De novo sequencing and detection of secondary metabolite gene clusters of Penicillium griseofulvum.

De Novo Sequencing and Detection of Secondary Metabolite Gene Clusters of *Penicillium Griseofulvum*

Houda Banani[1], Marina Marcet-Houben[2,3], Ana-Rosa Ballester[4], Pamela Abbruscato[5], Luis González-Candelas[4], Toni Gabaldón[2,3,6] and Davide Spadaro[1]

1) Dept. Agricultural, Forestry and Food Sciences (DISAFA), University of Torino, Grugliasco, Italy
2) Bioinformatics and Genomics Programme. Centre for Genomic Regulation (CRG), Barcelona, Spain
3) Universitat Pompeu Fabra (UPF), Barcelona, Spain
4) Instituto de Agroquímica y Tecnología de Alimentos (IATA-CSIC), Valencia, Spain
5) Rice Genomics Unit, Parco Tecnologico Padano, Lodi, Italy
6) Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Keywords: Patulin, griseofulvin, apple, food safety, mycotoxin, HPLC

Abstract

The genus *Penicillium* comprises different economically important species with the ability to produce a wide array of secondary metabolites. Among different *Penicillium* species, *P. griseofulvum* is worldwide distributed and has been associated with blue mold decay. In the present study, the complete genome of *P. griseofulvum* strain PG3 isolated from rotted apples harvested in Italy was sequenced and some important secondary metabolites clusters present in PG3 were reported. The PG3 estimated genome size was of 29.3 Mb, and the phylogenetic analysis at the whole-genome level revealed that *P. griseofulvum* is branched off after the divergence of *P. oxalicum* and before *P. chrysogenum.* Genome-wide analysis of PG3 genes uncovered a putative gene cluster for patulin biosynthesis. *In vitro* results clearly confirmed that PG3 is a high patulin producer. In addition to patulin, we detected a functional griseofulvin gene cluster.This study will enable to gain insight into secondary metabolite synthesis in *P. griseofulvum* and assess its potential applications in biotechnology and threats for food safety.

INTRODUCTION

P. griseofulvum Dierckx (syn. *P. patulum* Bain.; *P. urticae* Bain.) is worldwide distributed and it has been isolated from fruits, decaying plants, cereal grains and animal feed (Shim et al., 2006). *P. griseofulvum* has been associated with blue mold decay in storage apple fruits, which is considered as one of the most important postharvest diseases of pome fruits worldwide (Pianzzola et al., 2004).

Besides the economic losses, *P. griseofulvum* may represent a potential health risk because of its ability to produce mycotoxins such as patulin. The genes forming the patulin cluster were characterized in *P. expansum* and *Aspergillus clavatus* (Artigot et al., 2009; Tannous et al., 2014; Ballester et al., 2015; Li et al., 2015); however no information is yet available about the composition of the patulin cluster in *P. griseofulvum*. This information is needed to clearly understand the mechanisms leading to patulin production in this fungus and to define strategies for patulin control.

P. griseofulvum is also known to produce a wide array of important useful

secondary metabolites, including griseofulvin (Samson RA et al., 2004; Shim et al., 2006) which has been in use for many years in medical and veterinary applications (Finkelstein et al., 1996). The griseofulvin biosynthetic gene cluster consists of 13 putative genes and has been reported in *P. aethiopicum* (Chooi et al., 2010; Cacho et al., 2013, 2015), but it is still not known the genes forming the griseofulvin cluster in *P. griseofulvum*.

To gain insight into secondary metabolite clusters in *P. griseofulvum* and assess its biotechnological potential and define better the threats for food safety, we have sequenced for the first time its genome.

MATERIALS AND METHODS

Penicillium griseofulvum Dierckx (syn: *P. urticae* Bainier) strain PG3 (deposited at Centraalbureau voor Schimmelcultures, with Accession number CBS 140421) was obtained from rotten apples harvested in Piedmont, Northern Italy.

Total DNA was extracted from the strain PG3 as previously described by Ballester and collaborators (Ballester et al., 2015) and then DNA concentration and purity were checked by a spectrophotometer (Nanodrop 2000, Thermo Scientific, Wilmington, USA). The genome of *P. griseofulvum* PG3 was sequenced at the Genomics Platform of the Parco Tecnologico Padano using the Illumina MiSeq technology. SPAdes was used to assemble the *P. griseofulvum* genome (Bankevich et al., 2012) and the genes encoded in the genome were predicted by Augustus trained with *Aspergillus nidulans* (Keller et al., 2011). A phylome, designed as the complete collection of phylogenetic trees for each gene encoded in a genome, was reconstructed for *P. griseofulvum*. Fourteen other species were included in the phylome. These comprised the other sequenced *Penicillium* genomes (*P. chrysogenum, P. oxalicum, P. roqueforti, P. camemberti, P. expansum, P. digitatum* and *P. italicum*) and members of the *Aspergillus* and *Talaromyces* clade. Species tree was reconstructed using the method of gene concatenation. RaxML was used to reconstruct the species tree using the PROTGAMMALG model (Stamatakis et al., 2005). A collection of 114 secondary metabolism clusters were used to look for homologous clusters in the *P. griseofulvum* genome following the method described in Ballester and collaborators (Ballester et al., 2015). PG3 colony diameter (mm) was measured for up to 10 days of growth, and then patulin and griseofulvin production were analyzed by high-performance liquid chromatography (HPLC).

RESULTS AND DISCUSSION

Genome sequencing and comparative genomics

The genome of *P. griseofulvum* strain PG3 was sequenced. Table 1 shows the final statistics of the genome assembly, which is composed of 363 contigs, 14 of which were larger than 100 kb. The estimated genome size was of 29.3 Mb. Gene annotation showed that 9,631 proteins were encoded in the genome. We compared the genome of *P. griseofulvum* with the genomes of 14 other fully-sequenced *Penicillium* and *Aspergillus* species. To determine the phylogenetic position of *P. griseofulvum* in relation with other sequenced species, we reconstructed a species tree based on the concatenation of 2,134 genes that were found to be single copy in all considered species. Our results show that *P. griseofulvum* is branched off between *P. chrysogenum* and *P. oxalicum* (Figure 1A).

Genome-wide analysis of *P. griseofulvum* PG3 genes revealed two putative gene

clusters for patulin and griseofulvin biosynthesis

The presence of secondary metabolites clusters in PG3 was analyzed by searching for homologs of about 114 gene clusters present in the database. A patulin gene cluster is identified for the first time in *P. griseofulvum*, PG3 strain, containing 15 genes gathered together ordered similarly to the patulin cluster of *P. expansum* strain PEXP (Ballester et al., 2015) (Figure 1B). This result confirms that the changes in gene order observed between the cluster in *A. clavatus* and the cluster in *P. griseofulvum* and *P. expansum* happened before the two *Penicillium* species diverged.

Besides the patulin gene cluster, we found a griseofulvin gene cluster (Figure 1C). This gene cluster was originally described in *P. aethiopicum* and consists of 13 genes (Chooi et al., 2010).

When we compared the griseofulvin gene cluster of *P. aethiopicum* with the one found in PG3, we found that three genes (*gsfR2, gsfK* and *gsfH)* were not located within the PG3 griseofulvin gene cluster: *gsfR2* codes for a putative transcription factor, *gsfK*, encodes a putative NAD(P)-dependent oxidoreductase, and *gsfH* codes for a isochorismatase-like protein. However these proteins have homologs in a different position of the genome and therefore we could not discard the option that they are still playing a function in the synthesis of griseofulvin.

Patulin and griseofulvin production *in vitro* by PG3

Patulin and griseofulvin production by *P. griseofulvum* PG3 was quantified *in vitro* for up to 10 days. After 3 days of incubation *in vitro*, PG3 produces considerable amount of patulin (about 446.86 µg/plate), then increased significantly to reach 3498.71 µg/plate at day 10 (figure 2B). Interestingly, these concentrations are in the same range as the patulin production by *P. expansum* strain PEXP which has similar genes number and order as the putative PG3 patulin cluster (Ballester et al., 2015), although PG3 exhibited distinct differences in colony morphology and slower growth kinetics compared with PEXP (figure 2A/B).

As mentioned before, *P. griseofulvum* is known to produce thesecondary metabolite griseofulvin that has been used for many years in medical and veterinary applications.In the present work, griseofulvin production by PG3 was investigated, and the results clearly confirm that PG3 produced significant levels of griseofulvin *in vitro*, which increased over the time to reach at the day 10 about 215 µg/plate (Figure 2C).

The production of griseofulvin by the newly sequenced *P. griseofulvum* PG3 strain indicates that the lack of three genes within the gene cluster is apparently not affecting the synthesis of griseofulvin. However we do not know whether or not these genes are nevertheless involved in the synthesis of this compound.

CONCLUSIONS

In this study, we sequenced and annotated for the first time the genome of the postharvest pathogen *P. griseofulvum* PG3 isolated from rotted apples in Italy. Our data suggest that PG3 genome size is 29.3 Mb, and the phylogenetic analysis at the whole-genome level revealed that *P. griseofulvum* is branched off after the divergence of *P. oxalixum* and before *P. chrysogenum.*

Then our analyses uncovered two important secondary metabolite gene clusters in PG3: a griseofulvin cluster and a patulin gene clusterthat is similar to the cluster identified in *P. expansum*.

Finally, *in vitro* analysis conducted in the present study revealed that PG3

produces a considerable amount of patulin and griseofulvin.

Our data give insight into secondary metabolites synthesis in *P. griseofulvum* PG3 and assess its potentiality in term of threats for food safety, but also pave the way for its future biotechnological applications such as using PG3-knocked out of PKS of the patulin cluster for producing griseofulvin.

Literature cited

Artigot, M.P., Loiseau, N., Laffitte, J., Mas-Reguieg, L., Tadrist, S., Oswald, I.P. and Puel, O. 2009. Molecular cloning and functional characterization of two CYP619 cytochrome P450s involved in biosynthesis of patulin in *Aspergillus clavatus*. Microbiology 155:1738–1747.

Ballester, A., Marcet-Houben, M., Levin, E., Sela, N., Selma-Lázaro, C., Carmona, L., Wisniewski, M., Droby, S., González-Candelas, L. and Gabaldón, T. 2015. Genome , Transcriptome , and Functional Analyses of *Penicillium expansum* Provide New Insights Into Secondary Metabolism and Pathogenicity. Mol. Plant-Microbe Interact. 28:232–248.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A. V, Sirotkin, A. V, Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19:455–477.

Cacho, R.A., Chooi, Y.-H., Zhou, H. and Tang, Y. 2013. Complexity generation in fungal polyketide biosynthesis: a spirocycle-forming P450 in the concise pathway to the antifungal drug griseofulvin. ACS Chem. Biol. 8:2322–2230.

Cacho, R.A., Tang, Y. and Chooi, Y.-H. 2015. Next-generation sequencing approach for connecting secondary metabolites to biosynthetic gene clusters in fungi. Front. Microbiol. 5:774.

Chooi, Y.-H., Cacho, R. and Tang, Y. 2010. Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. Chem. Biol. 17, 483–494.

Finkelstein, E., Amichai, B. and Grunwald, M.H. 1996. Griseofulvin and its uses. Int. J. Antimicrob. Agents 6:189–194.

Frisvad, J.C. and Samson, R.A. 2004. Polyphasic taxonomy of *Penicillium* subgenus Penicillium A guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins. Stud. Mycol. 49:1–174.

Keller, O., Kollmar, M., Stanke, M. and Waack, S. 2011. A novel hybrid gene prediction method employing protein multiple sequence alignments. Bioinformatics 27:757–763.

Li, B., Zong, Y., Du, Z., Chen, Y., Zhang, Z., Qin, G., Zhao, W. and Tian, S. 2015. Genomic characterization reveals insights into patulin biosynthesis and pathogenicity in *Penicillium* species. Mol. Plant Microbe Interact. 28:635–647.

Pianzzola, M.J., Moscatelli, M. and Vero, S. 2004. Characterization of *Penicillium* isolates associated with blue mold on apple in Uruguay. Plant Disease 88, 23–28.

Samson, R.A. and Frisvad, J.C. 2004. *Penicillium* subgenus *Penicillium*: new taxonomic schemes, mycotoxins and other extrolites. Stud. Mycol. 49:1-251.

Shim, S.H., Swenson, D.C., Gloer, J.B., Dowd, P.F. and Wicklow, D.T. 2006. Penifulvin A: a sesquiterpenoid-derived metabolite containing a novel dioxa[5,5,5,6]fenestrane ring system from a fungicolous isolate of *Penicillium griseofulvum*. Organic Lett. 8:1225–1228.

Stamatakis, A., Ludwig, T., and Meier, H. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 21:456–463.

Tannous, J., El Khoury, R., Snini, S.P., Lippi, Y., El Khoury, A., Atoui, A., Lteif, R., Oswal, I.P. and Puel, O. 2014. Sequencing, physical organization and kinetic expression of the patulin biosynthetic gene cluster from *Penicillium expansum*. Int. J. Food Microbiol. 189:51–60.

Table

Table 1. Genome assembly statistics for the *P. griseofulvum* PG3 genome.

	Penicillium griseofulvum PG3
Genome size	29.3 Mb
Number of contigs	363
Number of contigs > 100kb	14
N50	2.8 Mb
Number of Ns	322
GC content	0.47
Number of predicted proteins	9631
Average protein length	521 aa

Figures:

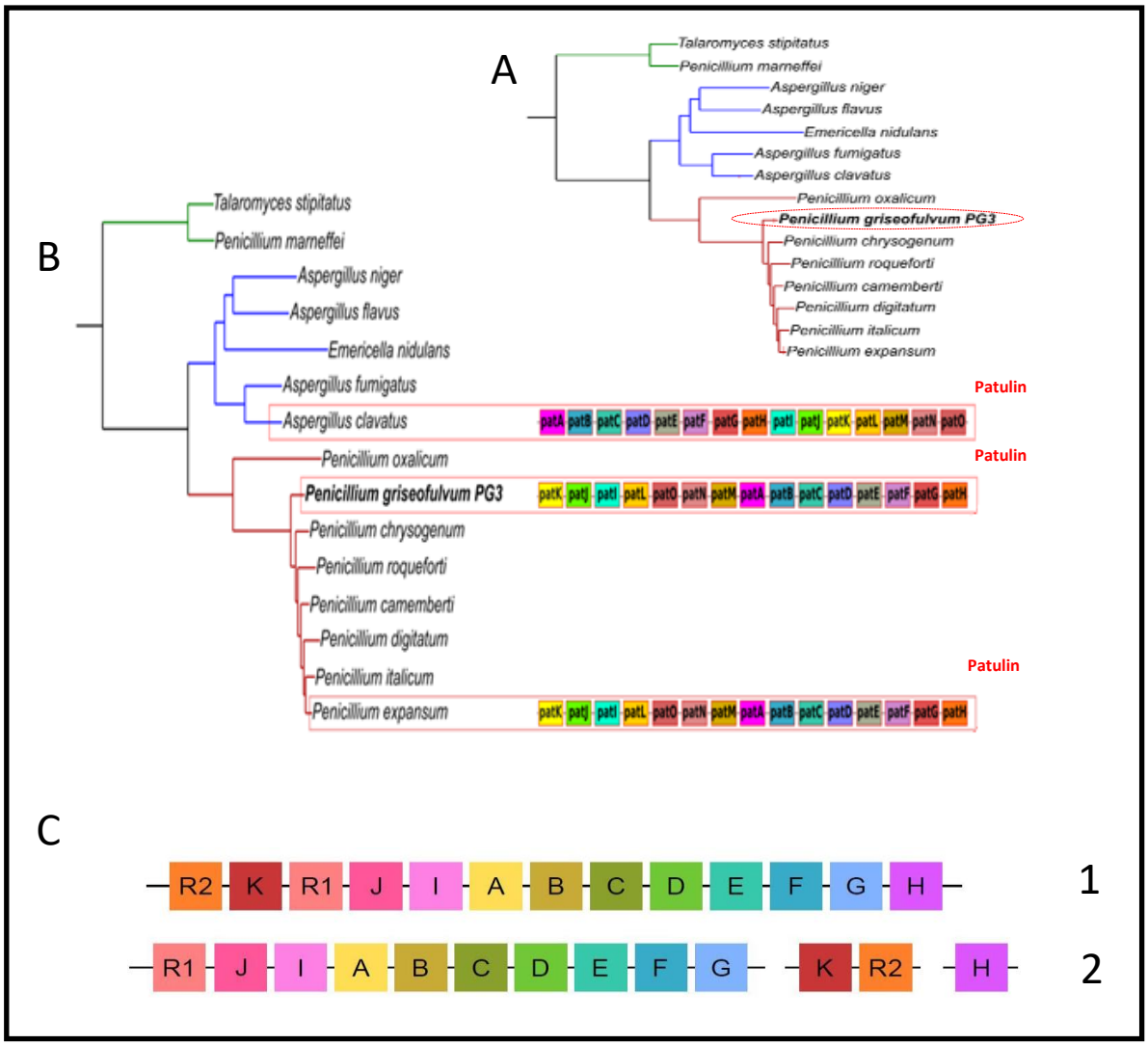

Fig. 1. Species tree showing the phylogenetic position of *P. griseofulvum* across the studied species and summary of the detected secondary metabolism clusters in PG3.
A) Maximum likelihood species tree derived through gene concatenation of 2134 single-copy genes that are present in PG3 and in the other 14 fully-sequenced *Penicillium* and *Aspergillus* considered species. All bootstrap values are maximal (100).
B) Comparison of patulin cluster genes in PG3 and in other phylogenetically close patulin producing species. The species used for comparison are *P. expansum* and *A. clavatus* which have a complete patulin gene cluster. Each gene is marked by a square and named according to their position in the original cluster described in *A. clavatus*.
C) Schematic representation of the griseofulvin gene cluster in *P. aethiopicum* (C1) and PG3 (C2). Each gene is named according to its position in the original cluster described in *P. aethiopicum*.

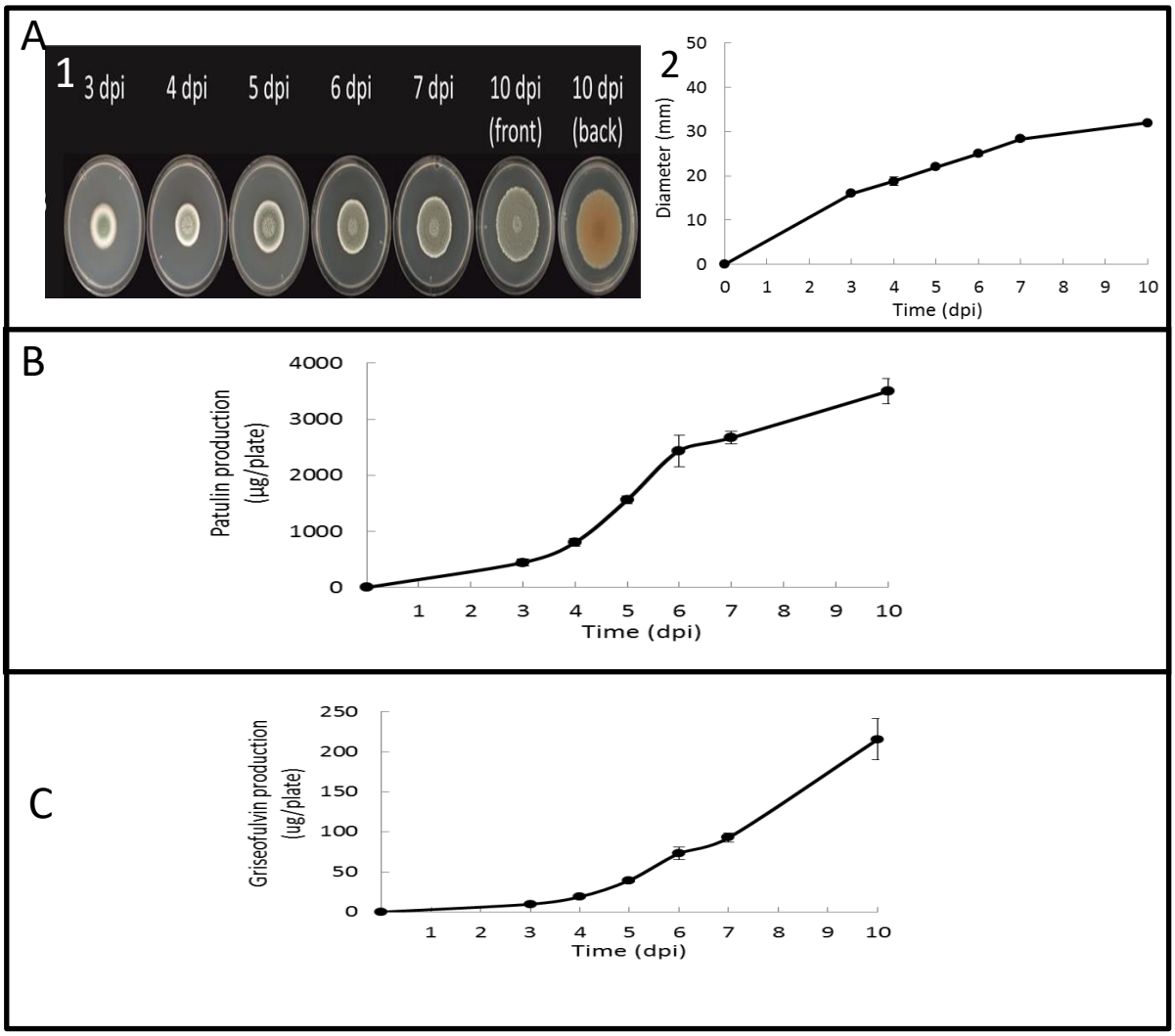

Fig. 2. Fungal growth and secondary metabolites production by PG3.
A) Colony view (A1) and diameter (A2) of *P. griseofulvum* strain PG3.
Spore suspensions of PG3 was inoculated on the center of 55 mm PDA plates and incubated for up to 10 days post inoculation (dpi) at 24°C in the dark. Error bars indicate standard deviations of three biological replicates.
B) Patulin production by *P. griseofulvum in vitro*. C) Griseofulvin production by *P. griseofulvum in vitro.*