Polyketide and non-ribosomal peptide synthetases in Aspergillus carbonarius genome:A strategy for identification of secondary metabolite clusters

Gallo A[1], Baker SE[23], Mulè G[1], Susca A[1], Logrieco A[1], Perrone G[1]

Motivation

One of the major aims of fungal genomics is the identification of genes involved in the biosynthesis of secondary metabolites, which include important mycotoxins. These compounds are relevant for human and animal health, so understanding their mechanism of production is important in order to limit contamination on food and feed. On the other hand, there is a significant opportunity for the discovery of novel bioactive natural products which could be exploited for their beneficial applications, such as antibiotics. Genes encoding enzymes likely to be involved in natural product biosynthesis can be readily located in sequenced genomes by use of computational sequence comparison tools. Many fungal secondary metabolites are polyketides and non ribosomal peptides and in recent years the genome analysis of many filamentous fungi has revealed an unexpectedly large number of genes encoding for polyketide synthases (PKS) and non ribosomal peptide synthetases (NRPS). The identification of these genes could aid in the prediction of secondary metabolite biosynthetic gene clusters. Coupling of genome sequencing with transcriptional analyses and genetic manipulation accelerates the elucidation of the biosynthetic pathway and the regulatory mechanism of production. The recent sequencing and annotation of Aspergillus carbonarius genome, generated in collaboration with the US DOE-Joint Genome Institute (JGI), will enable a variety of bioinformatic analyses in this important organism, which has been reported to be the main agent of ochratoxin A (OTA) contamination of wine, grapes, grape juice and dried vine fruits. OTA, a widespread mycotoxin produced by several species of Aspergillus and Penicillium, is a potent nephrotoxin, and also displays hepatotoxic, teratogenic and immunosuppressive properties. It has been classified in group 2B (a possible human carcinogen) by the International Agency for Research on Cancer. The OTA biosynthetic pathway has not yet been completely elucidated. So far the majority of the studies have been focused on Penicillium species and A. ochraceus, and, combined with the molecular structure of the mycotoxin, point toward a PKS as the key enzyme catalyzing the first step of OTA biosynthesis.

[1] National Research Council, Institute of Science of Food Production (ISPA) Bari, Italy [2] Pacific Northwest National Laboratory, Richland, Washington, US [3] DOE Joint Genome Institute, Walnut Creek, California, US

Methods

For the A. carbonarius sequencing project both the 454 and the Sanger sequencing technologies were used. The annotation was performed on the basis of a consensus gene set predicted by the JGI annotation pipeline, using a variety of cDNA-based, protein-based, and ab initio gene modelers, and a filtering based on homology and EST support. PKSs and NRPSs constitute a class of multifunctional proteins that use a very similar strategy for the biosynthesis of two distinct classes of natural products. They present complex modular structures. The main domains of each module of NRPSs are the adenylation (A), the peptyl carrier (T) and the condensation (C) domains; whereas the ketoacyl synthase (KS), the acyltransferase (AT) and the acyl carrier protein (ACP) represent the main domains for PKSs. To retrieve pks and nrps genes the consensus sequence of the principal domains was used as queries in BLAST search of the A. carbonarius genome assembly. Homology analysis of retrieved pks genes was carried out for the identification of putative pks genes involved in the biosynthesis of OTA. Analysis of the annotated genes belonging to the putative OTA biosynthetic cluster was initiated with the attempt to clarify the biosynthetic pathway.

Results

The genome assembly sequence of A. carbonarius and the automated annotation is now open to the public (http://genome.jgi-psf.org/Aspca1), while a new annotated version is currently restricted only to registered users (http://genome.jgi-psf.org/Aspca3). The genome sequence was the result of a hybrid assembly of 963 scaffolds spanning a total of 36.3 Mbp. A total of 11624 genes, with an average length of 2241 bp, were structurally and functionally annotated. The search tools available at the genome portal allowed us to establish the presence in A. carbonarius of 33 NRPS and 28 PKS encoding genes, most of which are located in clusters. The analysis of their domain and modular structures confirmed the high diversity of the two enzymes due to the difference of roles they may have in the fungal metabolism. A pks gene of about 8000 bp and encoding a protein of about 2500 aa showed a high similarity to the OTA putative pks gene of A. niger, which was identified on the basis of homology to the A. ochraceus pks involved in the biosynthesis of OTA. The putative OTA pks of A. carbonarius belongs to a cluster which includes, according to the annotation data, some genes which may have a role in the OTA synthesis mechanism. Among them, genes encoding a NRPS, a cytochrome P450 type monoxygenase and a halogenase are present, which interestingly match with the prediction made on the basis of the OTA molecular structure and with the results so far obtained in other OTA producing fungal species.

Contact e-mail

antonia.gallo@ispa.cnr.it

Supplementary information
This research was partially supported by Italian Minister of Research (MIUR) project (MBLab,DM19410), Fondo FAR - Legge 297/1999 Art. 12/lab