

Ortholog identification in genera of high genetic diversity and evolution

Rasmussen, Jane Lind Nybo; Vesth, Tammi Camilla; Frisvad, Jens Christian; Theobald, Sebastian; Kjærbølling, Inge; Andersen, Mikael Rørdam

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Rasmussen, J. L. N., Vesth, T. C., Frisvad, J. C., Theobald, S., Kjærbølling, I., & Andersen, M. R. (2016). Ortholog identification in genera of high genetic diversity and evolution. Abstract from 15th European Conference on Computational Biology, Den Haag, Netherlands.

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Ortholog identification in genera of high genetic diversity and evolution

Jane L. Nybo, Tammi Vesth, Jens C. Frisvad, Sebastian Theobald, Inge Kjærboelling and Mikael R. Andersen

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.

In this project we aim to compare the genomes of 300 species from the genus of the filamentous fungi - *Aspergillus*, which represent evolutionary changes over 200 million years [1]. Present-day's gene orthology prediction software combine genetic diversity with evolutionary distance. However, when the genetic variation within this genus is higher than the variation between fish and humans [2], identifying true orthologs becomes problematic. The cause of this high genetic diversity within such closely related species might be due to genetic events such as gene duplication and horizontal gene transfer. Here we tackle this problem by stringent the alignment parameters and loosen the linkage criteria. The approach is based on BLAST alignments with optimized cut-offs considering identity and coverage and clustering using single linkage.

The result of this procedure is a large number of orthologous protein families, that will allow for guilt-by-association-based detailed investigation into mapping of genotype to phenotype. The data will highlight the genomic and functional features defining the pan and core genome of individual strains, species, clades and the genus of *Aspergillus*.

[1] Galagan et al. (2005). *Nature*, 438:1105–1115.

[2] Fedorova et al. (2008). *PLoS Genetics*, 4(4), e1000046.