British Biotechnology Journal
10(2): 1-13, 2016, Article no.BBJ.19400
ISSN: 2231–2927, NLM ID: 101616695

SCIENCEDOMAIN *international*
www.sciencedomain.org

Genomic Analysis of Fungal Species No.11243 Mutant Strains Provides Insights into the Relationship between Mutations and High Productivity

Makoto Matsui[1], Tatsuya Yokoyama[1], Toshitaka Kumagai[2], Kaoru Nemoto[3], Goro Terai[4], Masayuki Machida[5], Takashi Shibata[1] and Sachiyo Aburatani[6*]

[1]Biotechnology Labs, Astellas Pharma Inc., Japan.
[2]Fermlab Inc., Japan.
[3]Technology Research Association of Highly Efficient Gene Design (TRAHED), Japan.
[4]Advanced Technology Research and Development Institute, INTEC Inc., Japan.
[5]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Japan.
[6]Biotechnology Research Institute for Drug Discovery, National Institute of Advanced Industrial Science and Technology (AIST), Japan.

Authors' contributions

This work was carried out in collaboration between all authors. Authors M. Matsui, TS and SA designed the whole study. Authors M. Matsui and TY wrote the protocol. Authors TY, KN, M. Machida and TS designed and performed the genome sequence analysis. Authors TK and GT managed the computational investigation of the genomic analysis. Author M. Matsui wrote the first draft of the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BBJ/2016/19400
Editor(s):
(1) Christopher Cullis, Francis Hobart Herrick Professor and Chair of Biology, Case Western Reserve University, USA.
Reviewers:
(1) Claudia Spampinato, Universidad Nacional de Rosario, Argentina.
(2) Nina Kulakova, Limnological Institute SB RAS, Russia.
Complete Peer review History: http://sciencedomain.org/review-history/11692

Original Research Article

Received 8th June 2015
Accepted 11th July 2015
Published 6th October 2015

ABSTRACT

Aims: FR901469 is a novel antifungal antibiotic produced by fungal species No.11243. Although we have several FR901469 high-producing mutant strains, the mechanism of high productivity is unclear. This study aims to unravel the relationship between mutations and FR901469 productivity.

**Corresponding author: Email: s.aburatani@aist.go.jp;*

Methodology: We performed genome sequence analysis of mutant strains and detected mutated genes. Subsequently, we classified mutated genes into functional categories and searched the categories in which mutated genes were accumulated as generations progressed.
Results: We found that genome regions of two scaffolds were amplified and one of those contained putative FR901469 biosynthesis gene cluster. Moreover, we detected totally 396 mutated genes from 14 mutant strains and the genes within the “Replication, recombination and repair”, “Signal transduction mechanisms” and “Transcription” categories accumulated this mutation.
Conclusion: Our study suggests that productivity improvement occurs via the following two mechanisms: the amplification of putative FR901469 biosynthesis gene cluster and mutations of genes categorized as “Replication, recombination and repair”, “Signal transduction mechanisms” and “Transcription”.

Keywords: FR901469; fungal sp. No.11243; genomic analysis; mutation.

1. INTRODUCTION

Microorganisms produce various secondary metabolites, and many of them are utilized for drug substances and precursors, such as the antibiotic penicillin, the immunosuppressant FK506 [1-2] and the precursor of the antifungal agent FR901379 [3]. There are 23,000 known active compounds from microorganisms, and this number has been growing [4]. Moreover, the development of next-generation sequencing (NGS) technologies and genome-mining methods has brought about efficient detection of novel secondary metabolites [5]. It is thought that new secondary metabolites will be discovered by technological innovation and that they will be important resources for drug discovery.

Microorganisms in wild-type strains can produce a small amount of their specific secondary metabolites in general. In terms of supply of drug substances and economic efficiency, improvements in secondary metabolite production are important for drug development. Conventional methods for increasing productivity in microorganisms rely on culture optimization and high-producing strain development. These approaches have been successful for increasing the amount of substances. In our former investigation, FR901379 production has been increased 30 times by mutant selection and medium optimization [6]. Moreover, we can also improve a strain with genetic technology in recent years. For example, the productivity of cephalosporin C was improved by introducing *cefT* [7] and *cefG* [8]. We can add or delete targeted characteristics of strains by genetic engineering, so the breeding efficiency is expected to improve. Because many genome sequences of microorganisms have been clarified, it is believed that gene recombination will become one of the main techniques for productivity improvement in microorganisms.

The substance FR901469 is a novel antifungal antibiotic produced by fungal species No.11243 (No.11243 only hereafter), which was isolated by screening for new beta-1,3-glucan synthase inhibitors [9-11]. In our industrial process development of FR901469, mutant strains have been obtained from the wild-type strain by UV irradiation. Among the mutant strains, strains with high productivity of FR901469 have been separated in house. These artificial strain selections have been performed for several generations. The genealogy of mutant strains is shown in Fig. 1(a), and the relative FR901469 productivity is shown in Fig. 1(b). The most productive mutant strain can produce FR901469 at approximately nine times the amount produced by the wild-type strain, but the mechanisms of high productivity in mutant strains remain unclear. If the mechanism of high productivity is clarified, we can apply the knowledge to further productivity improvement.

In this study, to unravel the relationship between mutations and FR901469 productivity, we performed genomic analysis of mutant strains. At first, we performed genome sequencing of all No.11243 mutants and identified mutated genes. Subsequently, we classified the mutant genes by function and discovered the gene functions in which the mutations were accumulated. Those functions in which mutations were accumulated were thought to be related to FR901469 productivity. Our mutation analysis reveals the relationship between mutation and FR901469 productivity.

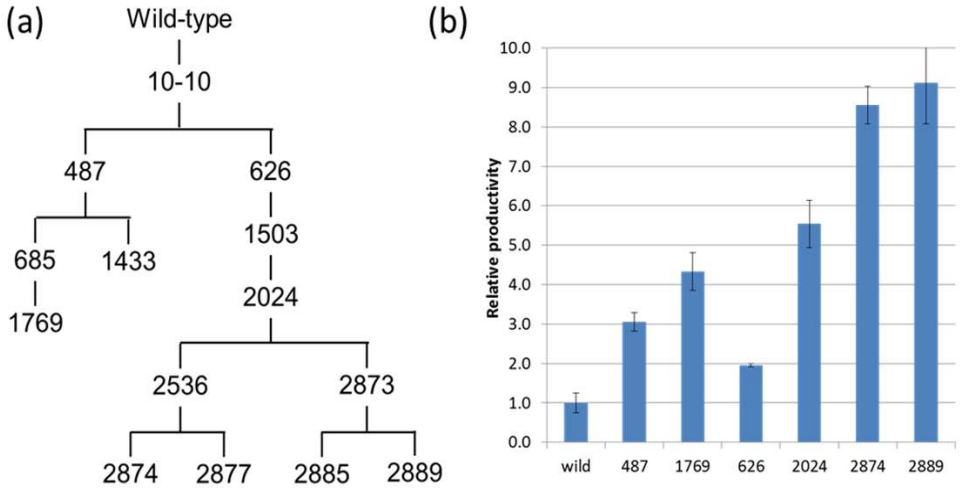

Fig. 1. No.11243 and its mutant strains

(a) Genealogy of mutant strains

(b) The relative FR901469 productivity of wild-type and mutant strains. Mean ± S.E.M = Mean values ± Standard error of means of three experiments

2. MATERIALS AND METHODS

2.1 Strains and Genome Sequence

Fungal species No.11243 (Taxonomy ID: 1603295, FERM BP-3373) was used as the wild-type strain. The mutant strains in Fig. 1(a) in the Astellas culture collection were obtained from the wild-type strain by UV irradiation. In the mutation treatment, we used bacterical lamp GL15 (Panasonic, Osaka, Japan) and irradiated UV to the cell until the survival rate became one to five percent. The genome sequence of the wild-type strain (DDBJ/EMBL/GenBank under the accession numbers DF938580 through DF938599) [12] was used as the reference genome.

2.2 Cultivation and Productivity Analysis of FR901469

A seed medium (20 ml), containing 4% soluble starch, 2% cottonseed flour (Sigma-aldrich, MO, USA), 1% soybean flour (Sigma-aldrich), 1% KH_2PO_4 and 0.2% $CaCO_3$ was poured into each of 100-ml Erlenmeyer flasks and sterilized at 121℃ for 20 minutes. A quadrangle piece with sides 5-mm of plate culture of fungus strain No.11243 or its mutants were inoculated to each medium and cultured under shaking condition (220 rpm) at 25℃ for five days.

A production medium (20 ml) consisting of 7% soluble starch, 1% glucose, 4% cottonseed flour, 6% corn steep liquor (Sigma-aldrich), 1% $(NH_4)_2SO_4$, 0.5% Bacto™ Peptone (BD, MD, USA), 0.44% K_2HPO_4, 1.0% KH_2PO_4, 0.1% antiform PE-L (Wako) and 0.1% antiform SI (Wako) was adjusted to pH 6.0 by adding 1M NaOH, poured into each of 100-ml Erlenmeyer flasks and sterilized at 121℃ for 20 minutes. Resultant seed culture broth (400 µL) was transferred to the production medium and cultured under shaking condition (220 rpm) at 25℃ for eight days.

To extract FR901469, 200 µL of 8-day culture was picked out into 1.5-mL Eppendorf Tube, mixed with 800 µL ethanol, and incubated for 30 minutes at room temperature. The debris was removed by centrifuging at 15,000 rpm at room temperature for 10 minutes, and the supernatant was filtered using filter units with a 0.22 µm pore size (Nacalai Tesque).

A 2-µL aliquot of extract was separated on an Acquity UPLC system (Waters, Milford, MA, USA) using a ACQUITY UPLC BEH C18 1.7 µm 2.1 × 50 mm reversed-phase column (Waters) and eluted with gradient of 36-46% aqueous acetonitrile (Wako) containing 0.1% formic acid (Wako) at a flow rate of 0.6 mL/min. The production of FR901469 was quantified by the peak area of absorbance at 210 nm.

All chemicals used were of analytical grade unless otherwise noted.

2.3 Genome Sequencing and Detection of Single Nucleotide Variants (SNVs), Insertions and Deletions (InDels)

Fragment libraries of genomic DNA from mutant strains were constructed with a SOLiD Library Construction kit (Life Technologies). We used the SOLiD 5500xl system (Life Technologies) to make single–end reads with a length of 50 bp. The quality of sequenced reads was checked and high-quality reads were mapped to the reference genome by bwa [13]. The SNVs and InDels were detected by VarScan 2 [14].

2.4 Calculating Relative Depth of Each Scaffold

The depth of each base was calculated by BEDTools [15]. The ratio of depth in the scaffold was calculated as follows:

$$R_S^W = \frac{D_N}{D_W}$$

where R_S^W is the ratio of depth of scaffold *N* in strain *W*. D_N is the average depth of scaffold *N* in strain *W*, and D_W is the average depth of whole genome in strain *W*.

2.5 Mutated Gene Detection and Annotation

We defined genes that have missense, nonsense, and frameshift mutations as mutated genes. The protein annotations of mutated genes were predicted from the annotations with the homologous protein (E-value < $1e\times10^{-5}$) using the SwissProt database. Furthermore, mutated genes were classified into functional categories in eggNOG by the following steps: (1) protein sequence was blasted against eggNOG v3.0 [16] database using a cut-off E-value > $1e\times10^{-20}$ and coverage < 80%; and (2) if there were homologous proteins, the mutated gene was categorized to its homologue's functional categories.

3. RESULTS

3.1 Genomic Characterization of No. 11243 and Its Mutants

To detect mutated bases, we sequenced all of the genomic DNA from No.11243 mutants using the SOLiD 5500xl system and mapped sequence reads to the reference wild-type genome. At this time, we noticed that the depths of scaffold0004 (accession no. DF938583) and scaffold0016 (accession no. DF938595) were distinctively higher than that of the scaffolds in most of the mutant strains, although the depths of all scaffolds were similar in the wild-type strain. The ratios of each scaffold depth in all strains are displayed in Table 1. In Table 1, the marked numbers indicated the highest and second highest values in each column. Except for the 685, 1769 and 2889 strains, the ratios of scaffold0004 and scaffold0016 show higher values than that of other scaffolds. In the 685, 1769 and 2889 strains, the ratios of scaffold0004 and scaffold0016 show truly higher value than that of most of the others, but some of the other scaffolds also have high scores.

Next, we counted SNVs and InDels for detecting mutation genes. Furthermore, we calculated the variant frequencies of SNVs. In this paper, we defined "Variant frequency of SNV" as the proportion of variant nucleotides to all of the nucleotides among the reads mapped to the site of SNV. The average variant frequencies in each scaffold are presented in Fig. 2. Although the average variant frequency was approximately 0.5 in almost scaffolds, those of scaffold0004 and scaffold0016 are 0.35 and 0.34, respectively. Because the lengths of scaffold0019 and scaffold0020 are short, none and only one mutated base was detected in these scaffolds, respectively.

3.2 Mutated Genes in Each Strain

Mutated genes were counted in genealogy. Missense, nonsense, and frameshift mutations were defined as mutated genes. The number of annotated, not annotated and sum of mutated genes is listed in Fig. 3(a). The smallest number of genes was mutated from the 2024 to 2873 strains. On the other hand, the most genes were mutated from the 10-10 to 626 strains. A total of 397 mutated genes were detected in all mutant strains. The relationship between mutation frequency and genealogy is shown in Fig. 3(b). The numbers of mutated genes in one generation tended to decrease as the generations progressed. Mutated genes were annotated using the Swiss-Prot database, and we found that 280 genes have homologues in this database. The annotations are listed in Supplementary Table 1.

3.3 Functional Categories of Mutated Genes

We classified the mutated genes into functional categories of eggNOG. We calculated the ratio of mutated genes in each functional category. Then, the ratios between the parental strain and mutant strain were compared. If the ratio in the mutant strain is higher than that of the parent strain, then the mutated genes are considered to be in the functional category that is increased in the strain. The increasing and decreasing pattern of ratios is displayed in Table 2. We focused on three strain lines starting from the wild-type to the 1769, 2874 and 2889 strains. When the ratios of mutated genes increased sequentially more than three times, those mutated genes defined as being accumulated in that functional category. From this table, the genes with accumulated mutations fell within the "Replication, recombination and repair", "Signal transduction mechanisms" and "Transcription" categories.

We focused on three functional categories in which mutated genes were accumulated. A total of 33 mutated genes were annotated to "Replication, recombination and repair", and 14 of those genes were involved in DNA repair. RTT107, ESO1, XRCC2, POL4 and CDC48 are relevant to double-strand break repair [17-21]. RAD25, RAD3, RAD14, Replication factor C subunit 1, RHP42, HRQ1 and Exodeoxyribonuclease 1 have roles in nucleotide excision repair [22-28]. PIN4 and RAD53 are related to cell cycle progression and the DNA damage response [29-30].

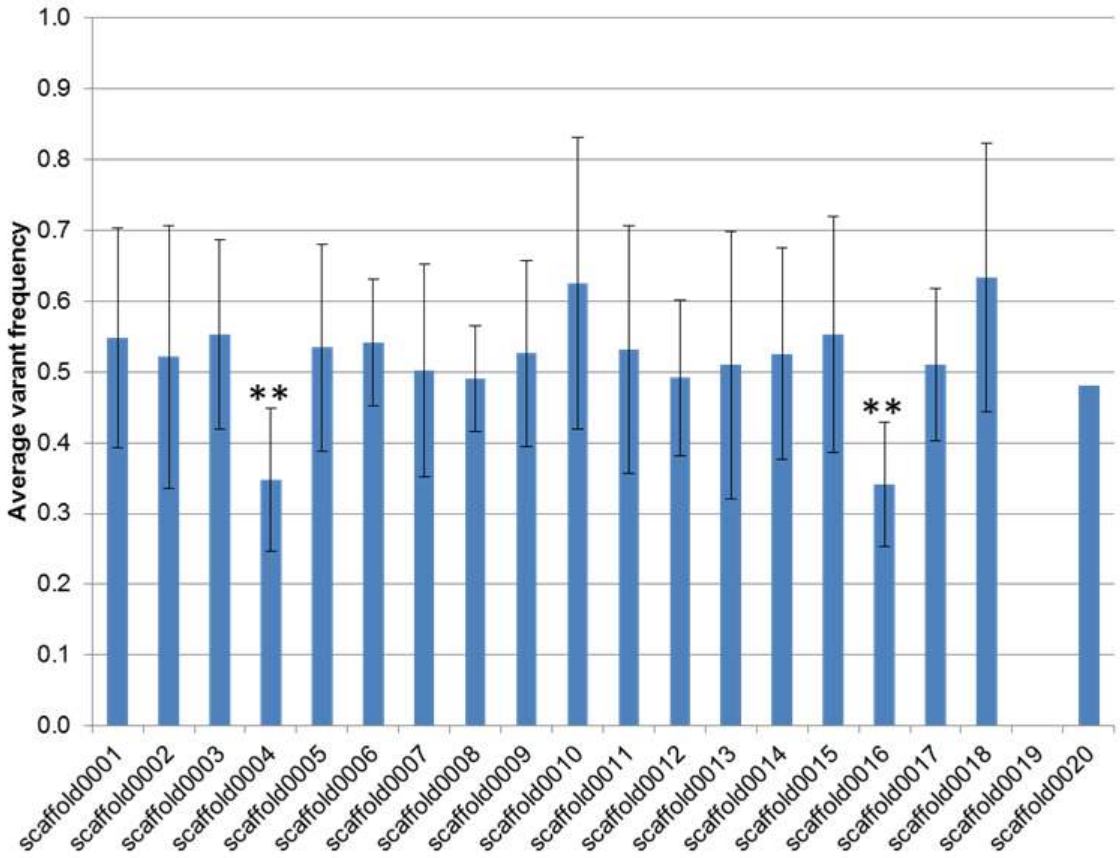

Fig. 2. Average variant frequency in each scaffold

The average variant frequency in each scaffold
Significant from scaffold0001,
*** P < 0.001. Mean ±S.D.M = Mean values ± Standard deviation of means*
Scaffold0019 has no SNV, so there was no bar graph, and scaffold0020 has only one SNV, so there was no error bar

Table 1. The ratio of depth in each scaffold

	Wild-type	10-10	487	626	685	1433	1503	1769	2024	2536	2873	2874	2877	2885	2889
scaffold0001	1.0	1.0	1.0	0.9	0.9	1.0	0.9	1.2	1.0	1.0	1.0	0.9	0.9	1.0	**1.3**
scaffold0002	1.0	1.0	0.9	0.9	0.9	0.9	0.9	0.9	1.0	1.0	1.0	1.0	1.0	1.0	0.9
scaffold0003	1.0	1.0	0.9	0.9	0.9	0.9	0.9	0.9	1.0	1.0	1.0	0.9	0.9	1.0	0.9
scaffold0004	1.0	**1.4**	**1.4**	**1.8**	**1.3**	**1.4**	**1.4**	**1.3**	**1.4**	**1.5**	**1.2**	**1.4**	**1.4**	**1.2**	1.2
scaffold0005	1.0	1.0	1.0	0.9	0.9	0.9	0.9	0.9	0.9	1.0	1.0	0.9	0.9	1.0	0.9
scaffold0006	1.0	1.0	1.0	0.9	1.0	0.9	1.0	0.9	1.0	1.0	1.0	1.0	1.0	1.0	0.9
scaffold0007	1.0	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
scaffold0008	1.0	1.0	1.0	0.9	1.0	0.9	1.0	0.9	1.0	1.0	1.0	1.0	1.0	1.0	0.9
scaffold0009	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
scaffold0010	1.0	0.9	0.9	0.9	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0	0.9
scaffold0011	1.0	1.0	1.0	0.9	0.9	0.9	1.0	0.9	0.9	1.0	1.0	1.0	1.0	1.0	0.9
scaffold0012	1.0	0.9	1.0	0.9	1.0	1.0	1.0	**1.3**	1.0	1.0	1.0	1.0	1.0	1.0	0.9
scaffold0013	1.0	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0	0.9	0.9	0.9	0.9
scaffold0014	1.1	1.0	1.0	1.0	1.2	1.3	1.0	1.2	1.0	1.0	1.0	1.0	1.0	1.0	1.2
scaffold0015	1.0	1.0	0.9	0.9	0.9	0.9	0.9	1.1	0.9	0.9	0.9	0.9	0.9	0.9	0.9
scaffold0016	1.0	**1.4**	**1.5**	**1.5**	**1.5**	**1.5**	**1.6**	**1.4**	**1.6**	**1.1**	**1.6**	**1.6**	**1.6**	**1.6**	**1.6**
scaffold0017	1.0	0.9	1.0	0.9	0.9	0.9	0.9	0.9	0.9	0.9	1.0	0.9	0.9	0.9	**1.3**
scaffold0018	1.0	0.9	1.0	0.9	**1.3**	1.3	1.0	**1.3**	0.9	1.0	1.0	1.0	1.0	1.0	0.9
scaffold0019	1.0	1.0	0.9	0.9	1.1	0.9	0.9	0.8	0.9	0.9	0.9	0.9	0.9	0.8	0.8
scaffold0020	1.1	1.0	0.9	0.8	0.8	0.8	0.8	1.1	0.8	0.8	0.8	0.8	0.8	0.8	0.7

This table shows the ratios of depth in each scaffold.
The marked numbers indicated the highest and the second highest values in each column

Table 2. The increasing and decreasing patterns of mutated genes in each functional category

Functional category	Wild to 1769			Wild to 2874					Wild to 2889				
	10-10 | 487	487 | 685	685 | 1769	10-10 | 626	626 | 1503	1503 | 2024	2024 | 2536	2536 - 2874	10-10 | 626	626 | 1503	1503 | 2024	2024 | 2873	2873 | 2889
Cell cycle control, cell division, chromosome partitioning (224)	+	+	-	+	+	-	-	-	+	+	-	-	-
Cell motility (15)	+	-	-	-	+	-	-	-	-	+	-	-	-
Cell wall/membrane/envelope biogenesis (172)	+	-	-	+	-	-	-	-	+	-	-	-	+
Cytoskeleton (131)	+	-	+	+	-	-	-	-	+	-	-	-	-
Defense mechanisms (79)	-	-	-	+	-	-	-	-	+	-	-	-	-
Intracellular trafficking, secretion, and vesicular transport (383)	-	-	-	-	-	-	+	-	-	-	-	-	-
Nuclear structure (25)	-	+	-	-	+	-	-	-	-	+	-	-	-
Posttranslational modification, protein turnover, chaperones (608)	+	+	-	-	+	+	-	+	-	+	+	-	-
Signal transduction mechanisms (482)	**+**	**+**	**+**	**+**	**+**	**+**	-	-	**+**	**+**	**+**	-	-
Extracellular structures (13)	-	-	-	-	-	-	-	-	-	-	-	-	-
Chromatin structure and dynamics (138)	+	+	-	-	+	-	-	+	-	+	-	-	-
Replication, recombination and repair (386)	**+**	**+**	**+**	-	-	**+**	**+**	**+**	-	-	+	-	-
RNA processing and modification (280)	-	+	+	+	-	+	-	-	+	-	+	-	-
Transcription (572)	**+**	**+**	**+**	-	**+**	**+**	**+**	**+**	-	+	+	-	-
Translation, ribosomal structure and biogenesis (415)	-	+	+	+	+	-	-	-	+	+	-	-	+
Amino acid transport and metabolism (627)	+	-	-	-	+	-	-	+	-	+	-	-	-
Carbohydrate transport and metabolism (569)	+	-	-	+	-	+	-	-	+	-	+	-	+
Coenzyme transport and metabolism (202)	-	-	-	-	-	-	-	+	-	-	-	-	-
Energy production and conversion (491)	-	+	-	-	-	-	+	+	-	-	-	-	+
Inorganic ion transport and metabolism (383)	+	-	+	-	-	+	-	-	-	-	+	-	-
Lipid transport and metabolism (515)	-	+	-	+	-	-	+	-	+	-	-	-	-
Secondary metabolites biosynthesis, transport and catabolism (424)	+	+	-	+	-	-	-	+	+	-	-	-	+
Nucleotide transport and metabolism (116)	-	-	-	-	-	-	-	-	-	-	-	-	-

This table shows the increasing and decreasing patterns of mutated genes in each functional category. The number of whole genes categorized within each functional category is indicated in parentheses. The plus symbol indicates that the ratio of mutated genes in that functional category increased in the mutant strain.
The cells in which three plus symbols appeared sequentially are highlighted

(a)

Strains		Number of mutated genes			
Parent	Mutant	A	NA	A+NA	Sum
wild-type	10-10	10	4	14	14
10-10	487	30	12	42	56
10-10	626	53	16	69	83
487	685	27	6	33	89
487	1433	14	8	22	78
626	1503	41	13	54	137
685	1769	38	15	53	142
1503	2024	12	7	19	156
2024	2536	7	4	11	167
2024	2873	1	4	5	161
2536	2874	17	5	22	189
2536	2877	7	4	11	178
2873	2885	9	10	19	180
2873	2889	13	9	22	183
Total		280	116	396	-

(b)

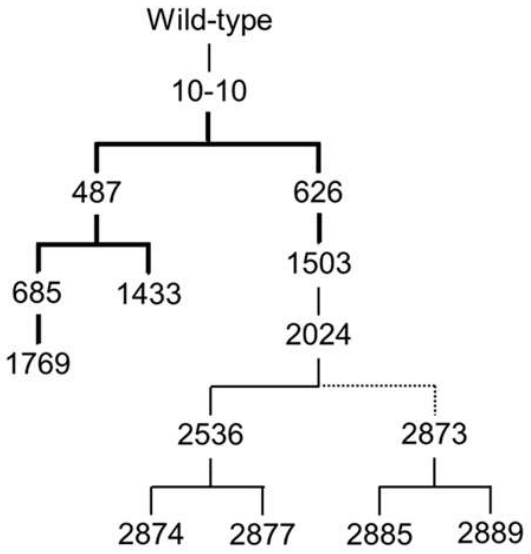

Fig. 3. The number of mutated genes in each strain

(a) The number of mutated genes. A: The number of annotated genes by Swiss-Prot. NA: The number of un-annotated genes. A+NA: Sum of annotated and un-annotated mutated genes. Sum; Accumulated sum of mutated genes from wild-type strain.

(b) Genealogy of mutant strains. The line types show the number of mutated genes.

Bold line: more than 30 genes were mutated between two strains, Thin line: from 10 to 30 genes were mutated between two strains, Broken line: less than 10 genes were mutated between two strains

A total of 38 mutated genes were annotated to "Signal transduction mechanisms". Many of these genes are involved in the nutrient sensing and the MAPK pathway. A total of 12 mutated genes were predicted as nutrient sensing-related genes, and 6 of these genes were predicted as TOR signaling (Target of Rapamycin signaling)-related genes. TOR signaling is a highly conserved signaling pathway. It regulates cellular responses toward nutrient availability, especially nitrogen and amino acid accessibility [31]. They included RHO1, PKH1, LST8, FKBP12-associated protein 1, TSC11 and FHL1 [32-37]. The remaining six genes were related to glucose sensing. SYG1, SOK2 and Guanine nucleotide-binding protein subunit alpha are involved in cAMP signaling [38-40]. PP1-2 and SIP3 are related to the SNF1 pathway [41-42]. Pyruvate dehydrogenase phosphatase regulates pyruvate dehydrogenase activity via dephosphorylation [43]. MAPK pathways (Mitogen-Activated Protein Kinase pathways) are signaling transduction pathways that regulate various physiological responses. A total of nine mutated genes were predicted MAPK pathway-related genes. They included STE11, PBS2, PTP2, HRK1, NIK1, RHO1, BCK1, PKH1 and MAK1 [44-52]. Two of these genes are common to TOR signaling (RHO1 and PKH1). In addition, STE11, PBS2, PTP2, HRK1 and NIK1 are related to the Hog1 pathway, and RHO1, BCK1 and PKH1 are related to the Mpk1/Slt2 Pathway. MAK1 relates to oxidative stress-induced MAPK pathway in *Schizosaccharomyces pombe*.

A total of 45 mutated genes were annotated to "Transcription", but many of them were also annotated "Replication, recombination and repair" and "Signal transduction mechanisms" (STE11, NIK1, FKBP12-associated protein 1, BCK1, SOK2, HRK1, PKH1, PBS2, FHL1, Regulator of Ty1 transposition protein 107, RAD25, RAD3, KIN28 and RAD53). Among the remaining genes, many genes were related to the transcription mechanism itself. eIF4A and TFIID subunit 2 are involved in transcription initiation [53-54]. MOT1, General negative regulator of transcription subunit 4, and SREP are related to TATA-binding [55-57]. SSU72, RPC2, RNA polymerase II subunit A C-terminal domain phosphatase and

DNA-directed RNA polymerase II subunit RPB2 are components of RNA polymerase [58-61]. Pre-mRNA-processing protein 45 and RSE1 work at RNA splicing [62-63]. There are six other genes that have putative transcription factor domains. Sterol regulatory element-binding protein 1 has a HLH domain [64], pacC/RIM101 has a zinc finger domain [65], and the other four putative transcription factors, annotated uncharacterized transcriptional regulatory protein and ACU-15, have fungal transcription factor regulatory middle homology regions [66]. These genes are expected to regulate transcription as transcription factors.

4. DISCUSSION

In this study, we found that the genome region of two scaffolds, scaffold0004 and scaffold0016, increased in the 10-10 strain (table 1). Interestingly, there is a putative FR901469 biosynthesis gene cluster in scaffold0004. In the case of the penicillin-producing fungus *Penicillium chrysogenum*, it was reported that the amplification of the penicillin biosynthesis gene cluster is shown in high-penicillin-producing industrial strains [67-68]. Likewise, it is believed that the amplification of the FR901469 biosynthesis gene cluster causes high productivity. Additionally, the average variant frequencies of scaffold0004 and scaffold0016 are about one third (Fig. 3), so it was suggested that the genome region of each of these scaffolds has three copies in genome. This suggestion is in agreement with the fact that the depths of mapping reads in scaffold0004 and scaffold0016 are higher than those in other scaffolds (Table 1).

We found that mutated genes in three functional categories were accumulated. These functional categories were accumulated in several strain lines, and genes within other categories had no tendency to be accumulated. Thus, it was expected that these three functional categories related to strain development. In mutated genes categorized as “Replication, recombination and repair”, many of the genes are relevant to DNA repair. Strain mutagenesis of No.11243 was caused by UV irradiation, and DNA repair mechanisms are believed to protect these genes from mutation. Therefore, mutations in DNA repair genes lead to other gene mutations. This other gene mutation will be important for generating productivity improvement strains. Thus, DNA repair gene mutations are related to productivity.

Mutated genes of “Signal transduction mechanisms” were also accumulated. Many of them were related to nutrient sensing and the MAPK pathway. Nutrient sensing regulates cell growth and metabolism according to nutrient availability. In general, secondary metabolites are produced in only special conditions, and one of the conditions may be the nutritional status of the cell. There are several reports of relationships between nutrient sensing genes and secondary metabolite production [69-70]. The mutation of genes related to nutrient signal transduction might change nutrient utilization and might lead to productivity improvement.

In yeasts and filamentous fungi, it is known that the MAPK pathway regulates sexual development and environmental stress responses [71-72]. Several studies have reported that the MAPK pathway was also related to secondary metabolites, for example, ochratoxin A in *Penicillium verrucosum* [73], sterigmatocystin in *Aspergillus nidulans* [74] and eumelanin in *Neurospora crassa* [75]. The MAPK pathway regulates many cell functions, so it was expected that the disturbance of the MAPK pathway by mutation would lead to FR901469 productivity. Genes categorized as “Transcription” were also accumulated. Generally, secondary metabolites are produced in only special conditions, and it is expected that controlling gene expression is one of the regulatory mechanisms for production. Several studies have reported that the production of secondary metabolites is regulated by transcription factors, for example, aflatoxin in *Aspergillus flavus* [76]. It was expected that mutations of transcription genes changed the transcription regulation of the FR901469 production, leading to productivity improvement.

5. CONCLUSION

In this study, we performed genomic analysis of FR901469 high-producing mutant strains of No.11243. Our study suggests that high productivity occurred via two mechanisms: the amplification of the putative FR901469 biosynthesis gene cluster and mutations of genes categorized as “Replication, recombination and repair”, “Signal transduction mechanisms” and “Transcription”. This result shed light on the mutations that occur in the genome during conventional strain development. These sequence analyses at the genome level will be useful for breeding new mutant strains.

ACKNOWLEDGEMENTS

This work was supported by the commission for Development of Artificial Gene Synthesis Technology for Creating Innovative Biomaterial from the Ministry of Economy, Trade and Industry (METI), Japan.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Kino T, Hatanaka H, Hashimoto M, Nishiyama M, Goto T, Okuhara M, et al. FK-506, a novel immunosuppressant isolated from a *Streptomyces*. I. Fermentation, isolation, and physico-chemical and biological characteristics. J Antibiot (Tokyo). 1987;40(9):1249-55.
2. Kino T, Hatanaka H, Miyata S, Inamura N, Nishiyama M, Yajima T, et al. FK-506, a novel immunosuppressant isolated from a *Streptomyces*. II. Immunosuppressive effect of FK-506 *in vitro*. J Antibiot (Tokyo). 1987;40(9):1256-65.
3. Fujie A, Iwamoto T, Sato B, Muramatsu H, Kasahara C, Furuta T, et al. FR131535, a novel water-soluble echinocandin-like lipopeptide: Synthesis and biological properties. Bioorg Med Chem Lett. 2001;11(3):399-402.
4. Demain AL. Importance of microbial natural products and the need to revitalize their discovery. J Ind Microbiol Biotechnol. 2014;41(2):185-201.
5. Bachmann BO, Van Lanen SG, Baltz RH. Microbial genome mining for accelerated natural products discovery: Is a renaissance in the making? J Ind Microbiol Biotechnol. 2014;41(2):175-84.
6. Kanda M, Tsuboi M, Sakamoto K, Shimizu S, Yamashita M, Honda H. Improvement of FR901379 production by mutant selection and medium optimization. J Biosci Bioeng. 2009;107(5):530-4.
7. Ullán RV, Liu G, Casqueiro J, Gutiérrez S, Bañuelos O, Martín JF. The cefT gene of *Acremonium chrysogenum* C10 encodes a putative multidrug efflux pump protein that significantly increases cephalosporin C production. Mol Genet Genomics. 2002; 267(5):673-83.
8. Gutiérrez S, Velasco J, Marcos AT, Fernández FJ, Fierro F, Barredo JL, et al. Expression of the cefG gene is limiting for cephalosporin biosynthesis in *Acremonium chrysogenum*. Appl Microbiol Biotechnol. 1997;48(5):606-14.
9. Fujie A, Iwamoto T, Muramatsu H, Okudaira T, Nitta K, Nakanishi T, et al. FR901469, a novel antifungal antibiotic from an unidentified fungus No.11243. I. Taxonomy, fermentation, isolation, physico-chemical properties and biological properties. J Antibiot (Tokyo). 2000;53(9): 912-9.
10. Fujie A, Iwamoto T, Muramatsu H, Okudaira T, Sato I, Furuta T, et al. FR901469, a novel antifungal antibiotic from an unidentified fungus No.11243. II. In vitro and *in vivo* activities. J Antibiot (Tokyo). 2000;53(9):920-7.
11. Fujie A, Muramatsu H, Yoshimura S, Hashimoto M, Shigematsu N, Takase S. FR901469, a novel antifungal antibiotic from an unidentified fungus No. 11243. III. Structure determination. J Antibiot (Tokyo). 2001;54(7):588-94.
12. Matsui M, Yokoyama T, Nemoto K, Kumagai T, Terai G, Arita M, et al. Genome sequence of fungal species No.11243, which produces the antifungal antibiotic FR901469. Genome Announc. 2015;3(2). pii: e00118-15.
13. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics. 2009;25(14): 1754-60.
14. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. Var Scan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012;22(3): 568-76.
15. Quinlan AR, Hall IM. BED Tools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.
16. Powell S, Szklarczyk D, Trachana K, Roth A, Kuhn M, Muller J, Arnold R, Rattei T, Letunic I, Doerks T, Jensen LJ, von Mering C, Bork P. egg NOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. Nucleic Acids Res. 2012;40(Database issue):D284-9.
17. Leung GP, Lee L, Schmidt TI, Shirahige K, Kobor MS. Rtt107 is required for recruitment of the SMC5/6 complex to DNA double strand breaks. J Biol Chem. 2011;286(29):26250-7.

18. Tanaka K, Yonekawa T, Kawasaki Y, Kai M, Furuya K, Iwasaki M, et al. Fission yeast Eso1p is required for establishing sister chromatid cohesion during S phase. Mol Cell Biol. 2000;20(10):3459-69
19. Cartwright R, Tambini CE, Simpson PJ, Thacker J. The XRCC2 DNA repair gene from human and mouse encodes a novel member of the recA/RAD51 family. Nucleic Acids Res. 1998;26(13):3084-9.
20. Tseng HM, Tomkinson AE. Processing and joining of DNA ends coordinated by interactions among Dnl4/Lif1, Pol4, and FEN-1. J Biol Chem. 2004;279(46): 47580-8.
21. Bergink S, Ammon T, Kern M, Schermelleh L, Leonhardt H, Jentsch S. Role of Cdc48/p97 as a SUMO-targeted segregase curbing Rad51-Rad52 interaction. Nat Cell Biol. 2013;15(5): 526-32.
22. Guzder SN, Sung P, Bailly V, Prakash L, Prakash S. RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription. Nature. 1994; 369(6481):578-81.
23. Sung P, Guzder SN, Prakash L, Prakash S. Reconstitution of TFIIH and requirement of its DNA helicase subunits, Rad3 and Rad25, in the incision step of nucleotide excision repair. J Biol Chem. 1996;271(18): 10821-6.
24. Guzder SN, Sommers CH, Prakash L, Prakash S. Complex formation with damage recognition protein Rad14 is essential for *Saccharomyces cerevisiae* Rad1-Rad10 nuclease to perform its function in nucleotide excision repair *in vivo*. Mol Cell Biol. 2006;26(3):1135-41.
25. Cullmann G, Fien K, Kobayashi R, Stillman B. Characterization of the five replication factor C genes of *Saccharomyces cerevisiae*. Mol Cell Biol. 1995;15(9): 4661-71.
26. Marti TM, Kunz C, Fleck O. Repair of damaged and mismatched DNA by the XPC homologues Rhp41 and Rhp42 of fission yeast. Genetics. 2003;164(2): 457-67.
27. Choi DH, Min MH, Kim MJ, Lee R, Kwon SH, Bae SH. Hrq1 facilitates nucleotide excision repair of DNA damage induced by 4-nitroquinoline-1-oxide and cisplatin in *Saccharomyces cerevisiae*. J Microbiol. 2014;52(4):292-8.
28. Tran PT, Fey JP, Erdeniz N, Gellon L, Boiteux S, Liskay RM. A mutation in EXO1 defines separable roles in DNA mismatch repair and post-replication repair. DNA Repair (Amst). 2007;6(11):1572-83.
29. Pike BL, Yongkiettrakul S, Tsai MD, Heierhorst J. Mdt1, a novel Rad53 FHA1 domain-interacting protein, modulates DNA damage tolerance and G(2)/M cell cycle progression in *Saccharomyces cerevisiae*. Mol Cell Biol. 2004;24(7):2779-88.
30. Allen JB, Zhou Z, Siede W, Friedberg EC, Elledge SJ. The SAD1/RAD53 protein kinase controls multiple checkpoints and DNA damage-induced transcription in yeast. Genes Dev. 1994;8(20):2401-15.
31. Loewith R, Hall MN. Target of rapamycin (TOR) in nutrient signaling and growth control. Genetics. 2011;189(4):1177-201. DOI: 10.1534/genetics.111.133363. Review.
32. Yan G, Lai Y, Jiang Y. The TOR complex 1 is a direct target of Rho1 GTPase. Mol Cell. 2012;45(6):743-53.
33. Mirisola MG, Taormina G, Fabrizio P, Wei M, Hu J, Longo VD. Serine- and threonine / valine-dependent activation of PDK and Tor orthologs converge on Sch9 to promote aging. PLoS Genet. 2014;10(2): e1004113.
34. Chen EJ, Kaiser CA. LST8 negatively regulates amino acid biosynthesis as a component of the TOR pathway. J Cell Biol. 2003;161(2):333-47.
35. Kunz J, Loeschmann A, Deuter-Reinhard M, Hall MN. FAP1, a homologue of human transcription factor NF-X1, competes with rapamycin for binding to FKBP12 in yeast. Mol Microbiol. 2000;37(6):1480-93. Erratum in: Mol Microbiol. 2001;39(4):1107.
36. Ho HL, Shiau YS, Chen MY. *Saccharomyces cerevisiae* TSC11/AVO3 participates in regulating cell integrity and functionally interacts with components of the Tor2 complex. Curr Genet. 2005;47(5): 273-88.
37. Martin DE, Soulard A, Hall MN. TOR regulates ribosomal protein gene expression via PKA and the Fork head transcription factor FHL1. Cell. 2004; 119(7):969-79.
38. Spain BH, Koo D, Ramakrishnan M, Dzudzor B, Colicelli J. Truncated forms of a novel yeast protein suppress the lethality of a G protein alpha subunit deficiency by interacting with the beta subunit. J Biol Chem. 1995;270(43):25435-44.
39. Ward MP, Gimeno CJ, Fink GR, Garrett S. SOK2 may regulate cyclic AMP-dependent protein kinase-stimulated growth and

pseudohyphal development by repressing transcription. Mol Cell Biol. 1995;15(12): 6854-63.

40. Nakafuku M, Obara T, Kaibuchi K, Miyajima I, Miyajima A, Itoh H, et al. Isolation of a second yeast *Saccharomyces cerevisiae* gene (GPA2) coding for guanine nucleotide-binding regulatory protein: studies on its structure and possible functions. Proc Natl Acad Sci. USA. 1988;85(5):1374-8.
41. Tu J, Carlson M. The GLC7 type 1 protein phosphatase is required for glucose repression in *Saccharomyces cerevisiae*. Mol Cell Biol. 1994;14(10):6789-96.
42. Lesage P, Yang X, Carlson M. Analysis of the SIP3 protein identified in a two-hybrid screen for interaction with the SNF1 protein kinase. Nucleic Acids Res. 1994;22(4):597-603.
43. Gey U, Czupalla C, Hoflack B, Rödel G, Krause-Buchholz U. Yeast pyruvate dehydrogenase complex is regulated by a concerted activity of two kinases and two phosphatases. J Biol Chem. 2008;283(15): 9759-67.
44. Posas F, Witten EA, Saito H. Requirement of STE50 for osmostress-induced activation of the STE11 mitogen-activated protein kinase kinase kinase in the high-osmolarity glycerol response pathway. Mol Cell Biol. 1998;18(10):5788-96.
45. Posas F, Saito H. Osmotic activation of the HOG MAPK pathway via Ste11p MAPKKK: Scaffold role of Pbs2p MAPKK. Science. 1997;276(5319):1702-5.
46. Wurgler-Murphy SM, Maeda T, Witten EA, Saito H. Regulation of the *Saccharomyces cerevisiae* HOG1 mitogen-activated protein kinase by the PTP2 and PTP3 protein tyrosine phosphatases. Mol Cell Biol. 1997;17(3):1289-97.
47. Kim SY, Ko YJ, Jung KW, Strain A, Nielsen K, Bahn YS. Hrk1 plays both Hog1-dependent and -independent roles in controlling stress response and antifungal drug resistance in Cryptococcus neoformans. PLoS One. 2011;6(4):e18769.
48. Hagiwara D, Takahashi-Nakaguchi A, Toyotome T, Yoshimi A, Abe K, Kamei K, et al. NikA/TcsC histidine kinase is involved in conidiation, hyphal morphology, and responses to osmotic stress and antifungal chemicals in *Aspergillus fumigatus*. PLoS One. 2013;8(12):e80881.
49. Boorsma A, De Nobel H, Ter Riet B, Bargmann B, Brul S, Hellingwerf KJ, et al. Characterization of the transcriptional response to cell wall stress in *Saccharomyces cerevisiae*. Yeast. 2004; 21(5):413-27.
50. Lee KS, Irie K, Gotoh Y, Watanabe Y, Araki H, Nishida E, Matsumoto K, Levin DE. A yeast mitogen-activated protein kinase homolog (Mpk1p) mediates signalling by protein kinase C. Mol Cell Biol. 1993;13(5): 3067-75.
51. Roelants FM, Torrance PD, Bezman N, Thorner J. Pkh1 and Pkh2 differentially phosphorylate and activate Ypk1 and Ykr2 and define protein kinase modules required for maintenance of cell wall integrity. Mol Biol Cell. 2002;13(9):3005-28.
52. Buck V, Quinn J, Soto Pino T, Martin H, Saldanha J, Makino K, Morgan BA, Millar JB. Peroxide sensors for the fission yeast stress-activated mitogen-activated protein kinase pathway. Mol Biol Cell. 2001;12(2): 407-19.
53. Blum S, Mueller M, Schmid SR, Linder P, Trachsel H. Translation in Saccharomyces cerevisiae: Initiation factor 4A-dependent cell-free system. Proc Natl Acad Sci USA. 1989;86(16):6043-6.
54. Sanders SL, Garbett KA, Weil PA. Molecular characterization of *Saccharomyces cerevisiae* TFIID. Mol Cell Biol. 2002;22(16):6000-13.
55. Auble DT, Wang D, Post KW, Hahn S. Molecular analysis of the SNF2/SWI2 protein family member MOT1, an ATP-driven enzyme that dissociates TATA-binding protein from DNA. Mol Cell Biol. 1997;17(8):4842-51.
56. Cade RM, Errede B. MOT2 encodes a negative regulator of gene expression that affects basal expression of pheromone-responsive genes in *Saccharomyces cerevisiae*. Mol Cell Biol. 1994;14(5): 3139-49.
57. Haas H, Angermayr K, Stöffler G. Molecular analysis of a *Penicillium chrysogenum* GATA factor encoding gene (sreP) exhibiting significant homology to the *Ustilago maydis* urbs1 gene. Gene. 1997;184(1):33-7.
58. Krishnamurthy S, He X, Reyes-Reyes M, Moore C, Hampsey M. Ssu72 Is an RNA polymerase II CTD phosphatase. Mol Cell. 2004;14(3):387-94.
59. James P, Whelen S, Hall BD. The RET1 gene of yeast encodes the second-largest subunit of RNA polymerase III. Structural

analysis of the wild-type and ret1-1 mutant alleles. J Biol Chem. 1991;266(9):5616-24.
60. Archambault J, Chambers RS, Kobor MS, Ho Y, Cartier M, Bolotin D, et al. An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA. 1997; 94(26):14300-5.
61. Sweetser D, Nonet M, Young RA. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc Natl Acad Sci USA. 1987; 84(5):1192-6.
62. Albers M, Diment A, Muraru M, Russell CS, Beggs JD. Identification and characterization of Prp45p and Prp46p, essential pre-mRNA splicing factors. RNA. 2003;9(1):138-50.
63. Caspary F, Shevchenko A, Wilm M, Séraphin B. Partial purification of the yeast U2 snRNP reveals a novel yeast pre-mRNA splicing factor required for pre-spliceosome assembly. EMBO J. 1999;18(12):3463-74.
64. Yokoyama C, Wang X, Briggs MR, Admon A, Wu J, Hua X, Goldstein JL, Brown MS. SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. Cell. 1993;75(1):187-97.
65. Tilburn J, Sarkar S, Widdick DA, Espeso EA, Orejas M, Mungroo J, Peñalva MA, Arst HN Jr. The aspergillus pacc zinc finger transcription factor mediates regulation of both acid- and alkaline-expressed genes by ambient pH. EMBO J. 1995;14(4): 779-90.
66. MacPherson S, Larochelle M, Turcotte B. A fungal family of transcriptional regulators: The zinc cluster proteins. Microbiol Mol Biol Rev. 2006;70(3):583-604. Review.
67. Fierro F, Barredo JL, Díez B, Gutierrez S, Fernández FJ, Martín JF. The penicillin gene cluster is amplified in tandem repeats linked by conserved hexanucleotide sequences. Proc Natl Acad Sci USA. 1995; 92(13):6200-4.
68. Newbert RW, Barton B, Greaves P, Harper J, Turner G. Analysis of a commercially improved *Penicillium chrysogenum* strain series: Involvement of recombinogenic regions in amplification and deletion of the penicillin biosynthesis gene cluster. J Ind Microbiol Biotechnol. 1997;19(1):18-27.
69. Brakhage AA. Molecular regulation of beta-lactam biosynthesis in filamentous fungi. Microbiol Mol Biol Rev. 1998 Sep;62(3): 547-85. Review.
70. Tudzynski B. Nitrogen regulation of fungal secondary metabolism in fungi. Front Microbiol. 2014;5:656. DOI:10.3389/fmicb.2014.00656. eCollection. Review.
71. Chen RE, Thorner J. Function and regulation in MAPK signaling pathways: lessons learned from the yeast *Saccharomyces cerevisiae*. Biochim Biophys Acta. 2007;1773(8):1311-40.
72. Román E, Arana DM, Nombela C, Alonso-Monge R, Pla J. MAP kinase pathways as regulators of fungal virulence. Trends Microbiol. 2007;15(4):181-90. Review.
73. Stoll D, Schmidt-Heydt M, Geisen R. Differences in the regulation of ochratoxin A by the HOG pathway in *Penicillium* and *Aspergillus* in response to high osmolar environments. Toxins (Basel). 2013;5(7): 1282-98.
74. Bayram Ö, Bayram ÖS, Ahmed YL, Maruyama J, Valerius O, Rizzoli SO, Ficner R, Irniger S, Braus GH. The *Aspergillus nidulans* MAPK module AnSte11-Ste50-Ste7-Fus3 controls development and secondary metabolism. PLoS Genet. 2012;8(7):e1002816.
75. Park G, Pan S, Borkovich KA. Mitogen-activated protein kinase cascade required for regulation of development and secondary metabolism in *Neurospora crassa*. Eukaryot Cell. 2008;7(12):2113-22.
76. Flaherty JE, Payne GA. Over expression of aflR leads to upregulation of pathway gene transcription and increased aflatoxin production in *Aspergillus flavus*. Appl Environ Microbiol. 1997;63(10):3995-4000.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/11692