

Application of next-generation genome and transcriptome based methods for the exploration of secondary metabolites from marine fungi for the treatment of cancer

Abhishek Kumar and Frank Kempken

Department of Genetics & Molecular Biology in Botany, Institute of Botany,

Christian-Albrechts-University at Kiel, Kiel, Germany

Email: akumar@bot.uni-kiel.de | fkempken@bot.uni-kiel.de

Fungi of marine origin are potent groups of secondary metabolite producers. However, they are not well characterized and underutilised in terms of biotechnological applications. We aim for sustainable exploration of marine fungal isolates and their encoding natural products as drugs against cancer under the EU-funded project marine fungi (www.marinefungi.eu). Besides isolation of new fungal strains from unique marine habitats, the molecular development of effective producer strains is in the focus. Genomes of selected candidate strains originating from our unique strain collection of marine fungi are currently characterized with respect to secondary metabolite production.

Next-generation sequencing (NGS) techniques have changed the facets of genomics and its application. We have established the genomic sequences from three marine isolates, *Scopulariopsis brevicaulis*, *Pestalotiopsis* sp. and *Calcariosporium* sp. by the use of different next-generation sequencing methods (Roche 454, Illumina and ion-torrent).

We report on different properties of genome assemblies and annotations for these fungi. Several gene families and superfamilies have been analyzed to explore genetic peculiarities of these species along with repeats and transposable element contents. The assembled genome of *Scopulariopsis brevicaulis* is ~32 Mb in size with N50 equals to 88 kb and 935 contigs containing 16298 genes with average intron length equals to 129.4. During the annotation process, we were able to annotate 9340 genes (57.31 %) while 6958 genes (43.69 %) remained non-annotated in *Scopulariopsis brevicaulis* genome. 17 genes encoding for non-ribosomal peptide synthetases (NRPSs), 18 polyketide synthases (PKSs) and one gene encoding a hybrid NRPS-PKS were found. Similarly, the genome size for *Pestalotiopsis* sp. is ~46 Mb with N50 equals to 71.9 kb and 4186 contigs containing 23492 genes, which is surprisingly high for an ascomycete. The average intron length and the average intron per gene are 126.8 and 2.2, respectively. During annotation process, we annotated 60% genes of *Pestalotiopsis* genome with 44 NRPSs, 62 PKSs and 7 hybrid NRPS-PKS genes. The assembled genome size of *Calcariosporium* sp. is about 35 Mb genome with N50 equals to 91.9 kb and 2464 contigs containing 15459 genes. The percentage GC% for this genome is 50.7%. The average intron length and the average intron per gene are 121 and 2.1, respectively. During annotation process, we annotated 72% genes, while 28% genes remained non-annotated for *Calcariosporium* genome with 52 NRPSs, 66 PKSs and 7 hybrid NRPS-PKS genes.

Predicted genes are presently in process of validation using illumina based RNA-seq. We are also comparing wild type phenotypes with higher-yielding mutants of these fungi with special interest on specific natural compounds.