International Journal of Food Microbiology 200 (2015) 66–71

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

High sequence variations in the region containing genes encoding a cellular morphogenesis protein and the repressor of sexual development help to reveal origins of *Aspergillus oryzae*

Perng-Kuang Chang [a,*], Leslie L. Scharfenstein [a], Cesar D. Solorzano [b,1], Hamed K. Abbas [b], Sui-Sheng T. Hua [c], Walker A. Jones [b], Robert M. Zablotowicz [d]

[a] Southern Regional Research Center, Agricultural Research Service, U.S. Department of Agriculture, New Orleans, LA 70124, United States
[b] Biological Control of Pests Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Stoneville, MS 38776, United States
[c] Western Regional Research Center, Agricultural Research Service, U.S. Department of Agriculture, Albany, CA, 94710, United States
[d] Crop Production Systems Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Stoneville, MS 38776, United States

ARTICLE INFO

Article history:
Received 6 October 2014
Received in revised form 21 January 2015
Accepted 28 January 2015
Available online 7 February 2015

Keywords:
Aspergillus oryzae
Aspergillus flavus
Deletion
Genome
Recombination

ABSTRACT

Aspergillus oryzae and *Aspergillus flavus* are closely related fungal species. The *A. flavus* morphotype that produces numerous small sclerotia (S strain) and aflatoxin has a unique 1.5 kb deletion in the *norB-cypA* region of the aflatoxin gene cluster (i.e. the S genotype). Phylogenetic studies have indicated that an isolate of the nonaflatoxigenic *A. flavus* with the S genotype is the ancestor of *A. oryzae*. Genome sequence comparison between *A. flavus* NRRL3357, which produces large sclerotia (L strain), and S-strain *A. flavus* 70S identified a region (*samA-rosA*) that was highly variable in the two morphotypes. A third type of *samA-rosA* region was found in *A. oryzae* RIB40. The three *samA-rosA* types were later revealed to be commonly present in *A. flavus* L-strain populations. Of the 182 L-strain *A. flavus* field isolates examined, 46%, 15% and 39% had the *samA-rosA* type of NRRL3357, 70S and RIB40, respectively. The three types also were found in 18 S-strain *A. flavus* isolates with different proportions. For *A. oryzae*, however, the majority (80%) of the 16 strains examined had the RIB40 type and none had the NRRL3357 type. The results suggested that *A. oryzae* strains in the current culture collections were mostly derived from the *samA-rosA*/RIB40 lineage of the nonaflatoxigenic *A. flavus* with the S genotype.

Published by Elsevier B.V.

1. Introduction

Aspergillus oryzae genetically almost identical to *A. flavus* is classified as a separate species because of their morphological differences and economic and food safety concerns. It has been used as the koji starter culture for the production of fermented foods and alcoholic beverages, and as the host for the production of many industrial enzymes used in food processing (Machida et al., 2008). Its long history of safe use by the food industry has earned it GRAS (generally recognized as safe) status. In contrast, *Aspergillus flavus*, a ubiquitous soil saprophyte and a facultative plant pathogen, is frequently found on dead plant tissues and agricultural debris. Through the dissemination of airborne spores, this fungus can infect many crops including corn, cotton, peanuts, and nut trees. Infection by toxigenic *A. flavus* often leads to contamination with aflatoxin B_1, a potent carcinogenic compound which is toxic to animals and humans. Therefore, *A. flavus* is most associated with preharvest aflatoxin contamination of commercial crops. Because of its small spores and its ability to grow at 37 °C, *A. flavus* is also an opportunistic pathogen to immunocompromised people. Infection by *A. flavus* now is the second leading cause of invasive aspergillosis in humans (Krishnan et al., 2009).

A. oryzae strains are believed to be variants of *A. flavus* arising as a distinct population after hundreds of years of domestication under artificial fermentation conditions. They were selected for fast growth and increased activities of amylases and proteases to facilitate the degradation of macromolecules in fermentative substrates (Jorgensen, 2007). Commercial *A. oryzae* strains tend to have floccose aerial mycelia with various amounts of spores and produce few or no sclerotia, which are hardened mycelial masses that serve as over-winter reproductive forms and remain dormant until conditions are favorable for growth. Populations of *A. flavus* exhibit a great deal of phenotypic variability. Besides producing spores, many *A. flavus* field isolates are able to produce sclerotia. Although sclerotial morphology is a poor indicator of phylogeny, sclerotial size has been used to separate *A. flavus* isolates into two main groups called the L strain and the S strain. The majority of *A. flavus* isolates belong to the L strain whose average sclerotial size is greater than 400 μm. A much smaller portion of *A. flavus* isolates belong

* Corresponding author. Tel.: +1 504 286 4208; fax: +1 504 286 4419.
E-mail address: perngkuang.chang@ars.usda.gov (P.-K. Chang).
[1] Current address: Valent U. S. A. Corporation, Leland, Mississippi 38756, United States.

http://dx.doi.org/10.1016/j.ijfoodmicro.2015.01.021
0168-1605/Published by Elsevier B.V.

to the S strain whose sclerotial size is less than 400 μm in diameter (Cotty, 1989). L-strain *A. flavus* isolates vary considerably in their ability to produce aflatoxin and cyclopiazonic acid (CPA), an indole-tetramic acid neurotoxin (Chang et al., 2009). S-strain *A. flavus* isolates are very stable in aflatoxin production, but much less is known about their ability to produce CPA. A significant portion of L-strain *A. flavus* isolates collected from many parts of the world are nonaflatoxigenic (Horn and Dorner, 1999; Mauro et al., 2013; Perrone et al., 2014; Pildain et al., 2004; Wei et al., 2014). In contrast, nonaflatoxigenic S-strain isolates are rarely found (Orum et al., 1997).

A comparative analysis of the chemical profiles of *A. oryzae* and *A. flavus* under 15 growth conditions has revealed that they produce very different secondary metabolites (Rank et al., 2012). No *A. oryzae* isolate has been found to produce aflatoxin although some isolates are able to produce CPA (Chang et al., 2009; Kim et al., 2014). The culture morphology of *A. oryzae* in terms of aerial mycelia and spore production resembles more that of the L-strain *A. flavus* than that of the S-strain *A. flavus*. However, *A. oryzae* strains are phylogenetically closer to *A. flavus* S-strain and some nonaflatoxigenic L-strain isolates than to aflatoxigenic L-strain *A. flavus* isolates based on *omtA* SNPs and a unique deletion in the *norB-cypA* region of the aflatoxin gene cluster (Chang et al., 2006; Ehrlich et al., 2004; Geiser et al., 2000). The S strain is estimated to have diverged from an aflatoxigenic L strain between 1 and 3 million years ago (Ehrlich et al., 2005). Based on these genetic traits *A. flavus* is considered a polyphyletic assemblage of isolates while *A. oryzae* can be considered either a monophyletic or a polyphyletic assemblage (Chang et al., 2006; Geiser et al., 2000). The genetic diversity of *A. flavus* populations likely results either from intra-specific genetic exchange during the parasexual cycle when heterokaryons are formed by the fusion of vegetatively compatible strains or from sexual recombination or from both. These mechanisms have been conclusively demonstrated in laboratory conditions (Horn et al., 2009; Papa, 1973). Based on intactness and deletion size in the aflatoxin gene cluster commercial *A. oryzae* strains have been categorized into three groups (Kusumoto et al., 2000; Tao and Chung, 2014; Tominaga et al., 2006). Identical or similar deletion patterns have been widely reported for nonaflatoxigenic *A. flavus* isolates (Chang et al., 2005; Donner et al., 2010; Mauro et al., 2013; Wei et al., 2014; Yin et al., 2009). Deletions in these partial cluster strains are a recent occurrence and suggest gene loss (Moore et al., 2009).

The genome of *A. oryzae* RIB40 has been determined to be about 37 Mb and organized in eight chromosomes with 12,074 predicted genes (Machida et al., 2005). The genome of *A. flavus* NRRL3357, an L strain, also is about 37 Mb and contains 13,485 genes (Payne et al., 2006). Genome sequences of seven additional strains of *A. flavus* and *A. oryzae* though not yet well assembled are available (Gibbons et al., 2012; Umemura et al., 2012). A comparison of genomes of *A. oryzae* RIB40 and *A. oryzae* 3.042 indicates they are mostly collinear but with several translocations (Zhao et al., 2013). A transcriptomic study of *A. oryzae* RIB40 has identified more than 1,100 novel transcripts with significant expression levels (Wang et al., 2010). These novel *A. oryzae* transcripts likely have counterparts in *A. flavus* since more than 1,100 novel transcript also have been reported for *A. flavus* CA43, an S-strain isolate (Wu et al., 2014). The genome of another S-strain *A. flavus* isolate, 70S (ATCC MYA-384), has been sequenced by the Food and Feed Safety Research Unit at Southern Regional Research Center in collaboration with J. Craig Venter Institute. Detailed comparisons between NRRL3357 and 70S as well as between genomes of RIB40 and 70S have not been performed.

Previous genetic fingerprints, including the *aflR* aflatoxin pathway regulatory gene (Chang et al., 1995; Lee et al., 2006a), for the identification of *A. oryzae* used genes exclusively located in the aflatoxin gene cluster. In the present study, we found, through a comparative analysis of deletions, that genome sequences of NRRL3357, 70S and RIB40 contained a highly variable region that included genes encoding a cellular morphogenesis protein with a sterile alpha motif (*samA*) and the repressor of sexual development (*rosA*). We intended to use these regions to infer origins of *A. oryzae* strains. All three variable regions were present in *A. flavus* isolates collected from several agricultural fields and retail markets. The majority of known *A. oryzae* strains examined had the RIB40 type, and none had the NRRL3357 type. *A. oryzae* apparently was derived mostly from a special lineage of the nonaflatoxigenic S genotype.

2. Materials and methods

2.1. Fungal strains, culture conditions and genomic DNA preparation

A total of 182 *A. flavus* L-strain isolates, composed of both aflatoxigenic and nonaflatoxigenic strains, were collected from corn grown in Indiana, Texas, Louisiana, Alabama, North Carolina and Florida as well as from raw almond nuts purchased at retail stores in the east Bay Area, California (Table S1). Other aspergilli examined included 18 S-strain *A. flavus* isolates from various localities in the United States and 16 *A. oryzae* strains from various culture collections. Cultures for genomic DNA preparation were grown in a 2-mL microfuge tube, which contained 1.0 mL potato dextrose broth (PDB, EMD, Darmstadt, Germany), until they formed a thin layer of mycelial mat. Detailed procedures for DNA isolation using Disruptor Genie™ and the ZR Fungal/Bacterial DNA Kit™ (ZYMO RESEARCH, Orange County, California, USA) were as previously described (Chang et al., 2012).

2.2. Localization of deletions in the genomes of A. flavus 70S and A. oryzae RIB40

The *Aspergillus flavus* NRRL3357 genome database of the Broad Institute (http://www.broadinstitute.org/annotation/genome/aspergillus_group/Regions.html.), a locally built *A. flavus* 70S genome database (http://10.114.143.20:4567), and the *A. oryzae* RIB40 genome database of the National Institution of Technology and Evaluation, Japan (http://www.bio.nite.go.jp/dogan/project/view/AO) were used to compare corresponding regions in *A. flavus* and *A. oryzae*. BLAST searches and standard procedures of sequence alignment and analysis were performed to identify deletions in the genomes of 70S and RIB40.

2.3. Identification of genetic traits in A. flavus and A. oryzae strains using publically available Next-Generation-Sequencing (NGS) data

Sequence templates for sequence alignment and mapping were (i) the *A. parasiticus norB-cypA* region (GenBank accession number: AY371490) for classifying the sclerotial genotypes, S and L, based on unique deletion patterns, (ii) the *A. flavus* mating-type genes as a concatenated sequence containing *MAT1-1* (EU357934) and *MAT1-2* (EU357936), and (iii) the *samA-rosA* type as a concatenated sequence representing unique regions, each about 1.0 kb downstream of the breakpoint junction around the *rosA* region, from *A. flavus* NRRL3357 and 70S, and *A. oryzae* RIB40 (See Fig. 1). Raw Illumina genomic sequence reads were retrieved from NCBI Raw Sequence Archive (SRA, http://www.ncbi.nlm.nih.gov/sra). The accession number for the six *A. oryzae* strains of RIB331, RIB333, RIB537, RIB632, RIB642 and RIB949 is SRA052658. The accession number for the seven *A. flavus* strains of SRRC1273, SRRC1357, SRRC2112, SRRC2114, SRRC2524, SRRC2632, and SRRC2653 is SRA052664. The accession number for *A. flavus* NRRL1957 (ATCC16883, type strain) is SRX396791. The module of CLC Genomic Workbench (http://www.slsbio.com) — CLCbio > NGS Core Tools > Map Reads to Reference was used. BLAST search and manual sequence analysis on assembled genome sequences (contigs) of *A. oryzae* RIB326 (Umemura et al., 2012) and strain 3.042, a soy-sauce producing strain used in China (Zhao et al., 2013), were also performed to identify these genetic traits.

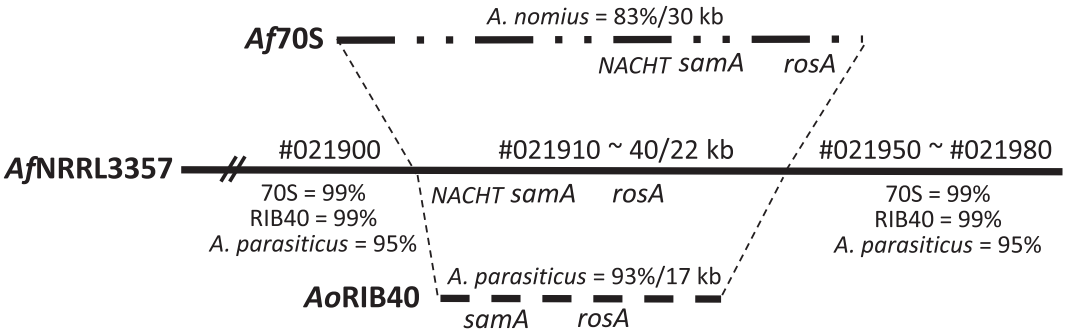

Fig. 1. Schematic representation of conserved regions and the highly variable region that contains the *samA* and *rosA* genes in *A. flavus* and *A. oryzae*. The *NACHT* gene encodes a nucleoside-triphosphatase domain, commonly found in animal, fungal and bacterial proteins. The *samA* gene encodes a cellular morphogenesis protein, which may play a role in the mitogen-activated protein kinase cascades, and the *rosA* gene encodes a Zn $(II)_2Cys_6$ DNA-binding protein and functions as a repressor of sexual development in *A. nidulans*. The relative locations of these genes in the genome of 70S, NRRL3357 and RIB40 are shown. Detailed sequence analyses and comparison are shown in Table S2 and Table S3. A summary of annotated genes in these highly variable regions is shown in Table S4.

2.4. PCR fingerprinting analysis

Oligonucleotide primers used to PCR fingerprint mating-type genes were derived from *A. flavus MAT1-1* and *MAT1-2*. The primers were MAT1-1F: 5′-ATGGAAACCACAGTGTCTCC and MAT1-1R: 5′-TCAACGAA TCTAGAGAAGTC, and MAT1-2F: 5′-ATCGAGAATGACGACTATAC and MAT1-2R: 5′-TTCTTCAGTAGCAGTCAGCA. For the identification of deletion patterns in the *norB-cypA* region the primers 5′-GTGCCCAGCATCTT GGTCCA and 5′-AGGACTTGATGATTCCTCGTC (Chang et al., 2005) were used. For distinguishing the three types of the *samA-rosA* region primers of 3357/70S: 5′-CACCTCTTGAGTGGCTACTA and 3357rosA: 5′-CATGTC GCGGGTTGGAATCAC were used to determine the NRRL3357 type; primers of 3357/70S and 70SrosA: 5′-ATCCCAAATCGTAGCTGAGC were used to determine the 70S type; and primers of rosAoryF: 5′-CCAGCAGTTACTTAGACATGATA and rosAoryR: 5′-CCACTCCTCAGCTG TCTCAGACT were used to determine the RIB40 type. PCR was performed using Plantinum® Blue PCR Supermix (Invitrogen, Carlsbad, California, USA) under the following conditions: the reaction mix was heated at 94 °C for 5 min and then subjected to 30 cycles consisting of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 2 min.

3. Results

3.1. Deletions in A. flavus 70S genome in reference to A. flavus NRRL3357 genome

In the previous transcriptomic study of an S-strain *A. flavus* isolate, CA42, about 2.6% (353 of 13,485) of the *A. flavus* NRRL3357 reference genes were not expressed across all samples (Chang et al., 2014). The data show that a majority of the non-expressing genes are located in small, continuous regions, which suggests possible deletions of these genes in the CA42 genome. An analysis of the only assembled S-strain *A. flavus* 70S genome sequence, considered to be equivalent to that of

Table 1
Deletions in genomes of *A. flavus* 70S and *A. oryzae* RIB40 in relation to *A. flavus* NRRL3357.

Non-expressing genes in *A. flavus* CA42[a]	Corresponding region in *A. flavus* NRRL3357[b]	Deletion in *A. flavus* AF70[c]	Deletion in *A. oryzae* RIB40[d]
AFLA_000010-000330	Contig 11: 1-10170	10.2 kb	YES
AFLA_000350-000510	Contig 11: 18869-32333	No	13.5 kb
AFLA_001120-001160	Contig 11: 180412-190907	10.5 kb	No
AFLA_005140-005270	Contig 11: 1344532-1391479	47.0 kb	YES
AFLA_005320-005350 Clus#3a	Contig 11: 1402890-1421012	10.5 kb	~16.5 kb
AFLA_009660-009710	Contig 15: 598561-618549	~16.9 kb	~12.0 kb
AFLA_010320-010350	Contig 15: 766903-774433	~7.5 kb	No
AFLA_021910-021940	Contig 03: 2478334-2500060	30.2 kb insert, low homology	17.2 kb, low homology
AFLA_040480-040670	Contig 05: 622130-685973	63.8 kb	YES
AFLA_040850-040870	Contig 05: 726572-734699	8.1 kb	No
AFLA_042610-042630	Contig 05: 1161765-1169415	~7.0 kb	No
AFLA_052730-052750	Contig 12: 1388966-1397379	8.4 kb	No
AFLA_062940-063020 Clus#20	Contig 06: 2195483-2214455	19.0 kb	YES
AFLA_063370-063410	Contig 06: 2292392-2304310	11.9 kb	No
AFLA_072510-072550	Contig 01: 97482-111669	13.5 kb	YES
AFLA_076170-076210	Contig 01: 1032946-1044974	12.0 kb	No
AFLA_077110-077220	Contig 01: 1300369-1332866	32.5 kb	YES
AFLA_078160-078180	Contig 01: 1555222-1557581	2.4 kb	No
AFLA_104590-104610	Contig 04: 121646-129044	~2.5 kb	No
AFLA_105420-105440 Clus#38	Contig 04: 331182-341939	10.8 kb	YES
AFLA_107900-107950	Contig 04: 978199-997558	19.4 kb	No
AFLA_110190-110210	Contig 04: 1606435-1619150	12.7 kb	No
AFLA_135420-135490 Clus#53	Contig 07: 1194215-1214553	20.3 kb	No
AFLA_138490-138510	Contig 07: 2028903-2036185	7.3 kb	No

Clus#: predicted secondary metabolite gene clusters.
[a] Gene IDs are designations of Kyoto Encyclopedia of Genes and Genomes (KEGG); see http://www.genome.jp/kegg/.
[b] *Aspergillus flavus* contigs are from the database of the Broad Institute; see http://www.broadinstitute.org/annotation/genome/aspergillus_group/Regions.html.
[c] Locally built *A. flavus* 70S genome sequence database was used; see http://10.114.143.20:4567.
[d] The *A. oryzae* genome sequence database used is maintained by the National Institution of Technology and Evaluation, Japan; see http://www.bio.nite.go.jp/dogan/project/view/AO.

CA42, revealed that three-fourths (268/353) of the NRRL3357 genes have no counterparts in 70S. Table 1 summarizes identified deletions in the 70S genome. The deletions occurred mostly in a few contigs, and the sizes ranged from a few kb to over 60 kb. One deletion each was found at the secondary metabolite gene clusters 3a, 20, 38, and 53. The corresponding regions in the *A. oryzae* RIB40 genome also were analyzed and showed less variability than 70S.

3.2. Identification of a highly variable region in A. flavus NRRL3357 and 70S, and A. oryzae RIB40

Previous studies suggested that *A. oryzae* is phylogenetically closer to S-strain *A. flavus* than to aflatoxigenic L-strain *A. flavus*. However, deletions found varied greatly in the genomes of 70S and RIB40 (Table 1). Thirteen deleted regions detected in the former were intact in the latter and the overall deletion pattern in RIB40 resembled that of NRRL3357 and not that of 70S. Strikingly, a detailed analysis of the presumably 22-kb deletion region corresponding to AFLA_021910 to AFLA_021940 in the 70S genome turned out to be a unique 30-kb sequence (Contig 111: 17685 to 47915) that shared no homology to the NRRL3357 sequence but had 83% identity to *A. nomius* sequence (Fig. 1; see Tables S2 and S3 for detailed analyses). In the corresponding location in the RIB40 genome, a 17-kb sequence (Supercontig 12: 2435815 to 2452851) that shared no homology to the NRRL3357 sequence but had 93% identity to *A. parasiticus* sequence was present instead. These highly variable regions in 70S and RIB40 were encompassed by sequences highly homologous (99% identity) to the NRRL3357 sequences (Figs. 1 and S1, Table S2). The 22-kb region of NRRL3357 contains four genes including those that encode a NACHT domain protein (AFLA_021910), a cellular morphogenesis protein with a sterile alpha motif (SAM) (AFLA_021920), RosA [repressor of sexual development (AFLA_021930)] and a conserved protein (Table S4). Despite the lack of DNA sequence homology, the 30-kb sequence of 70S contained the aforementioned three genes and additional four genes including one encoding a sugar transporter of the major facilitator superfamily (MFS, Table S4). In contrast to NRRL3357 and 70S, the NACHT-domain gene in the 17-kb sequence of RIB40 was deleted, which contains only three genes (Fig. 1 and Table S4).

3.3. Distribution of unique types of samA-rosA among A. flavus and A. oryzae isolates

Isolates of *A. flavus* have been grouped into S and L sclerotial genotypes based on distinct deletion patterns in the *norB-cypA* region. S genotype corresponds to a 1.5 kb deletion, and L genotype corresponds to a 1.0 kb deletion. Morphologically, *A. oryzae* strains resemble L-strain *A. flavus* isolates more than they resemble S-strain isolates. To correlate the morphological similarity with a possible genetic basis, we investigate the frequency of the three *samA-rosA* types along with genetic features, such as mating type and sclerotial genotype in *A. flavus* populations (Table 2). We found that among the 182 L-strain isolates surveyed by the PCR fingerprinting (Fig. 2) the *samA-rosA*/NRRL3357 type was the most prevalent (46%), the *samA-rosA*/70S type was the least frequently found (15%), and the *samA-rosA*/RIB40 type accounted for the remaining L-strain isolates (39%)

We also examined the aforementioned genetic traits of 16 *A. oryzae* strains and 18 aflatoxigenic S-strain *A. flavus* isolates (Table 3). The distribution of mating-type genes in the *A. oryzae* and S-strain *A. flavus* isolates was comparable (The ratio of mating-type 1–1 to mating type 1–2 was approximately 5:4) to that of L-strain *A. flavus* isolates. For the *samA-rosA* region, the majority (80%) of the *A. oryzae* strains had the RIB40 type, and no strain had the NRRL3357 type. In contrast, like the L-strain *A. flavus* isolates, the three *samA-rosA* types were commonly present in the S-strain *A. flavus* isolates, with a ratio of 5:2:11 for NRRL3335:70S:RIB40. The other genome sequence analyzed *A. flavus* isolates, which included SRRC1273, SRRC1357, SRRC2112, SRRC2114, SRRC2524, SRRC2632, SRRC2653 and NRRL1957 (ATCC16883, type strain), also contained the three types (data not shown).

Table 2
Distribution of genetic traits among L-strain *A. flavus* isolates.[a]

Mating type	*norB-cypA* deletion[b]	*samA-rosA* region[c]	Total
MAT1-1 (100)	S (31)	NRRL3357	12
		70S	8
		RIB40	11
	L (69)	NRRL3357	36
		70S	8
		RIB40	25
MAT1-2 (82)	S (53)	NRRL3357	28
		70S	7
		RIB40	18
	L (29)	NRRL3357	8
		70S	4
		RIB40	17

[a] The isolates consist of aflatoxigenic and non-aflatoxigenic isolates.
[b] The isolates were grouped into S and L sclerotial genotypes based on distinct deletion patterns in the *norB-cypA* region. S genotype corresponds to a 1.5 kb deletion, and L genotype corresponds to a 1.0 kb deletion.
[c] The total numbers for the type of NRRL3357, 70S and RIB40 are 84, 27 and 71, respectively.

4. Discussion

A. oryzae is believed to be a domesticated ecotype of *A. flavus* because genome sequences of RIB40 and NRRL3357 share greater than 99% identity (Gibbons and Rokas, 2013). Studies have suggested that *A. oryzae* is derived from a clade of nonaflatoxigenic L-strain *A. flavus* containing the *norB-cypA* S-genotype and *omtA* SNPs that are identical to those of the S strain (Chang et al., 2006; Geiser et al., 2000). This precursor clade to *A. oryzae* thus most likely shared the same ancestor as the S strain but subsequently lost its aflatoxin producing ability. It seems that this clade had not undergone extensive deletion since regions corresponding to deletions found in 70S are mostly intact in *A. oryzae* RIB40 (Table 1). Two forces in evolution may have operated independently and shaped the genomes of the S- and the nonaflatoxigenic L-strain *A. flavus*: (i) spontaneous transitions and transversions and (ii) genetic drift (gene deletion) due to a change in population size or for adaptation to living niches. The less frequently encountered S-strain *A. flavus* is estimated to have diverged from an aflatoxigenic L-strain *A. flavus* like NRRL3357 between 1 and 3 million years ago (Ehrlich et al., 2005). The aforementioned nonaflatoxigenic L-strain clade including *A. oryzae* must have evolved after the divergence and was subjected to niche selection pressures similar to those on NRRL3357. These selection pressures, however, must be quite different from those on the S strain based on the deleted regions found in their genomes (Table 1).

The overall sequence homology in the aflatoxin gene clusters of *A. flavus* AF13, an L-strain isolate, and 70S is greater than 99% (Ehrlich et al., 2005). Homology of these *A. flavus* clusters to that of *A. parasiticus* is 96% and to that of *A. nomius* is 91%, which supports that *A. flavus* is

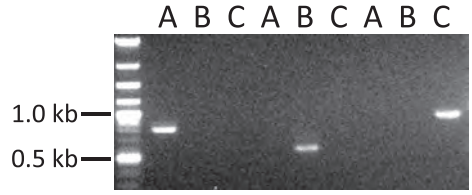

Fig. 2. PCR fingerprinting products correspond to the three types of the *samA-rosA* region in the *A. flavus* populations. Sequence-specific primers at each side of a sequence breakpoint (Fig. S1) were used in PCR (see Materials and methods). A: the NRRL3357 type, which yields an 838-bp fragment; B: the 70S type, which yields a 575-bp fragment, and C: the RIB40 type, which yields a 933-bp fragment.

Table 3
Comparison of genetic traits between *A. oryzae* and aflatoxigenic S-stain *A. flavus*.[a]

A. oryzae	Group[b]	Mating type	norB-cypA deletion	samA-rosA region[c]
RIB40	I	1-1	S	RIB40
RIB326	III	1-2	Δ[d]	RIB40
RIB331	I	1-1	S	RIB40
RIB333	I	1-2	S	RIB40
RIB537	II	1-1	Δ	RIB40
RIB632	II	1-1	Δ	RIB40
RIB642	I	1-1	S	RIB40
RIB949	II	1-2	Δ	70S
SRRC302	I	1-1	S	RIB40
SRRC304	I	1-1	S	RIB40
SRRC493	I	1-1	S	RIB40
SRRC2044	I	1-2	S	70S
SRRC2098	II	1-1	Δ	70S
SRRC2103	II	1-2	Δ	RIB40
3.042	I	1-2	S	RIB40
FGSC A815	II	1-2	Δ	RIB40
A. flavus				
AF12S		1-2	S	RIB40
AF70S		1-1	S	70S
CA28		1-2	S	70S
CA42		1-2	S	RIB40
CA43		1-1	S	RIB40
CA44		1-1	S	NRRL3357
GA10-18s		1-1	S	NRRL3357
IN173		1-2	S	NRRL3357
NC21		1-2	S	RIB40
NC22		1-1	S	RIB40
NC69		1-2	S	RIB40
NC80		1-1	S	RIB40
NC113		1-1	S	NRRL3357
NC128		1-1	S	NRRL3357
NC145		1-2	S	RIB40
NC152		1-2	S	RIB40
NC165		1-1	S	RIB40
VA4-36s		1-1	S	RIB40

[a] The mating type genes, *norB-cypA*, *sam-rosA* are located on chromosome 6, chromosome 3 and chromosome 4, respectively.
[b] The group numbers for the RIB strains were obtained from the Table S1 (Tominaga et al., 2006) and some confirmed by genome sequence analysis, for SRRC302 and 3.042 from genome sequence analysis, for other SRRC strains from a previous study (Chang et al., 2006), and for FGSC A815 from a PCR fingerprinting analysis with sequence-specific primers.
[c] Except for the types of the RIB strains, SRRC302 and 3.042, which were determined by genome sequence analysis, all others were determined by PCR.
[d] The *norB-cypA* region is deleted in group II and III *A. oryzae* strains.

more distantly related to *A. nomius* than to *A. parasiticus*. The estimated divergence of *A. nomius* from *A. parasiticus* is between 25 and 55 Ma and *A. parasiticus* from *A. flavus* between 8 and 17 Ma. In this study, the finding in 70S that its *samA-rosA* region is only homologous to that of *A. nomius* and not to that of NRRL3357 suggests a more complex evolutionary history than previously thought. Similarly, the *samA-rosA* region of RIB40 is homologous only to that of *A. parasiticus* and not to that of *A. flavus* NRRL3357. Although the genomes of L-strain *A. flavus* and *A. oryzae* are collinear, certain region(s) of *A. oryzae* might originate from other species, such as *A. parasiticus*. In both cases early active sexual recombination events among closely related groups and nonlinear nucleotide substitution in the course of evolution may have led to current species classified. Alternatively, assuming linearity of nucleotide substitution, recombination between ancestors of L-strain *A. flavus* and *A. nomius* that gave rise to the S-strain *A. flavus* may have occurred long before recombination between ancestors of L-strain *A. flavus* and *A. parasiticus* that gives rise to *A. oryzae*; for 70S the sequence identity in the *samA-rosA* region compared to that of NRRL3357 is only 83% but for RIB40 it is 96% (Fig. 1).

Implication of recombination in the L- and the S-strain populations examined in this study is based on loci that span three chromosomes. The mating-type (*MAT*) genes, *norB-cypA*, and *samA-rosA* are located on chromosomes 6, 3, and 4, respectively. All possible combinations of these loci are present in both populations, which is in accord with there being an extensive history of recombination (Carbone et al., 2007; Moore et al., 2009). The results also support the conclusion of sexual recombination in *A. flavus* populations although most previous evidence came from the analysis of aflatoxin gene clusters after mating under laboratory conditions (Moore et al., 2013; Olarte et al., 2012). In contrast, despite the occurrence of both mating-type genes in the *A. oryzae* strains, these strains exclusively have the S (1.5 kb deletion in the *norB-cypA* region) genotype. In addition, most of them have the *samA-rosA*/RIB40 type but none have the *samA-rosA*/NRRL3357 type. These results suggest that current A. *oryzae* strains belong to and are selected from certain lineage(s) of *A. flavus*. *A. oryzae* strains have been categorized into three groups. Group I has an intact aflatoxin gene cluster, but groups II and III each has a large deletion that results in a loss of more than half of the gene cluster (Kusumoto et al., 2000). Among the 122 group I strains deposited at the National Research Institute of Brewing (NRIB), 98 have the S genotype, 5 have the L genotype (1.0 kb deletion) and 19 have no deletion (Tominaga et al., 2006). The 19 non-deletion strains have been reclassified as *A. sojae* because these strains like *A. parasiticus* have the intact *norB-cypA* region but do not produce aflatoxins (Sato et al., 2011). Therefore, 95% of the group I strains have the S genotype.

In the NRIB *A. oryzae* culture collection 77 of 210 are group II strains. Of the six group II RIB strains and the FGSC strain A815 we examined, two have the *samA-rosA*/70S type (Table 3). One of the two, RIB949, is phylogentically distant from other seven RIB strains based on whole-genome SNPs (Gibbons et al., 2012). Therefore, group II represented by RIB62 (Lee et al., 2006b) may have been derived from another lineage. *A. flavus* isolates with the identical translocation in the deleted aflatoxin gene cluster region found in RIB62 have been identified including TX13-5 and TX21-5 (Chang et al., 2005), AF051 (Jiang et al., 2009), V5F-13 (Chen et al., 2011), CA12 and CA32 (Hua et al., 2012), and M34 and M65 (unpublished results for the latter five isolates). Group III strains account for only about 4% of the NRIB collection. A genome sequence comparison between RIB326 of group III and RIB40 indicates extensive deletions (3.3%) in the genome of the former (Umemura et al., 2012). Nonetheless, both have the *samA-rosA*/RIB40 type. Therefore, with the exception of a limited number of *A. oryzae* strains having the *samA-rosA*/70S type, the RIB40 type is the representative type of *A. oryzae* strains. Consistently, studies based on other genetic fingerprints, such as *omtA* SNPs and *norB-cypA* deletion also indicate that two phylogenetically different lineages exist in *A. oryzae* strains (Chang et al., 2006; Ehrlich et al., 2004).

Our findings that *A. oryzae* strains contain either mating-type gene, belong exclusively to the S genotype, and most have the *samA-rosA*/RIB40 type but none have the *samA-rosA*/NRRL3357 type suggest that selection of *A. oryzae* occurred long after cessation of sexual recombination. Although it is highly speculative, it is possible that the genetic fingerprints identified so far for *A. oryzae* strains may be coincidentally associated with certain original *A. flavus* group(s) having biological traits that are desirable for fermentation. Evidence of niche specialization for *A. flavus* groups has been obtained (Sweany et al., 2011). The original isolation of *Aspergillus* strains from "ear of rice" may contribute to the selection of current *A. oryzae* strains (Machida et al., 2008). In summary, variations in the *samA-rosA* region can serve as another molecular criterion for identifying and categorizing the relationships among *A. oryzae* isolates when there is a question about an isolate's relationship to *A. flavus* such as in the case of the so called *Aspergillus flavus*/*oryzae* complex (Hong et al., 2013; Lee et al., 2014).

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijfoodmicro.2015.01.021.

References

Carbone, I., Jakobek, J.L., Ramirez-Prado, J.H., Horn, B.W., 2007. Recombination, balancing selection and adaptive evolution in the aflatoxin gene cluster of *Aspergillus parasiticus*. Mol. Ecol. 16, 4401–4417.

Chang, P.-K., Abbas, H.K., Weaver, M.A., Ehrlich, K.C., Scharfenstein, L.L., Cotty, P.J., 2012. Identification of genetic defects in the atoxigenic biocontrol strain *Aspergillus flavus* K49 reveals the presence of a competitive recombinant group in field populations. Int. J. Food Microbiol. 154, 192–196.

Chang, P.-K., Bhatnagar, D., Cleveland, T.E., Bennett, J.W., 1995. Sequence variability in homologs of the aflatoxin pathway gene *aflR* distinguishes species in *Aspergillus* section *Flavi*. Appl. Environ. Microbiol. 61, 40–43.

Chang, P.-K., Ehrlich, K.C., Fujii, I., 2009. Cyclopiazonic acid biosynthesis of *Aspergillus flavus* and *Aspergillus oryzae*. Toxins (Basel) 1, 74–99.

Chang, P.-K., Ehrlich, K.C., Hua, S.S., 2006. Cladal relatedness among *Aspergillus oryzae* isolates and *Aspergillus flavus* S and L morphotype isolates. Int. J. Food Microbiol. 108, 172–177.

Chang, P.-K., Horn, B.W., Dorner, J.W., 2005. Sequence breakpoints in the aflatoxin biosynthesis gene cluster and flanking regions in nonaflatoxigenic *Aspergillus flavus* isolates. Fungal Genet. Biol. 42, 914–923.

Chang, P.-K., Scharfenstein, L.L., Mack, B., Yu, J., Ehrlich, K.C., 2014. Transcriptomic profiles of *Aspergillus flavus* CA42, a strain that produces small sclerotia, by decanal treatment and after recovery. Fungal Genet. Biol. 68, 39–47.

Chen, Y.T., Lin, M.J., Yang, C.H., Ko, W.H., 2011. Characterization of a fungistatic substance produced by *Aspergillus flavus* isolated from soil and its significance in nature. New Biotechnol. 28, 679–683.

Cotty, P.J., 1989. Virulence and cultural characteristics of two *Aspergillus flavus* strains pathogenic on cotton. Phytopathology 79, 808–814.

Donner, M., Atehnkeng, J., Sikora, R.A., Bandyopadhyay, R., Cotty, P.J., 2010. Molecular characterization of atoxigenic strains for biological control of aflatoxins in Nigeria. Food Addit. Contam. 27, 576–590.

Ehrlich, K.C., Chang, P.-K., Yu, J., Cotty, P.J., 2004. Aflatoxin biosynthesis cluster gene *cypA* is required for G aflatoxin formation. Appl. Environ. Microbiol. 70, 6518–6524.

Ehrlich, K.C., Yu, J., Cotty, P.J., 2005. Aflatoxin biosynthesis gene clusters and flanking regions. J. Appl. Microbiol. 99, 518–527.

Geiser, D.M., Dorner, J.W., Horn, B.W., Taylor, J.W., 2000. The phylogenetics of mycotoxin and sclerotium production in *Aspergillus flavus* and *Aspergillus oryzae*. Fungal Genet. Biol. 31, 169–179.

Gibbons, J.G., Rokas, A., 2013. The function and evolution of the *Aspergillus* genome. Trends Microbiol. 21, 14–22.

Gibbons, J.G., Salichos, L., Slot, J.C., Rinker, D.C., McGary, K.L., King, J.G., Klich, M.A., Tabb, D.L., McDonald, W.H., Rokas, A., 2012. The evolutionary imprint of domestication on genome variation and function of the filamentous fungus *Aspergillus oryzae*. Curr. Biol. 22, 1403–1409.

Hong, S.B., Lee, M., Kim, D.H., Chung, S.H., Shin, H.D., Samson, R.A., 2013. The proportion of non-aflatoxigenic strains of the *Aspergillus flavus/oryzae* complex from meju by analyses of the aflatoxin biosynthetic genes. J. Microbiol. 51, 766–772.

Horn, B.W., Dorner, J.W., 1999. Regional differences in production of aflatoxin B_1 and cyclopiazonic acid by soil isolates of *Aspergillus flavus* along a transect within the United States. Appl. Environ. Microbiol. 65, 1444–1449.

Horn, B.W., Moore, G.G., Carbone, I., 2009. Sexual reproduction in *Aspergillus flavus*. Mycologia 101, 423–429.

Hua, S.S., McAlpin, C.E., Chang, P.-K., Sarreal, S.B., 2012. Characterization of aflatoxigenic and non-aflatoxigenic *Aspergillus flavus* isolates from pistachio. Mycotoxin Research 28, 67–75.

Jiang, J., Yan, L., Ma, Z., 2009. Molecular characterization of an atoxigenic *Aspergillus flavus* strain AF051. Appl. Microbiol. Biotechnol. 83, 501–505.

Jorgensen, T.R., 2007. Identification and toxigenic potential of the industrially important fungi, *Aspergillus oryzae* and *Aspergillus sojae*. J. Food Prot. 70, 2916–2934.

Kim, N.Y., Lee, J.H., Lee, I., Ji, G.E., 2014. An evaluation of aflatoxin and cyclopiazonic acid production in *Aspergillus oryzae*. J. Food Prot. 77, 1010–1016.

Krishnan, S., Manavathu, E.K., Chandrasekar, P.H., 2009. *Aspergillus flavus*: an emerging non-fumigatus *Aspergillus* species of significance. Mycoses 52, 206–222.

Kusumoto, K., Nogata, Y., Ohta, H., 2000. Directed deletions in the aflatoxin biosynthesis gene homolog cluster of *Aspergillus oryzae*. Curr. Genet. 37, 104–111.

Lee, C.Z., Liou, G.Y., Yuan, G.F., 2006b. Comparison of the *aflR* gene sequences of strains in *Aspergillus* section *Flavi*. Microbiology 152, 161–170.

Lee, J.H., Jo, E.H., Hong, E.J., Kim, K.M., Lee, I., 2014. Safety evaluation of filamentous fungi isolated from industrial Doenjang koji. J. Microbiol. Biotechnol. 24.

Lee, Y.H., Tominaga, M., Hayashi, R., Sakamoto, K., Yamada, O., Akita, O., 2006a. *Aspergillus oryzae* strains with a large deletion of the aflatoxin biosynthetic homologous gene cluster differentiated by chromosomal breakage. Appl. Microbiol. Biotechnol. 72, 339–345.

Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., Kusumoto, K., Arima, T., Akita, O., Kashiwagi, Y., et al., 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438, 1157–1161.

Machida, M., Yamada, O., Gomi, K., 2008. Genomics of *Aspergillus oryzae*: learning from the history of koji mold and exploration of its future. DNA Res. 15, 173–183.

Mauro, A., Battilani, P., Callicott, K.A., Giorni, P., Pietri, A., Cotty, P.J., 2013. Structure of an *Aspergillus flavus* population from maize kernels in northern Italy. Int. J. Food Microbiol. 162, 1–7.

Moore, G.G., Elliott, J.L., Singh, R., Horn, B.W., Dorner, J.W., Stone, E.A., Chulze, S.N., Barros, G.G., Naik, M.K., Wright, G.C., et al., 2013. Sexuality generates diversity in the aflatoxin gene cluster: evidence on a global scale. PLoS Pathog. 9, e1003574.

Moore, G.G., Singh, R., Horn, B.W., Carbone, I., 2009. Recombination and lineage-specific gene loss in the aflatoxin gene cluster of *Aspergillus flavus*. Mol. Ecol. 18, 4870–4887.

Olarte, R.A., Horn, B.W., Dorner, J.W., Monacell, J.T., Singh, R., Stone, E.A., Carbone, I., 2012. Effect of sexual recombination on population diversity in aflatoxin production by *Aspergillus flavus* and evidence for cryptic heterokaryosis. Mol. Ecol. 21, 1453–1476.

Orum, T.V., Bigelow, D.M., Nelson, M.R., Howell, D.R., Cotty, P.J., 1997. Spatial and temporal patterns of *Aspergillus flavus* strain composition and propagule density in Yuma County, Arizona soils. Plant Dis. 81, 911–916.

Papa, K.E., 1973. The parasexual cycle in *Aspergillus flavus*. Mycologia 65, 1201–1205.

Payne, G.A., Nierman, W.C., Wortman, J.R., Pritchard, B.L., Brown, D., Dean, R.A., Bhatnagar, D., Cleveland, T.E., Machida, M., Yu, J., 2006. Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. Med. Mycol. 44 (Suppl.), 9–11.

Perrone, G., Haidukowski, M., Stea, G., Epifani, F., Bandyopadhyay, R., Leslie, J.F., Logrieco, A., 2014. Population structure and aflatoxin production by *Aspergillus* Sect. *Flavi* from maize in Nigeria and Ghana. Food Microbiol. 41, 52–59.

Pildain, M.B., Vaamonde, G., Cabral, D., 2004. Analysis of population structure of *Aspergillus flavus* from peanut based on vegetative compatibility, geographic origin, mycotoxin and sclerotia production. Int. J. Food Microbiol. 93, 31–40.

Rank, C., Klejnstrup, M.L., Petersen, L.M., Kildgaard, S., Frisvad, J.C., Held Gotfredsen, C., Ostenfeld Larsen, T., 2012. Comparative chemistry of *Aspergillus oryzae* (RIB40) and *A. flavus* (NRRL 3357). Metabolites 2, 39–56.

Sato, A., Oshima, K., Noguchi, H., Ogawa, M., Takahashi, T., Oguma, T., Koyama, Y., Itoh, T., Hattori, M., Hanya, Y., 2011. Draft genome sequencing and comparative analysis of *Aspergillus sojae* NBRC4239. DNA Res. 18, 165–176.

Sweany, R.R., Damann Jr., K.E., Kaller, M.D., 2011. Comparison of soil and corn kernel *Aspergillus flavus* populations: evidence for niche specialization. Phytopathology 101, 952–959.

Tao, L., Chung, S.H., 2014. Non-aflatoxigenicity of commercial *Asepergillus oryzae* strains due to genetic defects compared to aflatoxigenic *Aspergillus flavus*. J. Microbiol. Biotechnol. 24, 1081–1087.

Tominaga, M., Lee, Y.H., Hayashi, R., Suzuki, Y., Yamada, O., Sakamoto, K., Gotoh, K., Akita, O., 2006. Molecular analysis of an inactive aflatoxin biosynthesis gene cluster in *Aspergillus oryzae* RIB strains. Appl. Environ. Microbiol. 72, 484–490.

Umemura, M., Koike, H., Yamane, N., Koyama, Y., Satou, Y., Kikuzato, I., Teruya, M., Tsukahara, M., Imada, Y., Wachi, Y., et al., 2012. Comparative genome analysis between *Aspergillus oryzae* strains reveals close relationship between sites of mutation localization and regions of highly divergent genes among *Aspergillus* species. DNA Res. 19, 375–382.

Wang, B., Guo, G., Wang, C., Lin, Y., Wang, X., Zhao, M., Guo, Y., He, M., Zhang, Y., Pan, L., 2010. Survey of the transcriptome of *Aspergillus oryzae* via massively parallel mRNA sequencing. Nucleic Acids Res. 38, 5075–5087.

Wei, D., Zhou, L., Selvaraj, J.N., Zhang, C., Xing, F., Zhao, Y., Wang, Y., Liu, Y., 2014. Molecular characterization of atoxigenic *Aspergillus flavus* isolates collected in China. J. Microbiol. 52, 559–565.

Wu, X., Zhou, B., Yin, C., Guo, Y., Lin, Y., Pan, L., Wang, B., 2014. Characterization of natural antisense transcript, sclerotia development and secondary metabolism by strand-specific RNA sequencing of *Aspergillus flavus*. PLoS One 9, e97814.

Yin, Y., Lou, T., Yan, L., Michailides, T.J., Ma, Z., 2009. Molecular characterization of toxigenic and atoxigenic *Aspergillus flavus* isolates, collected from peanut fields in China. J. Appl. Microbiol. 107, 1857–1865.

Zhao, G., Yao, Y., Wang, C., Hou, L., Cao, X., 2013. Comparative genomic analysis of *Aspergillus oryzae* strains 3.042 and RIB40 for soy sauce fermentation. Int. J. Food Microbiol. 164, 148–154.