativicutes stain Gram-negative and possess two membranes, the genome and proteome analysis presented here confirm their place within the (mainly) Gram positive phylum of the Firmicutes. Further studies are required to unveil the evolutionary history of the Veillonella and other Negativicutes.

General information
State: Published
Organisations: Center for Biological Sequence Analysis, Comparative Microbial Genomics, Department of Systems Biology, National Food Institute, Division of Food Microbiology, Division of Epidemiology and Microbial Genomics
Contributors: Vesth, T. C., Ozen, A., Andersen, S. C., Kaas, R. S., Lukjancenko, O., Bohlin, J., Nookaew, I., Wassenaar, T. M., Ussery, D.
Pages: 431-448
Publication date: 2013
Peer-reviewed: Yes

Publication information
Journal: Standards in Genomic Sciences
Volume: 9
Issue number: 2
ISSN (Print): 1944-3277
Ratings:
Web of Science (2019): Indexed yes
Web of Science (2018): Indexed yes
Scopus rating (2017): CiteScore 1.69 SJR 0.768 SNIP 0.629
Web of Science (2017): Impact factor 1.6
Web of Science (2017): Indexed yes
Scopus rating (2016): CiteScore 1.26 SJR 0.626 SNIP 0.511
Bayesian prediction of bacterial growth temperature range based on genome sequences

Background: The preferred habitat of a given bacterium can provide a hint of which types of enzymes of potential industrial interest it might produce. These might include enzymes that are stable and active at very high or very low temperatures. Being able to accurately predict this based on a genomic sequence, would thus allow for an efficient and targeted search for production organisms, reducing the need for culturing experiments. Results: This study found a total of 40 protein families useful for distinction between three thermophilicity classes (thermophiles, mesophiles and psychrophiles). The predictive performance of these protein families were compared to those of 87 basic sequence features (relative use of amino acids and codons, genomic and 16S rDNA AT content and genome size). When using naïve Bayesian inference, it was possible to correctly predict the optimal temperature range with a Matthews correlation coefficient of up to 0.68. The best predictive performance was always achieved by including protein families as well as structural features, compared to either of these alone. A dedicated computer program was created to perform these predictions. Conclusions: This study shows that protein families associated with specific thermophilicity classes can provide effective input data for thermophilicity prediction, and that the naïve Bayesian approach is effective for such a task. The program created for this study is able to efficiently distinguish between thermophilic, mesophilic and psychrophilic adapted bacterial genomes.
From Genome Sequence to Taxonomy - A Skeptic’s View

The relative ease of sequencing bacterial genomes has resulted in thousands of sequenced bacterial genomes available in the public databases. This same technology now allows for using the entire genome sequence as an identifier for an organism. There are many methods available which attempt to use genome sequences to classify bacteria, and the method of choice, as always, depends on the question asked and the particular need. For example, 16S rRNA can define a bacterial species, and relate species, genera, and higher orders into groups consistent with their known biological properties. However, distinguishing between strains of the same species requires additional information. The advantage of having the whole-genome sequence is that roughly a 1,000 times as much information is available, and this information can be used for rapid classification of strains, based on DNA sequence. This chapter reviews many commonly used methods and also describes potential pitfalls if used inappropriately, as well as which questions are best addressed by particular methods. After a brief introduction to the classical methods of taxonomy, a description of the bacterial genomes currently available is given, and then whole-genome-based methods are explored using three different data sets.

General information
State: Published
Organisations: Department of Systems Biology, Center for Biological Sequence Analysis, Novo Nordisk Foundation Center for Biosustainability, CFB - Metagenomic Systems Biology
Contributors: Özen, A. I., Vesth, T. C., Ussery, D.
Number of pages: 15
Publication date: 2012

Host publication information
Title of host publication: The Prokaryotes
Place of publication: 2012
Publisher: Springer
Editor: E., R.
DOIs: 10.1007/978-3-642-30194-0_11
Source: dtu
Source-ID: u::5474
Research output: Research - peer-review › Book chapter – Annual report year: 2012

On the Origins of a Vibrio species

Thirty-two genome sequences of various Vibrionaceae members are compared, with emphasis on what makes V. cholerae unique. As few as 1,000 gene families are conserved across all the Vibrionaceae genomes analysed; this fraction roughly doubles for gene families conserved within the species V. cholerae. Of these, approximately 200 gene families that cluster on various locations of the genome are not found in other sequenced Vibrionaceae; these are possibly unique to the V. cholerae species. By comparing gene family content of the analysed genomes, the relatedness to a particular species is identified for two unspeciated genomes. Conversely, two genomes presumably belonging to the same species have suspiciously dissimilar gene family content. We are able to identify a number of genes that are conserved in, and unique to, V. cholerae. Some of these genes may be crucial to the niche adaptation of this species.

General information
State: Published
Genomic taxonomy of vibrios

BACKGROUND: Vibrio taxonomy has been based on a polyphasic approach. In this study, we retrieve useful taxonomic information (i.e. data that can be used to distinguish different taxonomic levels, such as species and genera) from 32 genome sequences of different vibrio species. We use a variety of tools to explore the taxonomic relationship between the sequenced genomes, including Multilocus Sequence Analysis (MLSA), supertrees, Average Amino Acid Identity (AAI), genomic signatures, and Genome BLAST atlases. Our aim is to analyse the usefulness of these tools for species identification in vibrios. RESULTS: We have generated four new genome sequences of three Vibrio species, i.e., V. alginolyticus 40B, V. harveyi-like 1DA3, and V. mimicus strains VM573 and VM603, and present a broad analyses of these genomes along with other sequenced Vibrio species. The genome atlas and pangenome plots provide a tantalizing image of the genomic differences that occur between closely related sister species, e.g. V. cholerae and V. mimicus. The vibrio pangenome contains around 26504 genes. The V. cholerae core genome and pangenome consist of 1520 and 6923 genes, respectively. Pangenomes might allow different strains of V. cholerae to occupy different niches. MLSA and supertree analyses resulted in a similar phylogenetic picture, with a clear distinction of four groups (Vibrio core group, V. cholerae-V. mimicus, Aliivibrio spp., and Photobacterium spp.). A Vibrio species is defined as a group of strains that share > 95% DNA identity in MLSA and supertree analysis, > 96% AAI, <or = 10 genome signature dissimilarity, and > 61% proteome identity. Strains of the same species and species of the same genus will form monophyletic groups on the basis of MLSA and supertree. CONCLUSION: The combination of different analytical and bioinformatics tools will enable the most accurate species identification through genomic computational analysis. This endeavour will culminate in the birth of the online genomic taxonomy whereby researchers and end-users of taxonomy will be able to identify their isolates through a web-based server. This novel approach to microbial systematics will result in a tremendous advance concerning biodiversity discovery, description, and understanding.
Genome mining of species of Aspergillus and Penicillium for elucidation of the diversity and potential of carbohydrate degrading enzymes
Yang, T., PhD Student, Department of Biotechnology and Biomedicine
Andersen, M. R., Main Supervisor, Department of Biotechnology and Biomedicine
Vesth, T. C., Supervisor, Department of Biotechnology and Biomedicine
Liu, X., Supervisor
01/12/2018 → 30/11/2021
Project: PhD

Genome reduction of a filamentous fungus
Rothschild-Mancinelli, K., PhD Student, Department of Biotechnology and Biomedicine
Andersen, M. R., Main Supervisor, Department of Biotechnology and Biomedicine
Mortensen, U. H., Supervisor, Department of Biotechnology and Biomedicine
Vesth, T. C., Supervisor, Department of Biotechnology and Biomedicine
01/02/2018 → 31/01/2021
Award relations: Genome reduction of a filamentous fungus
Project: PhD

Genus-scale analysis of genre cluster evolution in fungi
Kjærbølling, I., PhD Student, Department of Biotechnology and Biomedicine
Andersen, M. R., Main Supervisor, Department of Biotechnology and Biomedicine
Mortensen, U. H., Supervisor, Department of Biotechnology and Biomedicine
Vesth, T. C., Supervisor, Department of Biotechnology and Biomedicine
Larsen, T. O., Supervisor, Department of Biotechnology and Biomedicine
Frandsen, R. J. N., Examiner, Department of Biotechnology and Biomedicine
Hallin, P. F., Examiner, Others
Wang, C. C. C., Examiner
Institut stipendie (DTU)
01/09/2015 → 14/01/2019
Award relations: Genus-scale analysis of genre cluster evolution in fungi
Project: PhD

Genus-level studies of gene dynamics for the Aspergillus genus
Theobald, S., PhD Student, Novo Nordisk Foundation Center for Biosustainability
Andersen, M. R., Main Supervisor, Department of Biotechnology and Biomedicine
Larsen, T. O., Supervisor, Department of Biotechnology and Biomedicine
Pedersen, A. G., Supervisor, Department of Health Technology
Vesth, T. C., Supervisor, Department of Biotechnology and Biomedicine
Weber, T., Examiner, Novo Nordisk Foundation Center for Biosustainability
Klitgaard, D. M. K., Examiner
Lebrun, M., Examiner
Lebrun, D. M. K., Examiner
Lebrun, M., Examiner
Samfinansierede - Virksomhed
01/12/2014 → 20/06/2018
Award relations: Genus-level studies of gene dynamics for the Aspergillus genus
Project: PhD

Genus-level studies of genome dynamics for the Aspergillus genus
Rasmussen, J. L. N., PhD Student, Department of Biotechnology and Biomedicine
Andersen, M. R., Main Supervisor, Department of Biotechnology and Biomedicine
Pedersen, A. G., Supervisor, Department of Health Technology
Vesth, T. C., Supervisor, Department of Biotechnology and Biomedicine
Lund, O., Examiner, National Food Institute
Lübeck, M., Examiner, Department of Systems Biology
Slot, J. C., Examiner
Slot, J. C., Examiner
Samfinansierede - Virksomhed
01/11/2014 → 20/06/2018
Award relations: Genus-level studies of genome dynamics for the Aspergillus genus
### Novel methods for assignment of protein function

**Vesth, T. C., PhD Student, Department of Biotechnology and Biomedicine**

**Ussery, D., Main Supervisor, Department of Systems Biology**

**Lagesen, K., Supervisor, Department of Systems Biology**

**Lund, O., Examiner, Department of Biotechnology and Biomedicine**

**Tolstrup, N., Examiner, Department of Chemistry**

**Meyer, F., Examiner**

Institut stipendie (DTU)

01/02/2011 → 02/04/2014

Award relations: Novel methods for assignment of protein function

Project: PhD

### Activities:

**17th European Conference on Computational Biology**

*Period: 8 Sep 2018 → 12 Sep 2018*

Tammi Camilla Vesth (Organizer)

Department of Biotechnology and Biomedicine

Section for Synthetic Biology

Network Engineering of Eukaryotic Cell factories

**Description**

Immense diversity found in secondary metabolite gene clusters in filamentous fungi and bacteria using comparative genomics

Tammi Vesth [1], Jens Frisvad [1], Ákos Kovács [1], Lars Jelsbak [1], Tilmann Weber [2], Scott E. Baker [3], Mikael R. Andersen [1]

1) Department of Bioengineering, Technical University of Denmark, Lyngby, Denmark

2) Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark

3) Joint Bioenergy Institute, Berkeley, CA, USA

Secondary metabolism in microorganisms is defined as non-life essential metabolism such as the production of mycotoxins and antibiotic compounds. This definition has to some extent been obscure as it is hard to define what is essential for the life of a microorganism in natural environments. Who is to say that the secondary metabolite penicillin produced by species of Penicillium is not essential for the life of the fungi. The compounds produced in these reactions are possible antimicrobial, biofuels, food spoilers and other important or valuable compounds. Secondary metabolism is, therefore, the subject of great curiosity for both biological and financial reasons. The number of known secondary metabolite compounds far exceed the number of characterized genes involved in these pathways and the effort required to elucidate the connection between metabolites and genes require time-consuming and expensive chemical characterization. Because of the benefits and challenges in this elucidation much effort has been given to the computational prediction of genes involved in secondary metabolism, with the most prominent software being AntiSMASH (bacteria and fungi) and SMURF (fungi). These methods provide valuable knowledge in the selection of gene targets for further analysis and can increase the speed of gene to compound association tremendously.

Here we present the prediction of secondary metabolism genes clusters and following comparative genomics analysis of species of Penicillium, Aspergillus, Bacillus, and Pseudomonas. We have predicted secondary metabolism in species 20 species from each of these 4 different groups (80 species) using antiSMASH and fungiSMASH and created families of clusters believed to produce similar compounds. The method relies partly on homology of individual genes but also on the analysis shows a tremendous diversity of clusters in all fours groups of species, but also shows that the fungi species have a much higher diversity than the bacteria compared to the general diversity. It is also seen that the number of predicted clusters keeps growing with new species illustrating the immense natural diversity and potential of these compounds. In the construction of cluster families, we test the effect of parameters on the families for bacterial and fungal data. The analysis method presented here illustrates how the prediction of secondary metabolite genes can be used for bacteria and fungi and shows how the methods must be adjusted to the type of microorganism. It also illustrates the vast diversity and potential in bioinformatics in the field of secondary metabolism and the further association of genes to compounds.

Degree of recognition: International

Links:
http://eccb18.org/

**Related event**

**17th European Conference on Computational Biology**  
08/09/2018 → 12/09/2018  
Athens, Greece  
Activity: Attending an event › Participating in or organising a conference

**Annual Danish Bioinformatics Conference 2018, Elixir**  
Period: 29 Aug 2018 → 30 Aug 2018  
Tammi Camilla Vesth (Organizer)  
Department of Biotechnology and Biomedicine  
Section for Synthetic Biology  
Network Engineering of Eukaryotic Cell factories  
Degree of recognition: National  
Links:  
http://elixir-node.cbs.dtu.dk/?page_id=2369

**Related event**

**Annual Danish Bioinformatics Conference 2018, Elixir**  
29/08/2018 → 30/08/2018  
Odense  
Activity: Attending an event › Participating in or organising a conference

**Wellcome Genome Campus Advanced Course: Fungal Pathogen Genomics, 13-18 May 2018**  
Period: 14 May 2018  
Tammi Camilla Vesth (Guest lecturer)  
Department of Biotechnology and Biomedicine  
Section for Synthetic Biology  
Network Engineering of Eukaryotic Cell factories

**Related external organisation**

**Wellcome Genome Campus**  
United Kingdom  
Activity: Talks and presentations › Guest lectures, external teaching and course activities at other universities

**Joint Genome Institute User Meeting 13**  
Period: 13 Mar 2018 → 16 Mar 2018  
Tammi Camilla Vesth (Organizer)  
Department of Biotechnology and Biomedicine  
Section for Synthetic Biology  
Network Engineering of Eukaryotic Cell factories

**Related event**

**Joint Genome Institute User Meeting 13: DOE JGI Genomics of Energy and Environment Meeting**  
13/03/2018 → 16/03/2018  
San Francisco, United States  
Activity: Attending an event › Participating in or organising a conference

**14th European Conference on Fungal Genetics**  
Period: 25 Feb 2018 → 28 Feb 2018  
Tammi Camilla Vesth (Organizer)
In this work, we present the new whole genome sequences, functional annotation and comparative analysis of 6 filamentous fungi of the Aspergillus genus. The new species belong to the sections of Sparsi (3), Ochraceorosei (1), Tanneri (1) and Rubustii (1).

Species of section Sparsi, A. funiculosus, A. implicatus and A. biplanus are found in warm soil climates and produce antimicrobial compounds and toxins such as kojic acid, auraglaucin, gregatins, funicin and sidins. The section Ochraceorosei (suggested 2009) consists of A. ochraceoroseus and A. rambelii. A. ochraceoroseus, these species produce the mycotoxin aflatoxin B1.

The analysis presented here include genome sequence quality, secondary metabolism potential, carbohydrate degradation potential, shared proteomes and species as well as section specific genes. Comparisons were made to other filamentous fungi, Penicillium (3), Neurospora (1) and Aspergillus (49 species, 32 from section Nigri). The species of Ochraceorosei have a much smaller number of predicted genes than the other species in the set (7,800-8,200). This is in comparison to some of the Nigri species with up to 18,000 genes. Section Sparsi species have a very wide range of predicted genes (9,000-15,000) while A. tanneri falls in the midrange (13,000). The large range of predicted genes illustrates the large diversity within these species.

Analyzing the CaZyme distribution of the 6 species revealed a diversity comparable to that of section Nigri. In the analysis of secondary metabolism, we find shared and conserved clusters within some sections while other sections have not associated clusters. Unique gene clusters are found in all the newly sequenced genomes, to the same extent as found in the Aspergilli in general.

The six new species provide additional information to the comparative genomics studies of Aspergillus and illustrate the large diversity and application of species in this genus.

Links:
http://www.ecfg14.org/

Related event

14th European Conference on Fungal Genetics
25/02/2018 → 28/02/2018
Haifa, Israel
Activity: Attending an event › Participating in or organising a conference
The filamentous fungal species of the Aspergillus genus are of broad interest to the scientific community including applied, medical and basic research. These fungi are prolific producers of native and heterologous proteins, organic acids, and secondary metabolites (including bioactives and toxins such as ochratoxin A). Because of these abilities, they represent a substantial economic interests in pharmaceutical, biotechnology, and bioenergy applications. In a project collaboration with the US Joint Genome Institute and JBEI we are de novo sequencing 300 different species of Aspergillus and establishing an online analysis platform for the scientific community, aspMine. The goal of this project is to develop a targeted tool to expand and improve our knowledge and expertise about this versatile group of fungi. At time of writing, 200 genomes are in various stages of sequencing and a bioinformatic pipeline has been established to analyze and store the data. This project covers a wide range of biologically interesting ideas surrounding the concept of speciation, such as genetic diversity, primary and secondary metabolism and proteome diversity. Complementary to the tools offered by FungiDB and JGI, the aspMine analysis resource offers tools for tracking genes and functions across species, allowing for investigation of shared genes and clusters across the genus as well as species- and clade-specific genes. The online platform also offers comparative analysis of secondary metabolism gene clusters with focus on synteny and functional conservation across species. The aspMine is implemented as a number of web applications created in R shiny, a graphical interface for analysis. The different tools are collected on a webpage which also includes method descriptions and relevant literature. The webpage is available from the beginning of 2016 and will be continually expanded. It is our goal to provide a comprehensive analysis platform for the community for comparative analysis of Aspergillus species.

Related event
Comparative genomics of eukaryotic microbes: Dissecting sources of evolutionary diversity
14/10/2017 → 19/10/2017
Activity: Attending an event › Participating in or organising a conference

Annual Danish Bioinformatics Conference 2017
Tammi Camilla Vesth (Organizer)
Department of Biotechnology and Biomedicine
Section for Synthetic Biology
Network Engineering of Eukaryotic Cell factories

Related event
Annual Danish Bioinformatics Conference 2017: Elixir
23/08/2017 → 25/08/2017
Odense, Denmark
Activity: Attending an event › Participating in or organising a conference

ETALEE 2017
Period: 23 May 2017 → 24 May 2017
Tammi Camilla Vesth (Organizer)
Department of Biotechnology and Biomedicine
Section for Synthetic Biology
Network Engineering of Eukaryotic Cell factories

Related event
ETALEE 2017: Exploring Teaching for Active Learning in Engineering Education 2017
23/05/2017 → 24/05/2017
Odense, Denmark
Activity: Attending an event › Participating in or organising a conference

Joint Genome Institute User Meeting 12
Period: 20 Mar 2017 → 22 Mar 2017
Tammi Camilla Vesth (Organizer)
Department of Biotechnology and Biomedicine
Section for Synthetic Biology
Network Engineering of Eukaryotic Cell factories

Related event

Joint Genome Institute User Meeting 12: DOE JGI Genomics of Energy and Environment Meeting
20/03/2017 → 23/03/2017
Walnut Creek, United States
Activity: Attending an event › Participating in or organising a conference

29th Fungal Genetics Conference
Period: 14 Mar 2017 → 19 Mar 2018
Tammi Camilla Vesth (Organizer)
Department of Biotechnology and Biomedicine
Section for Synthetic Biology
Network Engineering of Eukaryotic Cell factories

Description
Genetic diversity of 100+ Aspergillus species - the aspMine analysis resource.
T. C. Vesth1, J. L. Nybo (1), S. THEOBALD (1), R. P. DE VRIES (4), I. V. GRIGORIEV (3), S. E. BAKER2, M. R. ANDERSEN (1)

1) Department of Bioengineering, Technical University of Denmark, Lyngby, Denmark;
2) Joint Bioenergy Institute, Berkeley, CA, USA, Berkeley, CA, USA;
3) Joint Genome Institute, Walnut Creek, CA, USA, Walnut Creek, CA, USA;
4) Fungal Physiology, CBS-KNAW Fungal Biodiversity Centre, Utrecht, the Netherlands, Utrecht, The Netherlands.

The filamentous fungal species of the Aspergillus genus are of broad interest to the scientific community including applied, medical and basic research. These fungi are prolific producers of native and heterologous proteins, organic acids, and secondary metabolites (including bioactives and toxins such as ochratoxin A). Because of these abilities, they represent a substantial economic interests in pharmaceutical, biotechnology, and bioenergy applications. In a project collaboration with the US Joint Genome Institute and JBEI we are de novo sequencing 300 different species of Aspergillus and establishing an online analysis platform for the scientific community, aspMine. The goal of this project is to develop a targeted tool to expand and improve our knowledge and expertise about this versatile group of fungi. At time of writing, 200 genomes are in various stages of sequencing and a bioinformatic pipeline has been established to analyze and store the data. This project covers a wide range of biologically interesting ideas surrounding the concept of speciation, such as genetic diversity, primary and secondary metabolism and proteome diversity. Complementary to the tools offered by FungiDB and JGI, the aspMine analysis resource offers tools for tracking genes and functions across species, allowing for investigation of shared genes and clusters across the genus as well as species- and clade-specific genes. The online platform also offers comparative analysis of secondary metabolism gene clusters with focus on synteny and functional conservation across species. The aspMine is implemented as a number of web applications created in R shiny, a graphical interface for analysis. The different tools are collected on a webpage which also includes method descriptions and relevant literature. The webpage is available from the beginning of 2016 and will be continually expanded. It is our goal to provide a comprehensive analysis platform for the community for comparative analysis of Aspergillus species.

Related event

29th Fungal Genetics Conference
14/03/2017 → 19/03/2017
Pacific Grove, United States
Activity: Attending an event › Participating in or organising a conference

aspMine - online comparative analysis of species from the Aspergillus genus
Period: 5 Apr 2016
Tammi Camilla Vesth (Invited speaker)
Department of Systems Biology
Network Engineering of Eukaryotic Cell Factories

Description
The filamentous fungal species of the Aspergillus genus are of broad interest to the scientific community including applied, medical and basic research. These fungi are prolific producers of native and heterologous proteins, organic acids, and secondary metabolites (including bioactives and toxins such as ochratoxin A). Because of these abilities, they represent a substantial economic interests in pharmaceutical, biotechnology, and bioenergy applications. In a project collaboration with the US Joint Genome Institute and JBEI we are de novo sequencing 300 different species of Aspergillus and establishing an online analysis platform for the scientific community, aspMine. The goal of this project is to develop a targeted tool to expand and improve our knowledge and expertise about this versatile group of fungi. At time of writing, 200 genomes are in various stages of sequencing and a bioinformatic pipeline has been established to analyze and store the data. This project covers a wide range of biologically interesting ideas surrounding the concept of speciation, such as genetic diversity, primary and secondary metabolism and proteome diversity. Complementary to the tools offered by FungiDB and JGI, the aspMine analysis package offers tools for tracking genes and functions across species, allowing for investigation of shared genes and clusters across the genus as well as species- and clade-specific genes. The online platform also offers comparative analysis of secondary metabolism gene clusters with focus on synteny and functional conservation across species. The aspMine is implemented as a number of web applications created in R shiny, a graphical interface for analysis. The different tools are collected on a webpage which also includes method descriptions and relevant literature. The webpage is available from the beginning of 2016 and will be continually expanded. It is our goal to provide a comprehensive analysis platform for the community for comparative analysis of Aspergillus species.

Documents:
2016_ECFG_abstract

Related event

13th European Conference on Fungal Genetics: bridging fungal genetics, evolution and ecology
03/04/2016 → 06/04/2016
Paris, France
Activity: Talks and presentations › Conference presentations

Comparative genomics and gene cluster identification in 28 species of Aspergillus section Nigri.
Period: 2015
Tammi Camilla Vesth (Invited speaker)
Department of Systems Biology
Network Engineering of Eukaryotic Cell Factories

Description
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. While 8 individual species from this group has been whole- genome sequenced, the genetic basis for these diverse phenotypes remains largely unidentified.

In this study, we have de novo sequenced the genomes of 20 additional species of the section Nigri, thus allowing the genome comparison of all members of this important section of fungal species. Here we present the results of this large-scale genomic analysis where we have examined the core genome of these 28 species and identified variations in the genetic makeup of individual species and groups of species. In particular, we have found genes unique to Aspergillus section Nigri, as well as genes which are only found in subgroups of the section. Our analysis here correlates these genes to the phenotypes of the fungi.

Furthermore, we have predicted secondary metabolite gene clusters in all 28 species. We present here an overview of these gene clusters and how they are shared and vary between species. We also correlate the presence of gene clusters to presence of known fungal metabolites.

Documents:
2015_FGS_abstract

Related event

Fungal Genetics Confence
17/03/2015 → ...
Activity: Talks and presentations › Conference presentations