Fungal Genetics and Biology 46 (2009) S45–S52

Contents lists available at ScienceDirect

Fungal Genetics and Biology

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

A survey of nonribosomal peptide synthetase (NRPS) genes in *Aspergillus nidulans*

Hans von Döhren *

Technical University Berlin, Institute of Chemistry OE2, Franklinstr. 29, 10587 Berlin, Germany

ARTICLE INFO

Article history:
Received 28 May 2008
Accepted 27 August 2008
Available online 3 September 2008

Keywords:
Nonribosomal peptide synthetase
Polyketide synthase
Secondary metabolite
Gene cluster

ABSTRACT

The genome of *Aspergillus nidulans* carries 27 genes encoding nonribosomal peptide synthetase (NRPS) structures, although only five of these forming peptides and amino acid containing metabolites have been identified so far. This manuscript describes domain structures, substrate binding pockets and related genes and gene clusters and summarizes our current state of product prediction of fungal NRPS systems.

1. Introduction

Fungal genomics have advanced tremendously during the past years, major aims being their use as cell factories for protein and metabolite production, and their impact as pathogens in both food production and human health. Presumably most interactions of fungi with other microbes, plants and animals are significantly influenced or controlled by secondary metabolites including toxins like aflatoxins but also compounds which have found medical applications such as penicillins or cyclosporins (Miesik and Hoffmeister, 2008). The significant commercial value of such metabolites suggests the term *genomic mining* for the search for novel compounds within the now available sequence data (Bok et al., 2006).

Genome sequencing efforts on filamentous fungi have revealed an unexpectedly large number of secondary metabolite genes and gene clusters, especially polyketide synthases (PKS) and nonribosomal peptide synthetases (NRPS) (Galagan et al., 2005; Machida et al., 2005). First surveys have been conducted for NRPS genes of *Cochliobolus heterostrophus* (Lee et al., 2005), three Fusaria species (Tobiasen et al., 2007) and *A. fumigatus* (Stack et al., 2007), and PKS genes of *Gibberella zeae* (Gaffoor et al., 2005), as well as secondary metabolite clusters of *A. fumigatus* (Nierman et al., 2005), *A. niger* (Pel et al., 2007) and *Penicillium chrysogenum* (van den Berg et al., in press). Of special importance is the capacity of industrial fungi like *A. niger* and *A. oryzae* considered as safe with respect to the production of toxic metabolites (Frisvad et al., 2007). Comparative genomics (Rokas et al., 2007; Rokas and Galagan, in press) provides clues to the acquisition of genes and their lack of expression in domesticated strains, and also identifies strains described as separate species as ecotypes of a distinct species, such as *A. oryzae* as a domesticated *A. flavus* (Rokas et al., 2007). Future enhancing work on such strains aims at removing bad genes and adding good genes.

These large numbers of now predictable metabolites however do not match the number of known secondary metabolites (Hoffmeister and Keller 2007). While in *A. nidulans* 25 compounds have been described (Table 1), and some of those are not related to the strain FGSC A4 sequenced, 53 PKS genes and NRPS genes including putative members of these groups have been identified so far within the genome sequence (http://www.binf.gmu.edu/fseifudd/smurf/index1.html). A recent inventory of *Aspergillus* metabolites describing more than 450 compounds has been set up to define and differentiate recombinant *A. nidulans* strains (Kouskoumvekaki et al., 2008). Most of these compounds are unknown, and from seven biomarkers selected to discriminate 6-methylsalicylate (6-MSA) producing mutant strains even five were yet unidentified. These data seem to indicate that a large number of metabolites are present and actually await their discovery and structural elucidation. On the other hand transcriptional inventories of Ascomycetes carried out have provided evidence that many secondary metabolite genes have not been expressed under the respective cultivation conditions, and even shifts of conditions like solid-state fermentation did not lead to the expression of complex

Abbreviations: A, adenylate domain; AA, amino acid; AcOH-Trp, *N*-acetyl-tryptophan; C, condensation domain; d, dehydrogenase domain; DMAT, dimethyl-allyl-tryptophan transferase = aromatic prenyl transferase; DMATS, dimethyl-allyl-tryptophan synthase; E, epimerization domain; ETP, epipolythiodioxopiperazine; k, kinase domain; MSF, msf membrane transporter; NRPS, nonribosomal peptide synthetase; p, transporter domainL; PK, polyketide; PKS, polyketide synthase; T, thiolation or carrier domain; te, thioesterase domain; TAF, triacetylfusarinine.

* Fax: +49 3031424783.
E-mail address: doehren@chem.tu-berlin.de

1087-1845/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.fgb.2008.08.008

Table 1
Known metabolites isolated from *Aspergillus nidulans* peptides and amino acid derivatives

Compound	Structural type	Bioactivity
Protease inhibitor 3127	Lipopeptide	Protease inhibitor
Echinocandin	Lipocyclopeptide	Antifungal
Emericellamide*	Cyclopeptidolactone	
Triacetylfusigen	Depsipeptide	Siderophore
Fusarinine	Depsipeptide	Siderophore
Ferricrocin	Cyclopeptide	Siderophore
Terrequinone*	AA derivative	
Emerin	AA derivative?	
AcOH-Trp	AA derivative	
Aspyridone*	AA-PK hybrid	
Polyketides and diverse structural types		
Ascoquinone	Polyketide	Spore pigment
Sporogenic psi factor	Polyketide	
Asperline	Pyranone deriv.	Antibacterial, etc.
Microperfuranone	Polyketide	Plant growth regulator
Fonsecin	Polyketide	
Nidulalin A	Polyketide	Antitumor
Nidulin	Polyketide	Antibiotic
Nidurufin	Polyketide	
Asperthecin	Polyketide	Pigment
Benzophenone derivative	Polyketide?	
Tropolone derivative	Polyketide	Antibacterial, antifungal
Unguinol	Depsidone (PK)	
Variecoxanthone	Polyketide	
Austin	Terpenoide	
Cordycepin	Nucleotide	Cytostatic
Pentostatin	Nucleotide	

* Identified by genomic mining efforts.

metabolite genes (Tamano et al., 2008). Molecular genetic genome mining approaches so far have utilized gene disruption approaches (Chiang et al., 2008) or the manipulation of transcriptional regulators (Bok and Keller 2004; Bok et al., 2006).

In a microarray approach deletion and overexpression of the nuclear global regulator gene *laeA* identified several secondary metabolite clusters in *A. nidulans*, including the gene cluster for the production of terrequinone (Bok et al., 2006; Bouhired et al., 2007; Balibar et al., 2007). Terrequinone has been known as a metabolite of *A. terreus* (Wijeratne et al., 2003), and the respective gene cluster is also readily detected in the now available genome sequence. The significance of this discovery is the identification of a single module NRPS-like enzyme activating and dimerizing the alpha-keto acid indolepyruvate to bisindolylbenzoquinone, thus assigning functions for this class of proteins. LaeA-orthologs are readily detected in other fungal genomes (Schardl, 2007) and their roles are being investigated. Transcriptional analysis of a *LaeA* deletion mutant of *A. fumigatus* revealed the control of13 out of 22 secondary metabolite clusters by this factor Perrin et al., (2007). Likewise in *A. flavus* LaeA deletion and overexpression affected secondary metabolism and production of sclerotia (Kale et al., 2008).

Further regulatory genes are being identified, such as *mpkB*, a putative pheromone response mitogen-activated protein kinase, affecting sterigmatocystin, penicillin and terrequinone A synthesis, acting apparently through a link to LaeA expression (Atoui et al., 2008).

The specific activation of a cryptic gene cluster by overexpression of an associated transcriptional regulator (AN8414.3) led to the expression of the polyketide-amino acid metabolite aspyridone, so far unknown in *A. nidulans* (Bergmann et al., 2007). This application of a pathway-specific regulatory gene represents a new approach in drug discovery (Brakhage et al., 2008; Schneider et al., 2008).

Likewise the genomic inventory provides sound evidence for the absence of pathways, such as for epipolythiodioxopiperazines (ETP), found in *A. fumigatus, A. terreus, A. flavus*, and *A. niger*, but not in *A. nidulans* (Patron et al., 2007; Fox and Howlett 2008). Such losses of pathway specific gene clusters are attributed to differential gene loss (Fedorova et al., 2008).

To further advance genome mining efforts a bioinformatic approach has been applied to the NRPS gene inventory of the *A. nidulans* genome to describe and identify types of peptide synthetases, and possibly predict their products.

2. NRPS genes

2.1. Typing of NRPS genes

Phylogenetic analysis of NRPS genes and gene clusters has revealed a fairly conservative maintenance of these complex multistep biosynthetic proteins. Thus the organization of cyanobacterial cyclopeptide synthetases of the microcystin/nodularin type is thought to be conserved since about 1.5 billion years (Rantala et al., 2004). The structural organization of the key enzyme in penicillin biosynthesis, ACV synthetase, which has very likely been transferred horizontally from prokaryotic producers to Ascomycetes several hundred millions years ago, is present and unchanged in both kingdoms (Liras et al., 1998; van den Berg et al., in press). Both domain organization of the multienzymes and type of products tend to be conserved, although the distribution of genes is scattered, indicating frequent losses and possible acquisitions.

In a recent evaluation of NRPS genes of Aspergilli Cramer et al. (2006) listed major synthetases and based on a phylogenetic analysis of adenylate domains differentiated the groups (1) hybrid PKS/NRPS, (2) synthetases with *N*-methylated amino acids, (3) ETP toxin synthetases, (4) siderophore synthetase, (5) putative unknown siderophore synthetases, (6) penicillin synthetases, and (7) specific unknown synthetases of the oryzae/flavus group. This evaluation has been based on 88 Aspergilli NRPS and 8 additional NRPS comprising 216 adenylate domains. With respect to *A. nidulans* 14 NRPS were defined, 3 of which had been identified at that time, and 4 classified as hybrid PKS/NRPS (AN8412), *N*-methyl-amino acid containing synthetase (AN9226), siderophore synthetase involved in oxidative stress (AN6236), and penicillin biosynthesis (AN2621). In the meantime more data on fungal NRPS has become available mainly through genome sequencing efforts. Tools as substrate binding site analysis (Rausch et al., 2005), SMURF (The Secondary Metabolism Unknown Regions Finder, Nierman et al., 2007), and SYBIL (http://www.sybil.sourceforge.net/) have been applied to extend the bioinformatic description of fungal NRPS, and with respect to the *A. nidulans* genome more details have become available and are described below.

A phylogenetic analysis of 370 fungal adenylate domains revealed additional groups like peptaibol synthetases, small siderophore synthetases involved in iron acquisition, and aryl-capped peptides of the ergopeptide type. In addition, a distinct grouping of adenylate domains is observed with respect to their specificity, and not only the synthetase type (Zelakowska-Komon et al., 2007 and unpublished data).

2.2. Adenylate domain specificity

Likewise adenylate domains with identical or related substrate binding pockets are linked in distinct clusters in phylogenetic analysis (results not shown). Such clusters combine identical or related substrate specificities, as has e.g. been shown for peptaibol synthetases (Zelakowska-Komon et al., 2007). In attempts to identify a nonribosomal code, or broader speaking to identify substrates for yet unknown enzymes, work based on the first crystal structure available, that of the phenylalanine activating domain of gramicidin S synthetase, has defined contact side chains for amino acids

(Stachelhaus et al., 1999, Challis et al., 2000). This set of amino acids referred to as specificity code proved to be extremely useful in bacterial NRPS structures, and permits in many cases the prediction of peptide structures from unknown NRPS genes. However, none of the bacterial codes has any match in fungal adenylate domains, except for the case of the likely horizontal transfer of bacterial NRPS as described for ACVS, the penicillin precursor tripeptide synthetase (result not shown). Further assignment of fungal codes is expected to improve the predictability from sequence data. Interestingly different specificity codes are found for some amino acids indicating the evolution of different types of domains with related specificities, resembling in some way a degenerated specificity code.

Thus by combining both structural relationships of adenylate domains including their binding sites and their affiliation to structural types of NRPS provides some clues to the types of peptide products.

2.3. Compilation and update of A. nidulans NRPS

Key Secondary metabolite genes largely identified by SMURF (http://www.binf.gmu.edu/fseifudd/smurf/index1.html) have been compiled in Table 2. These include 26 putative PK synthases besides the known sterigmatocystin synthase (AN7825) and conidial pigment synthase (AN8209). Further included are 6 aromatic prenyltransferases, of which 3 are linked to putative biosynthetic clusters of PKS/NRPS genes. These include AN8514, which has been shown to be involved in terrequinone biosynthesis (Bok et al., 2006). The terrequinone biosynthetic cluster comprises the genes *tdiA-E*, with the key enzyme TdiA (AN8513), a single module NRPS-like protein, activating indole pyruvate formed by the aminotransferase TdiD (AN8516) from L-tryptophane, catalyses formation of the bis-indolebenzoquinone intermediate (Balibar et al., 2007; Schneider et al., 2007). Interestingly *tdiA* has not been picked up by this version of SMURF, and gene fragments like AN6962, a single condensation domain fragment, and AN8504, an 892 amino acid protein with a distorted adenylate domain, possible a sequencing error, are not considered. Further 24 NRPS-related genes have been detected, with the siderophore synthetase SidC (AN0607), ACV synthetase acvA (AN2621), aspyridone synthetase (AN8412) and emericellamide synthetase (AN2545) as identified encoded proteins. Together with *tdiA* 8 single module NRPS-like enzymes have been detected, which are identified by BLAST as acyl-CoA synthetases, and are unlikely to have amino acids as substrates, as the search algorithm for pocket analysis of NRPS domains fails (Rausch et al., 2005). The respective NRPS and related enzymes have been annotated and compiled in Table 3. In the following sections the entries are discussed.

2.4. AN0016 and AN1242: A family of peptide synthetases producing metabolites with two D-amino acids

A remarkable NRPS system is the AN0016-encoded type with the domain organization **ATECA*ATCATECTCT***, where **A** corresponds to an adenylate domain, **T** to a peptidyl carrier domain, **C** to a condensation domain, and **E** to an epimerisation domain. The second A-domain is defect, as indicated by **A***, missing the essential A3-motif (data not shown). The terminal T-domain has no serine-residue within the motif for posttranslational modification with CoA by 4′-phosphpopantetheine attachment. Surprisingly the AN1242 product with 6078 AA, also has a defective A-domain. All Stachelhaus/Challis substrate binding pockets (Rausch et al., 2005) are identical, but the two proteins have a low amino acid identity below 45%. All adenylate domains group together, respectively, in phylogenetic analysis (results not shown).

Table 2
Secondary metabolite related key genes in *A. nidulans*

Accession	Gene annotation	Abbreviation
AN0150.3	Polyketide synthase, putative	PKS
AN0523.3	Polyketide synthase, putative	PKS
AN0607.3	Nonribosomal peptide siderophore synthase sidc	NRPS
AN1242.3	Nonribosomal peptide synthase, putative	NRPS
AN1680.3	NRPS-like enzyme, putative	NRPS-Like
AN1784.3	Polyketide synthase, putative	PKS
AN2032.3	Polyketide synthase, putative	PKS
AN2035.3	Polyketide synthase, putative	PKS
AN2064.3	NRPS-like enzyme, putative	NRPS-Like
AN2545.3	Nonribosomal peptide synthase, putative	NRPS
AN2547.3	Polyketide synthase, putative	PKS
AN2621.3	*N*-(5-amino-5-carboxypentanoyl)-L-cysteinyl-D-valine synthase AcvA	NRPS
AN2924.3	NRPS-like enzyme, putative	NRPS-Like
AN3230.3	polyketide synthase, putative	PKS
AN3386.3	polyketide synthase, putative	PKS
AN3396.3	NRPS-like enzyme, putative	NRPS-Like
AN3495.3	NRPS-like enzyme, putative	NRPS-Like
AN3496.3	Nonribosomal peptide synthase, putative	NRPS
AN3612.3	Polyketide synthase, putative	PKS
AN4827.3	NRPS-like enzyme, putative	NRPS-Like
AN5318.3	NRPS-like enzyme, putative	NRPS-Like
AN6000.3	Polyketide synthase, putative	PKS
AN6236.3	Nonribosomal peptide synthetase, putative	NRPS
AN6431.3	Polyketide synthase, putative	PKS
AN6444.3	NRPS-like enzyme, putative	NRPS-Like
AN6448.3	Polyketide synthase, putative	PKS
AN6784.3	DMATS type aromatic prenyltransferase, putative	DMAT
AN6791.3	Polyketide synthase, putative	PKS
AN7071.3	Polyketide synthase, putative	PKS
AN7084.3	PKS-like enzyme, putative	PKS-Like
AN7489.3	Mitochondrial 3-oxoacyl-[acyl-carrier-protein]-synthase, putative	PKS-Like
AN7815.3	Sterigmatocystin biosynthesis fatty acid synthase alpha subunit, putative	PKS-Like
AN7825.3	Sterigmatocystin polyketide synthase	PKS
AN7838.3	PKS-like enzyme, putative	PKS-Like
AN7884.3	Nonribosomal peptide synthase, putative	NRPS
AN7909.3	Polyketide synthase, putative	PKS
AN8105.3	NRPS-like enzyme, putative	NRPS-Like
AN8209.3	Conidial yellow pigment biosynthesis polyketide synthase	PKS
AN8383.3	Polyketide synthase, putative	PKS
AN8412.3	Hybrid PKS-NRPS, aspyridone synthetase	HYBRID
AN8513.3	Bis-indolylbenzoquinone synthetase	NRPS-Like
AN8514.3	Bisindolylquinone prenyltransferase	DMAT
AN8910.3	Polyketide synthase, putative	PKS
AN9005.3	Polyketide synthase, putative	PKS
AN9226.3	nonribosomal peptide synthase, putative	NRPS
AN9244.3	nonribosomal peptide synthase, putative	NRPS
AN9291.3	MFS transporter/NRPS-like enzyme, putative	NRPS-Like
AN10430.3	Polyketide synthase, putative	PKS
AN10486.3	NRPS-like enzyme, putative	NRPS-Like
AN10576.3	Nonribosomal peptide synthase, putative	NRPS
AN11080.3	DMATS type aromatic prenyltransferase, putative	DMAT
AN11191.3	Polyketide synthase, putative	PKS
AN11194.3	DMATS type aromatic prenyltransferase, putative	DMAT
AN11202.3	DMATS type aromatic prenyltransferase, putative	DMAT

The related NRPS in *A. fumigatus* (6270 AA, Afu1g10380), *A. niger* (6311 AA, 118587) and *A. terreus* (5842 AA, EAU38874) all have functional second A-domains (**ATECAATCATECTCT***), while in *A. oryzae* (5199 AA, AO090038000390) this domain has been deleted (**ATECATCATECTCT***) (Table 4a).

That some sort of relation exists seems probable comparing the pocket lining residues, although none corresponds to a known substrate. A suggested product is a cyclotetrapeptide with two adjacent D-amino acid residues, possibly reduced to a tripeptide in case of *A. oryzae* and *A. nidulans*, but iterative processes cannot be excluded, and thus besides tetrapeptides even pentapeptides of the malformin type could be considered as products.

Table 3
NRPS and NRPS-related genes of *A. nidulans*

Gen-ID	Classification	Structural type[1]	Closest similarity	Pocket description	Related clusters
AN0016.3	Putative cyclotripeptide synthetase, 5936 aa	ATECAATCATECTCT	*A. terreus* ATEG_00228 (46% ID), 5842 AA *A. oryzae* XP_001821068 (59% ID), 4760 AA	A1- AASIAGVTK A2-unclear A3- DVLNVGAIIK A4- AMFVGAVYK	In all Aspergilli
AN0607.3	Siderophore synthase sidc, 4761 aa	ATCATCATCTCTC	A1- DPMMWMAINK A2- DVQHTITVVK A3- DPLSTGAIGK		In all Aspergilli
AN1242.3	Putative cyclotripeptide synthetase, 6077 aa	ATECAATCATECTCT	*A. fumigatus* AFUA_1G10380 (51% ID), 6269 AA	A1- AASIAGVTK A2-unclear A3- DVLNVGAIIK A4- AMFVGAVYK	In all Aspergilli
AN1680.3	NRPS-like enzyme, putative, no AA substrate	A*Tdk	*A. terreus* ATEG_06998 (72% ID), 1236 AA		In *A. oryzae* and *A. terreus*
AN2064.3	NRPS-like enzyme, putative, no AA substrate	A*d	*A. niger* An13g02460 (53% ID), 1038 AA		none
AN2545.3	Nonribosomal peptide synthase, putative, clustered with PKS AN2547.3, 7047 AS	TECATECATCATCATCATC	*A. nidulans* AN7884.3 (32 % ID), 7214 AA	A1-DIQGVLAMQK A2-DASQIGGIYK A3-DIHFVGAIAK A4-DLLVVAGILK A5-DIAILVAILK	None
AN2621.3	*N*-(5-amino-5-carboxypentanoyl)-L-cysteinyl-D-valine synthase AcvA, clustered with isopenicillin N synthase AN2622.3 and acyltransferase AN2623.3	ATCATCATEte	*A. oryzae* XP 001825447.1 (66 % ID) 3774 AA	A1- EPRNIVEFVK (Aad) A2- DHESDVGITK (Cys) A3- DFESTAAVYK (Val)	In *A. oryzae, A. flavus**, *P. chrysogenum*
AN2924.3	NRPS-like enzyme, putative, no AA substrate	A*d	*Sclerotinia sclerotiorum* SS1G_02020 (33% ID), 1043 AA		none
AN3396.3	NRPS-like enzyme, putative, no AA substrate, 939 AA	A*Td	*A. niger* An01g11790 (61% ID), 966 AA		*None*
AN3495.3	NRPS-like enzyme, putative, 1480 AA, linked with AN3496.3	CATd	*A. terreus* ATEG 09086 (36 % ID), 2307 AA	DLYSAVAVVK	None
AN3496.3 (38 % ID), 4423 AA	Nonribosomal peptide synthase, putatve, 2326 AA A1- DVSNVGSIIK A2- DAQNIAVINK	TCATCAT none	*Gibberella zeae* FG08209.1		
AN4827.3	NRPS-like enzyme, putative, no AA substrate, 1052 AA	Ad	*Botryotinia fuckeliana* BC1G_12512 (53% ID), 1952 AA		None
AN6236.3	TAF-type siderophore synthetase, putative, 2086 AA	ATCxTC	N. fischeri NFIA_005590 (64 % ID) 2083 AA	DVDGGGGIGK	In various Ascomycetes
AN6444.3	NRPS-like enzyme, putative , no AA substrate, acyl-CoA synthetase, 1040 AA	A*d	Coccidioides immitis CIMG_04688 (42% ID), 111172 AA		None
AN7884.3	Nonribosomal peptide synthase, putative, 7214 AA, linked to fatty acid synthase, dehydrogenase, branched chain aminotransferase, P450 hydroxylase, and transporters	ATCATECATCATCATCATd	*A. nidulans* AN2545.3 (32 % ID) 7047 AA	A1- DLGLSGMILK A2- DVGFTGCVYK A3- DALFMGGVYK A4- DAMMVGVMIK A5- DVAFTGSIWK A6- DVLFIGGVAK	None
AN8105.3	NRPS-like enzyme, putative, 1077 AA	ATd	*A.niger* An16g00600 (42 % ID), 1031 AA	DVWWLGAVFK	None
AN8412.3	Hybrid PKS-NRPS, ApdA, aspyridone synthetase, 3931 AA	KS-AT-CAd	*N. fischeri* NFIA 001530 (81% ID), 3763 AA	DMVICGCAAK (Tyr)	In *N. fischeri* and *A. niger*
AN8504.3	NRPS-like enzyme, putative, 892 AA	A*C	*Leptosphaeria maculans* SirP (31% ID), 2176 AA		
AN8513.3	Bis-indolylbenzoquinone synthetase, terrequinone biosynthesis, linked with aminotransferase, DMATS and methyltransferase, 950 AA	A*Tte	*Chaetomium globosumXP_001230203.1 (68% ID), 957 AA* N. fischeri *NFIA 094680 (58% ID), 994 AA*		In *A. terreus*
AN9129.3	Nonribosomal peptide synthase, putative, 950 AA	A*			*In N. fischeri, A. terreus, A. clavatus* and *A. fumigatus*

AN9226.3	Nonribosomal peptide synthase, putative, 2559 AA, Produces an N-methylated cyclodipeptide, flanked by 2 P450 enzymes and a dioxygenase	ATCAMTC	*N. fischeri NFIA* 043670 (51% ID), 2662 AA	A1- GMIIVAAGIK A2-DAFFYGCIVK		In *N. fischeri*
AN9243.3	Nonribosomal peptide synthase, putative, linked to AN9244.3, 948 AA	CA	*A. clavatus* ACLA_093780 (50% ID), 6199 AA	–LFIGGVF-		None
AN9244.3	Nonribosomal peptide synthase, putative, linked to AN9243.3, 2242 AA	ATCTCA	*A. clavatus* ACLA_093780 (50% ID), 6199 AA	A1- DVLFFGGIAK A2- DTGLGVSIMK		None
AN9291.3	MFS transporter/NRPS-like enzyme, putative, 1430 AA	TCp*	*A. oryzae* XP_001817985.1 (43% ID), 1654 AA			None
AN10297.3	=AN2403.3—NRPS-like enzyme, putative, no AA act, 1061 AA	A*d	*A. niger* An02g00840 (67% ID), 1056 AA			In *A. terreus*
AN10486.3	=AN3912.3—NRPS-like enzyme, putative, no AA ac, 1053 AAt	ATd	*A. oryzae* XP_001819027.1 (65% ID), 1068 AA			In all Aspergilli
AN10576.3	Nonribosomal peptide synthase, putative, 1704 AA	ATCC	*Gibberella zeae* FG02315.1 (31% ID), 7599 AA	DVDLQGVVAK		None

* Based on *A. flavus* genome sequence data 80.m03224.

[1] Domain notations: A-adenylate domain, T-carrier domain, C-condensation domain, *M-N*-methyltransferase domain, d-NAD-binding domain, k-kinase domain, p-transporter domain.

The to both synthetases related *A. fumigatus* NRPS Pes1 (Afu1g10380, 6270 AA) is flanked by the similar ABC transporter Afu1g10390 (1396 AA, 55% identity, 70% similarity) to the *A. nidulans* transporter (AN1240), and the MSF membrane transporter Afu1g10370 (565 AA). A related transporter protein of 1501 AA is CIMG_09753 [*Coccidioides immitis* RS] sequenced from BROAD, but adjacent genes are not disclosed yet (CH476730).

A conserved cluster analysis within Aspergilli reveals that this type of NRPS is always associated with an ABC transporter (Table 4b).

These data indicate that a cyclic peptide product may be formed and excreted. Doyle and colleagues have shown that Pes1 is indeed expressed in *A. fumigatus*, and confers protection against oxidative stress (Reeves et al., 2006). A knockout mutant showed reduced virulence and altered spore surface morphology.

2.5. Ferricrocin synthetase (AN0607)

SidA (AN0606) and SidC (AN0607) have been analysed in detail by Haas et al. (Eisendle et al., 2003, 2006). The adjacent long chain fatty acid CoA ligase encoded by AN0609, which is readily identified by NCBI-BLAST, is presumably involved in L-ornithine modification. SidC represents a type II ferrichrome synthetase also present in *Aspergillus fumigatus* (AFUB_016590, 4763 AA, 55% identity), *Aspergillus niger* (An06g01300, 4722 AA, 58% identity) *Aspergillus oryzae* (AO090023000528, 4760 AA, 59% identity), *A. terreus* (ATEG_05073, 4744 AA, 58% identity), *Neosartorya fischeri (NFIA_008170,* 4759 AA, 56% identity), and the basidiomycetes *Ustilago maydis* (UM01434.1, 4830 AA, 32% identity) and *Omphalotus olearius* (AAX49356, 4547 AA, 33% identity) (Schwecke et al., 2006). Some fungi like *Fusarium graminearum* and *Chaetomium globosum* indeed possess both type II and type IV ferrichrome synthetases, suggesting loss of type IV genes in Aspergilli.

2.6. The small siderophore synthetase SidD (AN6236)

The AN6236.3 product has a domain structure **ATCxTC**, where x represents a region with similarities to A domains, and also N- and C-terminal unusual regions. This group of synthetases termed SidD has in Ascomycetes the conserved substrate code DVDGGGGIGK and 55–65% identity at the AA level, and forms a distinct cluster in a neighbor-joining tree from CLUSTAL analysis of all TAF-like NRPS. Physiological roles of SidD have been studied in detail in *A. fumigatus* (Schrettl et al., 2007).

This conserved putative fusarinine type siderophore biosynthesis cluster composed of NRPS (AN6236), ABC-transporter (AN6237) and putative siderophore transporter (AN6238) is also found in *Aspergillus fumigatus, Aspergillus oryzae, Aspergillus terreus, Aspergillus clavatus* and *Neosartorya fischeri*. While the same predicted substrate code is also found in *Botryotinia fuckeliana* (AAX11420, 2058 AA) and *Chaetomium globosum* (EAQ85461, 1733 AA), similar built NRPS have slightly different codes: DVDHGGAVGK: *A. terreus*, *N. crassa*, *Coccidioides immitis* and *Magnaporthe grisea*; DVDGAGAVGK: *Cochliobolus heterostrophus*, *Alternaria brassiceae*, and *Phaeosphaeria nodorum*; DVDNGGGIGK: *Gibberella zeae*. Both types of small siderophore synthetases are present e.g. in *A. niger* and *P. chrysogenum*.

2.7. AN9226, an N-methyl-cyclodipeptide synthetase

AN9226.3 encodes a 2560 AA NRPS with the structure **ATCAMTC**. The ortholog of *N. fischeri* (NFIA043670, 2662 AA) picked up by SYBIL has the same structure. The region from 1600 to 2100, inserted in the second A domain, shows similarities to the cyclosporin synthetase *N*-methyltransferase regions (45% identity). The codes provide no clear picture, but the *N*-methyl AA is likely to

Table 4
Related NRPS systems of peptide synthetases AN0016 and AN1242

nid_5935	nid_6077	nig_6311	fumi_6269	ory_5199	ter_5842	Nfia_6269
(*4a*) *Predicted substrate pocket lining residues of the 7 related multienzmyes*						
DAASIAGVTK	DAASIAGVTK	DAVIIAAVAK	DAVSVAGVAK	DAVTVAGVTK	DAVSAVGVTK	DAVSVAGVAK
defect	defect	DMQSAWFSSK	DMNVAYLSSL	deleted	DVQSVYFST-	DMNIAYLSSL
DVLNVGAIIK	DVLNVGAIIK	DVLSTGAVAK	DVKSVGAVMK	DVSTIGAVAK	DVLSI-AILK	DVMSVGAVMK
DAMFVGAVYK	DAMFVGAVYK	DAMFVGGIFK	DAMFAGGIFK	DAMFAGGIFK	DAMF—GIFK	DAMFAGGIFK
		NRPS				ABC-transporter
(*4b*) *Adjacent transporters of these NRPS*						
A. fumigatus		Afu1g10380, Pes1, 6269 AA				Afu1g10390, 1396 AA
N. fischeri		NFIA_015290, Pes1, 6269 AA				NFIA_015280, 1468 AA
A. oryzae		AO090038000390, 5199 AA				910 AA, 1437 AA
A. nidulans		AN1242, 6078 AA				AN1240, 1116 AA
A. nidulans		AN0016, 5936 AA				AN0015, 1457 AA
A. clavatus		ACLA_025160, Pes1, 6263 AA				ACLA_025170, 1473 AA

be Gly/Ala/Val-related. The NRPS gene is flanked by two P450 enzymes and a dioxygenase.

3. PKS-NRPS clusters

AN2545.3 encodes a 7047 AA NRPS with the structure **TECATE-CATCATCATCATC**, and has been shown to be a member of the emericellamide cluster, comprising the PKS **AN2547.3** (2534 AA), the acyl-CoA synthetase **AN2549.3**, and the acyl carrier protein **AN2548.3** (Chiang et al., 2008). The PKS domain structure is KS-AT-DH-ER-KR-ACP.

Emericellamides A-F consist of an aliphatic hydroxy acid (C_{10} or C_{12}) elongated with aminoacyl residues, and finally undergoing lactonization presumably catalysed by the terminal condensation domain. The PKS provides the 2,4-dimethyl-3-hydroxy-(do)decanoyl-intermediate (stereochemistry not yet established), which has been proposed to be released, and not directly transferred to the NRPS, as both the acyl-CoA ligase **AN2549.3** and the acyl carrier protein **AN2548.3** have been shown by gene knockout technology to be required for emericellmide formation. The proposed path then includes reactivation of the (do)decanoyl intermediate and transfer to the ACP.

Unclear remains the function of the two epimerization domains: the first might act on the acyl compound, while the first amino acid is glycine. This might suggest that related analogs exist, where this position indeed has a D-configuration. The substrate codes DIQGVLAMQK, DASQIGGIYK, DIHFVGAIAK, DLLVVAGILK, DIAILVAILK have been assigned to Gly, L-Val, L-Leu, and L-Ala for the last two. It may seem strange that two different specificity codes are found for L-Ala. A comparison of the two L-alanine specific adenylate domains reveals an amino acid identity of only 44%, which is in a similar range as the other adenylation domains of this NRPS, clearly pointing to deviating types of binding pockets.

Interestingly emericellamides A and B have been identified before as a products of a marine *Aspergillus* isolated from the surface of the green algae induced by cocultivation with the marine *Actinobacteria salinispora* (Oh et al., 2007). In *A. nidulans* emericellamides are produced on glucose minimal medium.

AN7884.3 encodes a 7214 AA NRPS with the structure **ATCATE-CATCATCATCA*T*deh,** with deh denoting a dehydrogenase domain. Adjacent genes are a dehydrogenase, a P450 enzyme, the two subunits of a fatty acid synthase (FAS), a branched chain aminotransferase, an ABC transporter and a resistance protein (Jonathan Walton, personal communication). It is not sure if the FAS system contributes to the final peptide structure, but both *A. nidulans* and *A. oryzae* host 5 and 6 FAS, respectively, and some are presumably involved in secondary metabolite biosynthesis. The predicted product could be a reduced hexapeptide with an unusual starter amino acid.

4. Interacting NRPS systems

AN3495.3 encodes a 1480 AA NRPS with the structure **CATdeh**. At first sight Lys2 analogy has been suspected, and the gene classified into Lys biosynthesis and primary metabolism. BLAST, however, links it clearly to nonribosomal systems, with highest similarities with peptaibol synthetase genes because of the reduction domain. The code DLYSAVAVVK gives no clue. Associated is the NRPS **AN3496.3**, 2326 AA, with the structure TCATCAT. The codes DVSNVGSIIK and DAQNIAVINK are unknown. The initial T suggests N-terminal modification. The proposed co-product is a tripeptide with reduced C-terminal.

AN9243.3 (with a reading frame of 948 AA) and AN9244.3 (with a reading frame of 2242 AA) possibly represent a single 5-module synthetase, since there is strong similarity (50–60% id) to *A. clavatus* ACLA_093780 (6199 AA), with nearly identical pockets in positions 1 and 4, a recheck of sequence required.

5. One-module NRPS-like enzymes

Monomodular NRPS-related enzymes involved in secondary metabolism have so far not been well characterized. While Lys2 as primary metabolite enzyme is easily recognized by its high conservation of structure and the aminoadipate pocket signature, so far now only the TdiA-type containing an α-keto acid activating domain is known. The specificity code approach is generally not working, as amino acids often do not seem to be substrates. So far the pocket binding code approach has not been pursued for keto and hydroxy acids or acyl- and aryl-compounds in general. This is not an easy task, and deduction from GrsA-sequence alignment may lead to ambiguous results, as major signature sequences differ (Schneider et al., 2007).

AN1680.3 encodes a unique 4-domain protein with an acyl-CoA-related adenylate domain, a carrier domain, an NAD-binding domain, and a terminal gluconate kinase domain. Closely related are a 1236 AA protein of *A. terreus* (ATEG_06998), with 72% identity at the amino acid level, and a 973 AA protein of *A. oryzae* (XP001816557.1) with 68% identity. Both reside in orthologous clusters flanked by yet unidentified genes. The high similarity implies primary metabolic functions.

AN2064.3 encodes a 1038 AA protein with an acyl-CoA-related adenylate domain and an NAD-binding domain. Related genes are found in all Aspergilli with only 52–53% identity at the amino acid level. The environment of the *A. nidulans* gene is unique, while *N. fischeri, A. terreus, A. clavatus* and *A. fumigatus* share orthologous clusters with adjacent transcription factors of yet unknown function.

AN2924.3 encodes a unique acyl CoA synthetase with an NAD-binding domain of 984 AA. Only distantly related proteins can be found in *Sclerotinia sclerotiorum* (SS1G_02020), 33% identity, and *A.oryzae* (XP 001817724.1), 29% identity.

AN3396.3 encodes a 3-domain protein with an acyl-CoA-related adenylate domain, a carrier protein domain and an NAD-binding domain. An ortholog is found in *A. niger* An01g11790 (61% ID), 966 AA, with a different chromosomal environment.

AN4837.3 encodes a unique acyl CoA synthetase with an NAD-binding domain of 1051 AA, related proteins with 44–53% amino acid identity are found in *Gibberella zeae, Magnaporthe grisea, Sclerotinia sclerotiorum*, and *Botryotinia fuckeliana*.

AN6444.3 encodes a unique acyl CoA synthetase with an NAD-binding domain of 1040 AA. Only distantly related proteins can be found in *Coccidioides immitis* (CIMG_04688), 42% identity, and *A. terreus* (XP001212808.1), 41% identity at the AA level.

AN8105.3 encodes a 3-domain protein with an adenylate domain, a carrier protein domain and an NAD-binding domain of 1031 AA. Related proteins with only 40–42% amino acid identity are present in most Aspergilli with different genomic environments. The unknown substrate pocket code gives no clue.

AN8513.3, encodes TdiA, which has been identified as bis-indolylbenzoquinone synthetase, a key enzyme in terrequinone biosynthesis, clustered with an aminotransferase, DMATS and a methyltransferase (Bouhired et al., 2007). The 3-domain protein contains an adenylate domain activating indolepyruvate, a carrier protein domain and a thioesterase. An orthologous cluster is known in *A. terreus*.

AN9129.3 encodes a predicted 950 AA protein with an unusual adenylate domain. Related proteins are found in *N. fischeri, A. clavatus, A. fumigatus* and *A. terreus* with 55–58% amino acid identity. All proteins reside in conserved clusters, linked to an enoyl-CoA hydratase/isomerase family gene. In the other Aspergilli additional conserved genes include transcription factors, acetyltransferases, and FAD-binding proteins.

AN10297.3, identical to AN2403.3, encodes a 1061 AA acyl CoA synthetase with an NAD-binding domain

AN10486.3, identical to AN3912.3, encodes a 3-domain protein with an adenylate domain, a carrier protein domain and an NAD-binding domain of 1053 AA.

AN10576.3 encodes a unique one module NRPS with an additional condensation domain. The pocket binding code DVDLQGVVAK is unknown.

5.1. Mixed PKS/NRPS genes

AN8412.3 encodes a PKS-NRPS hybrid identified as aspyridone synthetase (Bergmann et al., 2007). A similar enzyme is involved in tenellin biosynthesis in *Cordyceps bassiana* (CAL69597, 4239 AA, 40% identity at the amino acid level) (Eley et al., 2007). Similar multienzymes are also known from *Penicillium expansum* (cytochalasin, CAO91861, 4013 AA, 42% identity), *A. oryzae* (AO09000-1000277, 3946 AA, 44% identity), and *N. fischeri* (NFIA001530, 3763 AA) with a surprising 81% identity, with yet unknown products. The substrate binding pockets DMVICGCAAK in AN8412 and DMVIPPCAAK in tenellin synthetase both define L-Tyr. The AN8412 protein and the orthologous *N. fischeri* enzyme have identical substrate binding pockets.

5.2. Unusual systems

AN9291.3 encodes a 3-domain protein of a carrier domain, a condensation domain and an efflux pump domain. While numerous efflux pump genes linked to NRPS genes are known, only one such fusion construct is known in *A. oryzae* (XP_001817985.1).

6. Conclusion

In principle the current state of understanding of NRPS organisation permits a rough guess about the type of product. The lack of interpretation of substrate binding pockets, which is well established in the bacterial systems, prevents the prediction of chemical properties. This is a significant drawback, and improvement is only slowly accomplished. The comparative analysis of gene clusters aided by SYBIL is a significant improvement of predictive power and a major tool in our future understanding of the plasticity of secondary metabolism.

Acknowledgements

Thanks to Jonathan Walton, Natalia Fedorova, Christian Hertweck, Nancy Keller, Dirk Hoffmeister, Russell Cox, Jens Frisvad, Robert Cramer, and Carl Balibar for the communication of data. The study has been supported by the EC-grant EUROFUNGBASE.

References

Atoui, A., Bao, D., Kaur, N., Grayburn, W.S., Calvo, A.M., 2008. *Aspergillus nidulans* natural product biosynthesis is regulated by *mpkB*, a putative pheromone response mitogen-activated protein kinase. Appl. Environ. Microbiol. 74, 3596–3600.

Balibar, C.J., Howard-Jones, A.R., Walsh, C.T., 2007. Terrequinone A biosynthesis through L-tryptophan oxidation, dimerization and bisprenylation. Nat. Chem. Biol. 3, 584–592.

Bergmann, S., Schümann, J., Scherlach, K., Lange, C., Brakhage, A.A., Hertweck, C., 2007. Genomics driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. Nat. Chem. Biol. 3, 213–217.

Bok, J.W., Keller, N.P., 2004. LaeA, a regulator of secondary metabolism in *Aspergillus spp.*. Eukaryot Cell 3, 527–535.

Bok, J.W., Hoffmeister, D., Maggio-Hall, L.A., Murillo, R., Glasner, J.D., Keller, N.P., 2006. Genomic mining for *Aspergillus* natural products. Chem. Biol. 13, 31–37.

Bouhired, S., Weber, M., Kempf-Sontag, A., Keller, N.P., Hoffmeister, D., 2007. Accurate prediction of the *Aspergillus nidulans* terrequinone cluster boundaries using the transcriptional regulator LaeA. Fungal Genet. Biol. 44, 1134–1145).

Brakhage, A.A., Schuemann, J., Bergmann, S., Scherlach, K., Schroeckh, V., Hertweck, C., 2008. Activation of fungal silent gene clusters: a new avenue to drug discovery. Prog Drug Res. 66, 1, 3–12.

Challis, G.L., Ravel, J., Townsend, C.A., 2000. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem. Biol. 7, 211–224.

Chiang, Y.M., Szewczyk, E., Nayak, T., Davidson, A.D., Sanchez, J.F., Lo, H.C., Ho, W.Y., Simityan, H., Kuo, E., Praseuth, A., Watanabe, K., Oakley, B.R., Wang, C.C., 2007. Molecular genetic mining of the *Aspergillus* secondary metabolome: discovery of the emericellamide biosynthetic pathway. Chem. Biol. 15, 527–532.

Cramer Jr., R.A., Stajich, J.E., Yamanaka, Y., Dietrich, F.S., Steinbach, W.J., Perfect, J.R., 2006. Phylogenomic analysis of non-ribosomal peptide synthetases in the genus *Aspergillus*. Gene 383, 24–32.

Eisendle, M., Oberegger, H., Zadra, I., Haas, H., 2003. The siderophore system is essential for viability of *Aspergillus nidulans*: functional analysis of two genes encoding l-ornithine *N*-5-monooxygenase (sidA) and a non-ribosomal peptide synthetase (sidC). Mol. Microbiol. 49, 275–359.

Eisendle, M., Schrettl, M., Kragl, C., Müller, D., Illmer, P., Haas, H., 2006. The intracellular siderophore ferricrocin is involved in iron storage, oxidative-stress resistance, germination, and sexual development in *Aspergillus nidulans*. Eukaryot. Cell 5, 1596–1603.

Eley, K.L., Halo, L.M., Song, Z., Powles, H., Cox, R.J., Bailey, A.M., Lazarus, C.M., Simpson, T.J., 2007. Biosynthesis of the 2-pyridone tenellin in the insect pathogenic fungus *Beauveria bassiana*. Chembiochem 8, 289–297.

Fedorova, N.D., Khaldi, N., Joardar, V.S., Maiti, R., Amedeo, P., Anderson, M.J., Crabtree, J., Silva, J.C., Badger, J.H., Albarraq, A., Angiuoli, S., Bussey, H., Bowyer, P., Cotty, P.J., Dyer, P.S., Egan, A., Galens, K., Fraser-Liggett, C.M., Haas, B.J., Inman, J.M., Kent, R., Lemieux, S., Malavazi, I., Orvis, J., Roemer, T., Ronning, C.M., Sundaram, J.P., Sutton, G., Turner, G., Venter, J.C., White, O.R., Whitty, B.R., Youngman, P., Wolfe, K.H., Goldman, G.H., Wortman, J.R., Jiang, B., Denning, D.W., Nierman, W.C., 2008. Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. PLoS Genet. 4, e1000046.

Fox, E.M., Howlett, B.J., 2008. Biosynthetic gene clusters for epipolythiodioxopiperazines in filamentous fungi. Mycol. Res. 112 (Pt 2), 162–169.

Frisvad, J.C., Smedsgaard, J., Samson, R.A., Larsen, T.O., Thrane, U., 2007. Fumonisin B2 production by *Aspergillus niger*. J. Agric. Food Chem. 55, 9727–9732.

Gaffoor, I., Brown, D.W., Plattner, R., Proctor, R.H., Qi, W., Tail, F., 2005. Functional Analysis of the Polyketide Synthase Genes in the Filamentous Fungus *Gibberella zeae* (Anamorph *Fusarium graminearum*). Eukaryot. Cell 4, 1926–1933.

Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.J., Wortman, J.R., Batzoglou, S., Lee, S.I., Baştürkmen, M., Spevak, C.C., Clutterbuck, J., Kapitonov, V., Jurka, J., Scazzocchio, C., Farman, M., Butler, J., Purcell, S., Harris, S., Braus, G.H., Draht, O., Busch, S., D'Enfert, C., Bouchier, C., Goldman, G.H., Bell-Pedersen, D., Griffiths-Jones, S., Doonan, J.H., Yu, J., Vienken, K., Pain, A., Freitag, M., Selker, E.U., Archer, D.B., Peñalva, M.A., Oakley, B.R., Momany, M., Tanaka, T., Kumagai, T., Asai, K.,

Machida, Nierman, W.C., Denning, D.W., Caddick, M., Hynes, M., Paoletti, M., Fischer, R., Miller, B., Dyer, P., Sachs, M.S., Osmani, S.A., Birren, B.W., 2007. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438, 1105–1115.

Hoffmeister, D., Keller, N.P., 2007. Natural products of filamentous fungi: enzymes, genes, and their regulation. Nat. Prod. Rep. 24, 393–416.

Kale, S.P., Milde, L., Trapp, M.K., Frisvad, J.C., Keller, N.P., Bok, J.W., 2008. Requirement of LaeA for secondary metabolism and sclerotial production in *Aspergillus flavus*. Fungal Genet Biol. 2008 Jul 11. [Epub ahead of print].

Kouskoumvekaki, I., Yang, Z., Jónsdóttir, S.O., Olsson, L., Panagiotou, G., 2008. Identification of biomarkers for genotyping Aspergilli using non-linear methods for clustering and classification. BMC Bioinformatics 9, 59.

Lee, B.-N., Kroken, S., Chou, D.Y.T., Robbertse, B., Yoder, O.C., Turgeon, B.G., 2005. Functional analysis of all nonribosomal peptide synthetases in *Cochliobolus heterostrophus* reveals a factor, NPS6, involved in virulence and resistance to oxidative stress. Eukaryot. Cell 4, 545–555.

Liras, P., Rodríguez-García, A., Martin, J.F., 1998. Evolution of the clusters of genes of β-lactam antibiotics: a model for evolutive combinatorial assembly of new β-lactams. Int. Microbiol. 1, 271–278.

Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., Kusumoto, K., Arima, T., Akita, O., Kashiwagi, Y., Abe, K., Gomi, K., Horiuchi, H., Kitamoto, K., Kobayashi, T., Takeuchi, M., Denning, D.W., Galagan, J.E., Nierman, W.C., Yu, J., Archer, D.B., Bennett, J.W., Bhatnagar, D., Cleveland, T.E., Fedorova, N.D., Gotoh, O., Horikawa, H., Hosoyama, A., Ichinomiya, M., Igarashi, R., Iwashita, K., Juvvadi, P.R., Kato, M., Kato, Y., Kin, T., Kokubun, A., Maeda, H., Maeyama, N., Maruyama, J., Nagasaki, H., Nakajima, T., Oda, K., Okada, K., Paulsen, I., Sakamoto, K., Sawano, T., Takahashi, M., Takase, K., Terabayashi, Y., Wortman, J.R., Yamada, O., Yamagata, Y., Anazawa, H., Hata, Y., Koide, Y., Komori, T., Koyama, Y., Minetoki, T., Suharnan, S., Tanaka, A., Isono, K., Kuhara, S., Ogasawara, N., Kikuchi, H., 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438, 1157–1161.

Misiek, M., Hoffmeister, D., 2007. Fungal genetics, genomics, and secondary metabolites in pharmaceutical sciences. Planta Med. 73, 103–115.

Nierman, W.C., Pain, A., Anderson, M.J., Wortman, J.R., Kim, H.S., Arroyo, J., Berriman, M., Abe, K., Archer, D.B., Bermejo, C., Bennett, J., Bowyer, P., Chen, D., Collins, M., Coulsen, R., Davies, R., Dyer, P.S., Farman, M., Fedorova, N., Fedorova, N., Feldblyum, T.V., Fischer, R., Fosker, N., Fraser, A., García, J.L., García, M.J., Goble, A., Goldman, G.H., Gomi, K., Griffith-Jones, S., Gwilliam, R., Haas, B., Haas, H., Harris, D., Horiuchi, H., Huang, J., Humphray, S., Jiménez, J., Keller, N., Khouri, H., Kitamoto, K., Kobayashi, T., Konzack, S., Kulkarni, R., Kumagai, T., Lafon, A., Latgé, J.P., Li, W., Lord, A., Lu, C., Majoros, W.H., May, G.S., Miller, B.L., Mohamoud, Y., Molina, M., Monod, M., Mouyna, I., Mulligan, S., Murphy, L., O'Neil, S., Paulsen, I., Peñalva, M.A., Pertea, M., Price, C., Pritchard, B.L., Quail, M.A., Rabbinowitsch, E., Rawlins, N., Rajandream, M.A., Reichard, U., Renauld, H., Robson, G.D., Rodriguez de Córdoba, S., Rodríguez-Peña, J.M., Ronning, C.M., Rutter, S., Salzberg, S.L., Sanchez, M., Sánchez-Ferrero, J.C., Saunders, D., Seeger, K., Squares, R., Squares, S., Takeuchi, M., Tekaia, F., Turner, G., Vazquez de Aldana, C.R., Weidman, J., White, O., Woodward, J., Yu, J.H., Fraser, C., Galagan, J.E., Asai, K., Machida, M., Hall, N., Barrell, B., Denning, D.W., 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature. 438, 115111-6. Erratum in: Nature 439 (7075), 502 (2006 January 26).

Nierman, W.C., Khaldi, N., Yu, J., Payne, G., Ma, L.-J., Fedorova, N., 2007. Computational and functional analysis of secondary metabolite clusters. Proc. 2007 Ann. Multi-Crop Aflatoxin/Fumonsin Elimination & Fungal Genomics Workshop, Atlanta 2007, p. 29.

Oh, D.C., Kauffman, C.A., Jensen, P.R., Fenical, W., 2007. Induced production of emericellamides A and B from the marine-derived fungus *Emericella* sp. in competing co-culture. J. Nat. Prod. 70, 515–520.

Patron, N.J., Waller, R.F., Cozijnsen, A.J., Straney, D.C., Gardiner, D.M., Nierman, W.C., Howlett, B.J., 2007. Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. BMC Evol. Biol. 7, 174.

Pel, H.J., de Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J., Turner, G., de Vries, R.P., Albang, R., Albermann, K., Andersen, M.R., Bendtsen, J.D., Benen, J.A., van den Berg, M., Breestraat, S., Caddick, M.X., Contreras, R., Cornell, M., Coutinho, P.M., Danchin, E.G., Debets, A.J., Dekker, P., van Dijck, P.W., van Dijk, A., Dijkhuizen, L., Driessen, A.J., d'Enfert, C., Geysens, S., Goosen, C., Groot, G.S., de Groot, P.W., Guillemette, T., Henrissat, B., Herweijer, M., van den Hombergh, J.P., van den Hondel, C.A., van der Heijden, R.T., van der Kaaij, R.M., Klis, F.M., Kools, H.J., Kubicek, C.P., van Kuyk, P.A., Lauber, J., Lu, X., van der Maarel, M.J., Meulenberg, R., Menke, H., Mortimer, M.A., Nielsen, J., Oliver, S.G., Olsthoorn, M., Pal, K., van Peij, N.N., Ram, A.F., Rinas, U., Roubos, J.A., Sagt, C.M., Schmoll, M., Sun, J., Ussery, D., Varga, J., Vervecken, W., van de Vondervoort, P.J., Wedler, H., Wösten, H.A., Zeng, A.P., van Ooyen, A.J., Visser, J., Stam, H., 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat. Biotechnol. 25, 221–231.

Perrin, R.M., Fedorova, N.D., Bok, J.W., Cramer, R.A., Wortman, J.R., Kim, H.S., Nierman, W.C., Keller, N.P., 2007. Transcriptional regulation of chemical diversity in *Aspergillus fumigatus* by LaeA. PLoS Pathog. 3, e50.

Rantala, A., Fewer, D., Hisbergues, M., Rouhiainen, L., Vaitomaa, J., Börner, T., Sivonen, K., 2004. Phylogenetic evidence for the early evolution of microcystin synthesis. Proc. Natl. Acad. Sci. USA 101, 568–573.

Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., Huson, D.H., 2005. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res.* 33, 5799–5808. Available from: http://www-ab.informatik.uni-tuebingen.de/toolbox/.

Reeves, E.P., Reiber, K., Neville, C., Scheibner, O., Kavanagh, K., Doyle, S., 2006. A nonribosomal peptide synthetase (Pes1) confers protection against oxidative stress in *Aspergillus fumigatus*. FEBS J. 273, 3038–3053.

Rokas, A., Payne, G., Fedorova, N.D., Baker, S.E., Machida, M., Yu, J., Georgianna, D.R., Dean, R.A., Bhatnagar, D., Cleveland, T.E., Wortman, J.R., Maiti, R., Joardar, V., Amedeo, P., Denning, D.W., Nierman, W.C., 2007. What can comparative genomics tell us about species concepts in the genus *Aspergillus*? Stud. Mycol. 59, 11–17.

Rokas, A., Galagan, J.E., in press. The *Aspergillus nidulans* genome and a comparative analysis of genome evolution in *Aspergillus*. In: The aspergilli: genomics, medical applications, biotechnology, and research methods. Osmani, S.A., Goldman, G.H., (Eds.), New York: CRC Press.

Schardl, C.L., 2006. A global view of metabolites. Chem. Biol. 13, 5–6.

Schneider, P., Weber, M., Rosenberger, K., Hoffmeister, D., 2007. A one-pot chemoenzymatic synthesis for the universal precursor of antidiabetes and antiviral bis-indolylquinones. Chem. Biol. 14, 635–644.

Schneider, P., Misiek, M., Hoffmeister, D., 2008. In vivo and in vitro production options for fungal secondary metabolites. Mol. Pharm. 5, 234–242.

Schrettl, M., Bignell, E., Kragl, C., Sabiha, Y., Loss, O., Eisendle, M., Wallner, A., Arst Jr., H.N., Haynes, K., Haas, H., 2007. Distinct roles for intra- and extracellular siderophores during *Aspergillus fumigatus* infection. PLoS Pathog. 3, 1195–1207.

Schwecke, T., Göttling, K., Durek, P., Dueñas, I., Käufer, N.F., Zock-Emmenthal, N., Staub, E., Neuhof, T., Dieckmann, R., von Döhren, H., 2006. Nonribosomal Peptide Synthesis in *Schizosaccharomyces pombe* and the architecture of ferrichrome type siderophore synthetases in fungi. Chembiochem 7, 612–622.

Stachelhaus, T., Mootz, H.D., Marahiel, M.A., 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. 6, 493–505.

Stack, D., Neville, C., Doyle, S., 2007. Nonribosomal peptide synthesis in *Aspergillus fumigatus* and other fungi. Microbiology (UK) 153, 1297–1306.

Tamano, K., Sano, M., Yamane, N., Terabayashi, Y., Toda, T., Sunagawa, M., Koike, H., Hatamoto, O., Umitsuki, G., Takahashi, T., Koyama, Y., Asai, R., Abe, K., Machida, M., 2008. Transcriptional regulation of genes on the non-syntenic blocks of *Aspergillus oryzae* and its functional relationship to solid-state cultivation. Fungal Genet. Biol. 45, 139–151.

Tobiasen, C., Ahmann, J., Ravnholt, K.S., Bjerrum, M.J., Grell, M.N., Giese, H., 2007. Nonribosomal peptide synthetase (*NPS*) genes in *Fusarium graminearum, F. culmorum* and *F. pseudograminearium* and identification of NPS2 as the producer of ferricrocin. Curr. Genet. 51, 43–58.

van den Berg, M.A., Albang, R., Albermann, K., Badger, J.H., Daran, J.-M., Driessen, A.J.M., Estrada, C.G., Fedorova, N.D., Harris, D., Heijne, W., Joardar, V., Kiel, J., Kovalchuk, A., Martin, J.F., Nierman, W.C., Nijland, J.G., Pronk, J.T., Roubos, J.A., van der Klei, I., van Peij, N.N.M.E., Veenhuis, M., von Döhren, H., Wagner, C., Wortman, J., Bovenberg, R.A.L., in press. Molecular analysis of a microbial strain improvement paradigm: Genome sequencing and analysis of *Penicillium chrysogenum* Wisconsin 54-1255, Nature Biotechnol.

Wijeratne, E.M., Turbyville, T.J., Zhang, Z., Bigelow, D., Pierson 3rd, L.S., VanEtten, H.D., Whitesell, L., Canfield, L.M., Gunatilaka, A.A., 2003. Cytotoxic constituents of *Aspergillus terreus* from the rhizosphere of Opuntia versicolor of the Sonoran Desert. J. Nat. Prod. 66, 1567–1573.

Zelakowska-Komon, M., Neuhof, T., Dieckmann, R., von Döhren, H., Herrera-Estrella, A., Kubicek, C.P., Druzhinina, I.S., 2007. The formation of atroviridin by *Hypocrea atroviridis* is condition-associated, and positively regulated by blue light and the G-protein GNA3. Eukaryot. Cell 6, 2232–2242.