Comparative analysis of genetic incompatibility in *Aspergillus niger* and *Podospora anserina*

Károly Pál

Promotor:	Prof. Dr. R. F. Hoekstra Hoogleraar in de Genetica Wageningen Universiteit
Co-promotors:	Dr. Ir. A. J.M. Debets Universitair Hoofddocent, Laboratorium voor Erfelijkheidsleer Wageningen Universiteit
	Dr. J. Varga Associate professor, University of Szeged, Hungary CBS Fungal Biodiversity Centre
Promotie Commissie:	Prof. Dr. J. A. van den Berg, Wageningen Universiteit Prof. Dr. Th. W. Kuyper, Wageningen Universiteit Prof. Dr. H.A.B. Wösten, Universiteit Utrecht Dr. R. A. Samson, Centraalbureau voor Schimmelcultures, Utrecht

Dit onderzoek is uitgevoerd binnen de *C.T. de Wit Graduate School*.

Károly Pál

Comparative Analysis of Genetic Incompatibility in *Aspergillus niger* and *Podospora anserina*

Proefschrift

Ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. Dr. M. J. Kropff,
in het openbaar te verdedigen
op woensdag 23 mei 2007
des namiddags te half twee in de Aula

Comparative Analysis of Genetic Incompatibility in *Aspergillus niger* and *Podospora anserina*

PhD-thesis Wageningen University, with references – with summaries in English and Dutch

Károly Pál, 2007

ISBN: 978-90-8504-655-4

Családomnak

Table of contents:

CHAPTER 1

General Introduction

Reproduction is of crucial importance for living organisms, from bacteria to humans. Except for mammals and birds, both sexual and asexual reproduction is known at almost every level of the phylogenetic tree. In this thesis fungal reproduction is in focus and questions related to this are discussed. The emphasis is on heterokaryon incompatibility genes and genes involved in mating, especially in the species *Aspergillus niger* and *Podospora anserina*.
In the coming paragraphs an overview of sexual and asexual reproduction and an overview of the genus *Aspergillus* and the genus *Sordaria* is given, including a detailed description of both *A. niger* and *P. anserina*. At the end of the chapter the outline of the thesis is delineated.

1.1 Sexual reproduction

From the appearance of the first organism until the appearance of mammals, reproduction is the question with highest priority. No life would exist on our planet if organisms were not able to copy themselves and their DNA, the blueprint of themselves. There are two different ways for fungi and many other organisms to produce offspring: clonal and sexual reproduction. In clonal processes, parents are genetically identical with their offspring, therefore both share the same evolutionary history. In contrast to that, during sexual reproduction meiotic recombination may lead to new combinations of traits. Therefore, different parts of the offspring's genome can have a different evolutionary history. Fungi are an exceptional group in terms of reproduction, since they are able to mix different reproductive strategies. In their case clonality is not always mitotic and asexual, for homothallic (self-fertilizing) fungi produce meiospores with an identical genetic background to the parents. At the same time, recombination is not strictly meiotic and sexual, since parental genomes can be mitotically mixed up during the parasexual cycle. In some cases even genetically different parents may simulate clonality via mating and recombination, if they frequently develop from sibling meiospores, like in pseudo-homothallism. Therefore, some authors define fungal clonality as production of offspring genomes that are genetically identical to their parents, and recombination is stated as a process in which the genomes of the progeny are mixtures of genetically different parental genomes (Taylor *et al.* 1999).
Both clonal and sexual reproduction have their advantages and disadvantages when compared to each other. Clonal reproduction generally costs less energy and resources, and is therefore 'cheaper' for the organism. But, deleterious mutations may occur and are inherited by the descendant. In contrast, sexual reproduction is an efficient tool to shuffle two parental genomes to promote the creation and spread of advantageous traits (i.e. parasite resistance) and to purge the genome of deleterious mutations. At the same it is more costly than clonal reproduction (two-fold cost of sex, mating-related costs, recombinational load). However, there are still continuing debates about the origin and advantages of sex (Hurst and Peck 1996, Barton and Charlesworth 1998, Cavalier-Smith 2002, Agrawal 2006, de Visser and Elena 2007)
The sexual processes in animals and plants are governed by megabase sized sexual chromosomes which contain many genes involved in sex determination and identity. In fungi, however, the mating processes are orchestrated by the MAT loci, which size is only a few hundred nucleotides.

Alleles at the MAT loci encode one or two proteins of key importance, which govern *in trans* the expression of sex determining genes on other chromosomes (Fraser and Heitman 2004).
The fungal mating locus evolved into two systems: bipolar and tetrapolar. In a bipolar system cell mating-type identity (generally 'a' and 'α' or 'a+' and 'a-') is defined by two alternative alleles (also referred to as idiomorphs since they do not share homology) of a single MAT locus, and for sexual reproduction mating partners must be of opposite mating-type. This system prevents selfing. In tetrapolar systems mating loci are organized in two unlinked regions in the genome (commonly 'a' and 'b'), and these regions must differ to undergo mating. In this system multiallelism can occur, therefore thousands of mating-types can be generated. In contrast to bipolar systems, tetrapolar systems more strongly promote outbreeding. Basidiomycota fungi may possess either bipolar or tetrapolar mating system, Ascomycota fungi are strictly bipolar (Fraser and Heitman 2004).
Many filamentous fungi are able to reproduce sexually and asexually, depending on the environment. In a rich environment fungi generally produce mitotic spores (anamorph state), but when conditions become unfavourable, they shift to propagate sexually (teleomorph state). There are two main types of fertilization in fungi: homothallic and heterothallic. Homothallic strains are self-fertile, bear both mating-type genes in the same colony (bipolar fungi, e.g. the ascomycete *Aspergillus nidulans*), whereas heterothallic strains are self-sterile and possess only one of the mating-type genes (e.g. the tetrapolar basidiomycete *Ustilago maydis*). In addition, there are many fungal species which reproduce strictly mitotically. Based on the fungal life cycles four categories can be set up: I. heterothallism with mitospores (e.g.. *Magnaporthe grisea*), II. heterothallism with no mitospores (e.g. *Heterobasidion annosum*), III. homothallism and pseudohomothallism with or without mitospores (e.g. *A. nidulans* and *P. anserina*) and IV. Only mitosporic reproduction (e.g. *Aspergillus niger*). Group I-III contain meiosporic taxa, group IV contains mitosporic taxa (Taylor *et al.* 1999).
There are numerous species which are known as only asexually propagating mitosporic strains without sexual cycle. On an evolutionary scale, asexual lineages do not appear to persist for a significant length of time, and asexual taxa presumably are an evolutionary endpoint (Taylor *et al.* 1999).
To know about the reproductive strategy of biologically important fungi (i.e. pathogens) can be crucial for e.g. disease management (Milgroom and Fry 1997). Observation of a high genetic diversity in a supposedly asexual fungus might be explained by meiotic exchanges in the near past or a cryptic sexual state. In addition to the genetical analysis to reveal genetic variability as a sign of meiotic recombination, the recently published fungal genomes open the possibility of data mining and hence *in silico* identification of genes involved in meiotic processes. For example, in different isolates of the opportunistic human pathogen *Aspergillus fumigatus*, which is known as an asexual species, both mating-type genes and meiosis-related genes were identified (Pöggeler 2002, Varga 2003, Dyer and Paoletti 2005, Paoletti *et al.* 2005).

1.2 The parasexual cycle and vegetative (in)compatibility

The lack of a sexual cycle does not mean automatically the lack of recombination. During the vegetative state of its life a fungus is able to undergo hyphal fusion, karyogamy and mitotic recombination, in the so called parasexual cycle (Pontecorvo 1956). Similarly to the mating processes, finding the appropriate partner for mitotic recombination is a crucial aspect of the parasexual cycle. A complicated cellular machinery and network of heterokaryon incompatibility genes are responsible for the acceptation/rejection of partners in parasexuality (Glass *et al.* 2000, Saupe 2000).

There are four steps in this parasexual cycle, which can be distinguished by the gene sets governing the steps (Table 1) (Leslie and Zeller 1996). The initial step, named pre-fusion, is controlled by genes involved in pheromone production and receptors and in heterokaryon self-incompatibility (*hsi* genes). The fusion step, in which the interhyphal anastomoses are formed and results in heteroplasmon, is governed by self/non-self recognition genes, like heterokaryon incompatibility (*het*) (Glass and Kuldau 1992) or vegetative incompatibility (*vic*) genes (Leslie 1993). Effects of *het*-genes can be influenced by modifier (*mod*) genes. After fusion the biochemical pathways leading to non-self recognition and cell death may be influenced by genes including suppressor (*sup*) genes, which can modify the signal. In the final step programmed cell death (PCD) genes initiate processes leading to apoptosis. In our work we investigated the presence/absence of *het*-, *mod*-, *sup*- and PCD genes in *A. niger* and related species (Chapter 2), and mainly focused on the class of *het*-genes truly involved in self/non-self recognition.

Table 1. Gene families involved in steps of vegetative incompatibility. Courtesy of Anne van Diepeningen.

Initiation		Self/Non-self Recognition		Maintenance / Signalling		Rejection
Pre-fusion	⇨	**Fusion**	⇨	**Post-fusion**	⇨	**Apoptosis**
hsi-genes		*het*-genes		*sup*-genes		PCD genes

Examples of similar self/non-self recognition in non-fungal eukaryotes are the S locus regulated sexual self-incompatibility system in plants (Nettancourt 1977), graft rejection in plants (Knox and Clarke 1980), the histo-incompatibility system in animals (Götze 1977, Brown and Eklund 1994) and also tissue rejection after organ transplantation in humans (Häyry *et al.* 1993) or the histo-incompatibility system in animals (Götze 1977, Brown and Eklund 1994). In fungi, self/non-self recognition is mediated by *het*-genes. Theoretically, any locus can be a *het*-gene, at which heteroallelism is not tolerated in a heterokaryon. In some fungi, mating-type genes also function as *het*-genes (Saupe 2000). Allelic and non-allelic incompatibility systems exist. In the first one, an allelic difference between two strains at one *het* locus is sufficient to trigger incompatibility, like it happens between *Neurospora crassa* isolates harboring different alleles of the *het-c* gene (Sarkar *et al.* 2002). In the second case, non-allelic incompatibility is a result of

interaction between two, different loci, like the *het-c*/*het-e* system in *P. anserina* (Saupe 2000).
If two isolates harbour the same set of *het*-genes, they are vegetatively compatible and belong to the same vegetative compatibility groups (VCGs) or heterokaryon compatibility groups (HCG's). The number of *het*-genes in a population defines the number of VCGs, e.g. in a population with 6 bi-allelic genes, in theory at least 2^6 VCGs are possible. Large numbers of *het*-genes have been observed in many fungal species (Coenen 1997): at least 8 and maximum 18 in *A. nidulans* (Anwar *et al.* 1993), 6 in *Cryphonectria parasitica* (Cortesi and Milgroom 1998), at least 11 in *N. crassa* (Glass *et al.* 2000), and 9 in *P. anserina* (Saupe 2000). The strength of the incompatibility reaction depends on the genes involved. In *A. nidulans* partial *het*-genes have been described that cause only weak vegetative incompatibility reactions (Coenen *et al.* 1994).
The function of incompatibility in fungi is not clear yet (Saupe 2000). Incompatibility is known from other eukaryotes, like graft rejection in plants (Knox and Clarke 1980) or tissue rejection after organ transplantation in humans (Häyry *et al.* 1993). There are two theories to explain the existence of these self/non-self recognition systems in fungi. One theory, the so called allorecognition hypothesis, proposes that it is a *bona fide* reaction between genetically different individuals, which may limit the spread of harmful cytoplasmic or nuclear elements (Caten 1972) or prevent resource plundering (Debets and Griffiths 1998). In these examples the role of *het*-genes is to preserve genetic individuality. The second, alternative, theory suggests that *het*-genes simply arose by evolutionary accidence. The existence of genes with dual function, like the *mat-a*/*mat-A* loci in *N. crassa* (Saupe 2000), supports the second theory. These genes sometimes behave as *het*-genes, another time fulfill cellular functions.
There is another, interesting aspect of these hypotheses. If *het*-genes are present to prevent formation of heterokaryons between genetically different individuals - like protection of a local population from 'invaders' -, then evolution acts towards the generation of new alleles, to construct a stable basis to avoid 'pollution' of the isolates belonging to the same VCG. This phenomenon has been studied for more than a decade in the *het-c*/*het-d* and *het-c*/*het-e* non-allelic systems in *P. anserina*. Four alleles of *het-c*, three of *het-d* and four of *het-e* were identified, which only interact with each other in certain combinations (Saupe *et al.* 1994). Recent studies showed that these non-allelic interactions compose a very dynamic system, in which *het-d* and *het-e* are under positive selection and presumably more alleles will be discovered (Paoletti *et al.* unpublished). However, population genetic modelling suggests that selection favouring more than a few alleles is expected to be extremely weak (Nauta and Hoekstra 1993). In contrast to positive selection in non-allelic systems, certain *het* loci are under maintaining/balancing selection. A well studied example is the *N. crassa het-c* (note that this is not homologous to the *P. anserina het-c*), with three known alleles. These alleles show trans-species polymorphism and balancing selection (Wu *et al.* 1998).

1.3 The Aspergilli and the black mould Aspergillus niger

The name of the genus was given by Pietro Antonio Micheli in 1729, when looking into the microscope the shape of conidiophores reminded him of the aspergillum (holy water sprinkler) (Raper and Fennel 1965). *Aspergillus* species are common in soil and plant debris. Macroscopically the strains are quite diverse; their colour varies from white, yellow, blue-green till black, depending on the species and sometimes the age and the morphology of mycelium, conidiophore and conidiospores varies. Microscopically the Aspergilli show some common features: hyphae are septate and hyaline and the conidiophores terminate in a vesicle at the apex. This vesicle is a typical organ for Aspergilli. A conidiophore of the black *A. niger* is shown in Figure 1.

The ascomycete *Aspergillus* genus contains many different fungi, which are mostly saprobic, but human and plant pathogens can be found in this genus, together with strains widely used in industry. So far 114 out of 186 *Aspergillus* species were found to produce only asexual mitotic conidiospores, like *A. flavus*, *A. fumigatus*, *A. niger* and *A. terreus*. The other 72 species produce sexual meiotic ascospores, sometimes next to conidiospores. These sexual Aspergilli mostly belong to nine genera, especially *Emericella*, *Eurotium* and *Neosartorya*. Phylogenetic analysis of 15 species revealed that sexual and asexual Aspergilli are closely related, and meiosis was independently lost and/or gained at least four times in *Aspergillus* among the sampled taxa. In addition, asexual lineages are often recently derived from sexual meiotic ancestors, and data do not support the existence of ancient asexual lineages (Geiser *et al.* 1996).

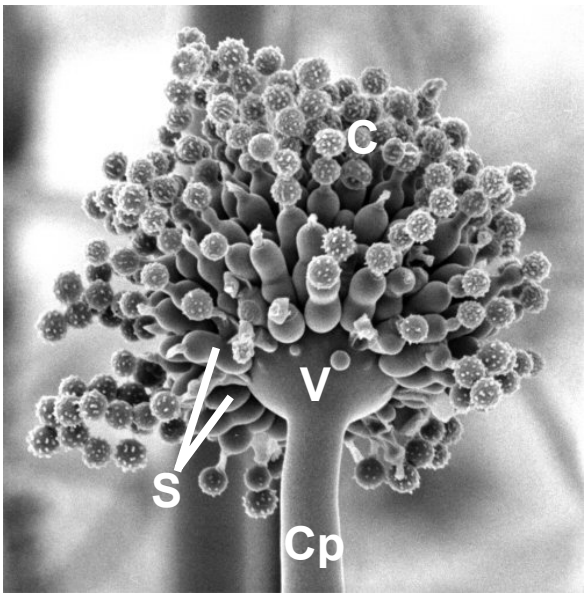


Figure 1. Conidium of *A. niger*. Picture is from the Fungal Cell Biology Group's homepage (Edinburgh University). Legend: Cp: conidiophore; V: vesicle; S: sterigmata; C: conidiospores.

There are about 20 *Aspergillus* species with an important impact on human health. These species can be divided into three groups based on the clinical settings: allergens, opportunistic pathogens and toxin-producers, and occur almost everywhere (i.e. in the air, on crops, nuts or in diary products) (Kovacs 2004)

.

The first black-spored *Aspergillus*, *Aspergillus niger*, was described by P. E. L. van Tieghem in 1867 (in the Annales des Sciences Naturelles, issue 8), as a fungus which is able to utilize the plant polymer tannin as a carbon source. Taxonomic synonyms of *A. niger* are *Sterigmatocystis niger* (Tieghem 1877), *A. fuscus* (Schiemann 1912), *A. niger var. altipes* (Schiemann 1912), *A. schiemanni* (Thom 1916), *A. cinnamomeus* (Schiemann), *A. fuliginosus* (Peck), *A. pyri* (English), and *A. welwitschiae* (Hennings) (MycoBank). Nowadays the black *Aspergillus* group is known to include at least 9 different species sharing the feature of dark brown to black conidiospores: *A. carbonarius, A. japonicus, A. aculeatus, A ellipticus, A. heteromorphus, A. niger, A. tubingensis*, *A. vadense* and *A. brasiliensis* (Kusters-van Someren *et al.* 1991, Mégnégneau *et al.* 1993, DeVries *et al.* 2005). Purely based on their morphology, three main groups can be distinguished: The *A. carbonarius* group, the *A. japonicus*/*A. aculeatus* group and the *A. niger* group (or '*A. niger* aggregate'). Further division of the *A. niger* aggregate was made on the basis of molecular characterization, like Restriction Fragment Length Polymorphisms (RFLPs), Random Amplification of Polymorphic DNA (RAPD) (Varga *et al.* 1993, 1994b, Kevei *et al.* 1996, Hamari *et al.* 1997) and restriction analysis of the ITS (internal transcribed spacer) region (Accensi *et al.* 1999). These analyses resulted in the recognition of three groups within the aggregate, the *A. niger* group, the *A. tubingensis* group and the *A. brasiliensis* group, as can be seen on Table 2 (van Diepeningen 1999). Also the other morphological groups can be further divided based on molecular markers. Some black Aspergilli are known to be opportunistic pathogens, but *A. niger* has the GRAS (generally recognised as safe) status and many strains of it and close relatives are used in industry.

Table 2. Recognisable species within the group of black Aspergilli on basis of morphology, ribosomal and mitochondrial RFLP data. [a]) Raper and Fennel, 1965. [b]) Kusters-van Someren *et al.*1991; Mégnégneau *et al.*, 1993. [c]) Kevei *et al.*, 1996. [d]) Varga *et al.*, 1993; 1994a. [e]) Hamari *et al.*, 1997. Courtesy of Anne van Diepeningen.

black Aspergilli					
morphology	*A. carbonarius*	*A. niger aggregate*			*A. japonicus* (Saito) / *A. aculeatus* (Iizuka)
sterigmata[a]:	biseriate	biseriate			single
conidia[a]:	large, multinucleate	relatively small			echinulate
RFLP		*A. niger*	*A. tubingensis*	*A. brasiliensis*	
rRFLP	C1-C2[bc]	I-I'[b]	II-II'[b]		J[be]
mtRFLP	C1[a-b] & C2[c]	1a-1e[d]	2a-2f[d]	3[d]	J1-J7[e]

Though most *Aspergillus* species are supposedly asexual, recent findings suggest that asexual species, like *A. fumigatus*, might have a cryptic sexual cycle (Varga 2003, Paoletti *et al.* 2005). Lack of sex may force a fungus to use an alternate pathway of recombination, the so called parasexual cycle (Figure 2). In the sexual life cycle reproduction is based on meiotically

produced ascospores, whereas in the asexual life cycle mitotically produced conidiospores are responsible for inheritance. If strains are heterokaryon compatible they may fuse to form heterokaryons and heteroplasmons in which recombination may occur (parasex). However, natural isolates of the *A. niger* aggregate show a high degree of incompatibility, strains are mostly unable to form heterokaryons in the parasexual cycle (van Diepeningen *et al.* 1997). This incompatibility blocks the transfer of genetic elements between different isolates and can be visualised as the inablility to form heterokaryons on selective media (van Diepeningen *et al.* 1997). Barrages as seen in *P. anserina* or the formation of aerial hyphae are not visible in these incompatible combinations.

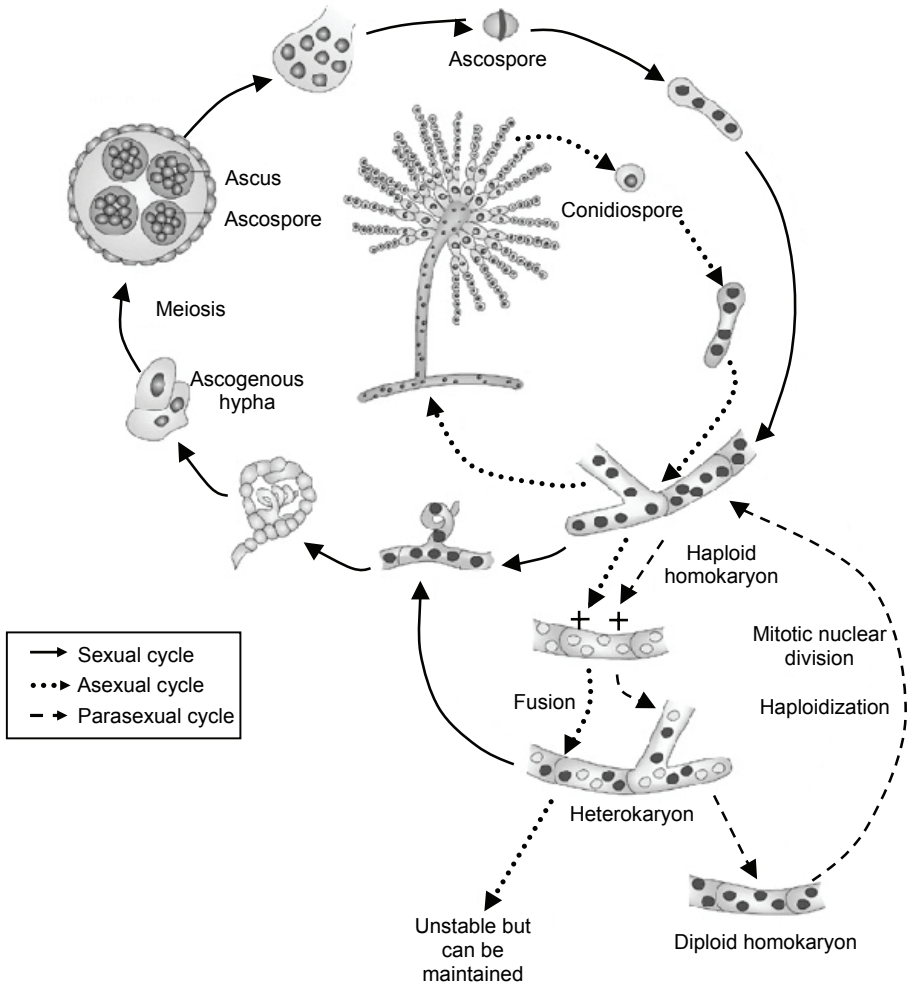


Figure 2. Life cycle of *A. nidulans*, as a model to represent sexual and asexual reproduction in Aspergilli. Picture was taken from Casselton and Zolan (2002).

Thus, the high degree of incompatibility brings up many questions, like: how is recombination possible between different isolates when both the sexual and parasexual cycles are missing or blocked? What advantages for *A. niger* and other Aspergilli are derived from the lack of recombination? Is there a cryptic sexual cycle or is sexuality for *A. niger* lost permanently? If sexuality is lost, when did it happen?

1.4 Sordariales and Podospora anserina

The order of the *Sordariales* is also a part of the Ascomycota phylum. Eight families belong to this order, but only a few of them are able to form asexual offspring. Typical, common features of the species belonging to this order are the presence of perithecia with a dark colour, and non-amyloid asci with an apical ring ("sphincter"), except for the family *Chaetomium*. The ascospores are amerosporous (1-celled). Species belonging to the *Sordariales* genus are mostly saprobic or coprophilic, and pathogenicity is very rare.

P. anserina belongs to the Lasiosphaeriaceae family and more particularly to the *Podospora* genus. The origin of the name is unknown, presumably the first time this species was published as *P. anserina* was in the journal Hedwigia (1883, issue 22), by Niessl. Synonyms are: *Sphaeria pauciseta* (Cesati 1852), *Malinvernia anserina* (Rabenhorst 1857), *Sordaria anserina* (Winter 1873), and *Pleurage anserina* (Kuntze) (MycoBank). *P. anserina* is a heterothallic or so-called pseudo-homothallic coprophilic species, growing on the dung of herbivores (Figure 3). The mycelia are usually dikaryotic, containing nuclei of both mating-types, although monokaryotic strains can be found as well. *P. anserina* has four ordered ascospores per ascus and is a commonly used model organism in fungal genetics. *P. anserina* does not have an asexual state.

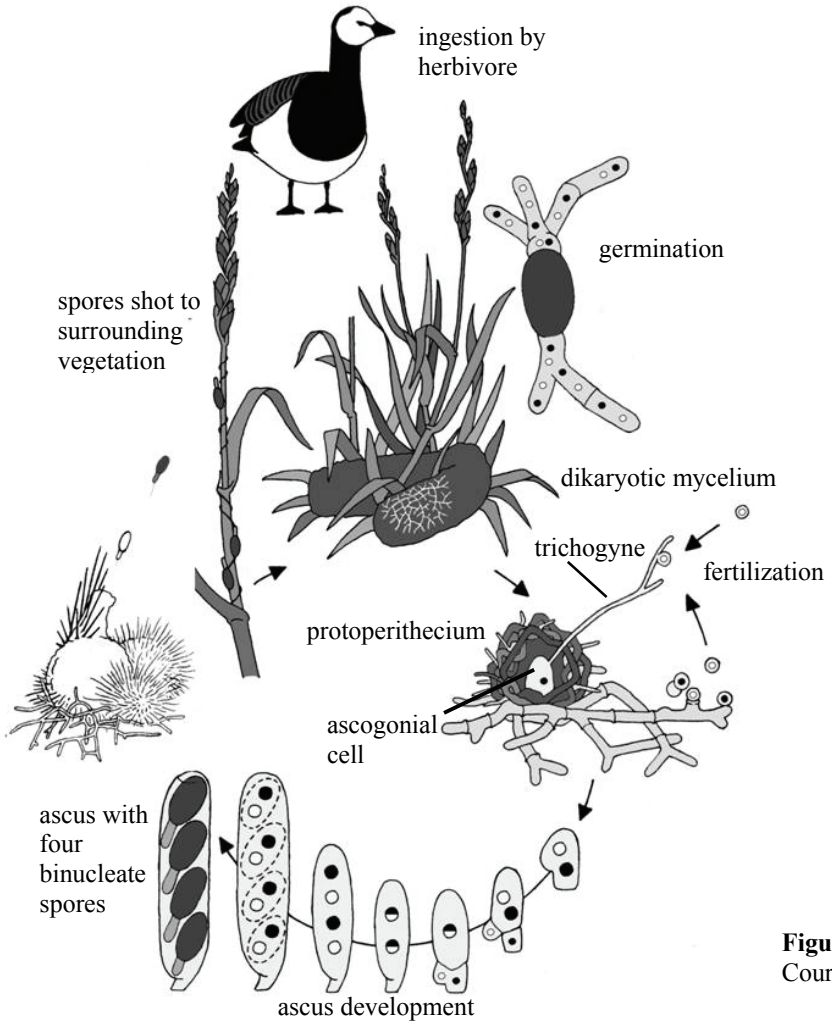


Figure 3. Life cycle of *P. anserina*. Courtesy of Marc F.P.M. Maas.

The fungus *Podospora anserina* is a model system for heterokaryon incompatibility. Both allelic (involving one locus) and non-allelic (involving several loci) heterokaryon incompatibility genes have been found. *P. anserina* bears at least nine *het* loci consisting of both types of incompatibility genes (Saupe 2000). Unlike in *A. niger*, an incompatibility reaction is macroscopically visible when two incompatible mycelia meet, as so called barrages (Figure 4). The *P. anserina* genome was recently sequenced (*Podospora anserina* Genome Project).

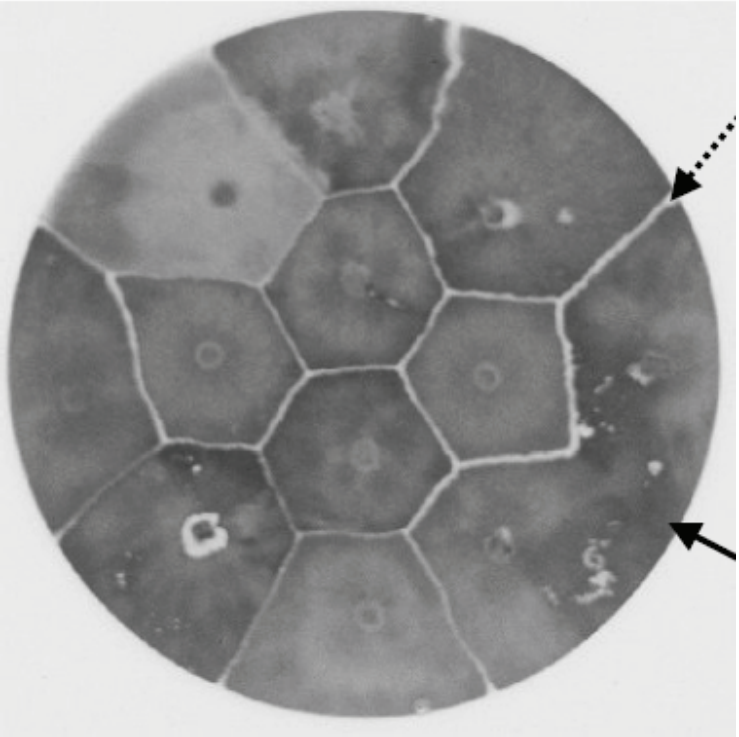


Figure 4. Barrage formation between different isolates of *P. anserina*. Top arrow: an incompatible reaction. Lower arrow: a compatible combination of strains shows no barrage. Picture courtesy of van der Gaag, 2005.

1.5 Outline of this thesis

The main aim of this study is to expand our knowledge of heterokaryon incompatibility in a sexual and an asexual filamentous ascomycete and of the mating capabilities of a species known as asexual. It deals with the *in silico* comparative analysis of fungal genomes, practical analysis of incompatibility related genes in numerous natural fungal isolates, but also with the experimental manipulation of a gene involved in incompatibility. The work is done with the ascomycetous species *Podospora anserina* (sexual) and *Aspergillus niger* (asexual).

Chapter 2: In this chapter a comparative analysis is made of eight fungal genomes, including both filamentous and yeast-like Ascomycetes. The aim was to find similarities and dissimilarities in incompatibility and apoptosis related genes between the different fungal genomes. This work was done with *in silico* analysis, which is an efficient tool to get a better look into the genetic background of vegetative incompatibility in *Aspergillus* species.

Chapter 3: Analysis of vegetative incompatibility in the black mould *A. niger* is impossible with the classical techniques. Natural isolates of this fungus are highly incompatible with each other, therefore recombination assays are not feasible. The *het-c* gene is a well studied locus in *N. crassa* (with 3 alleles) and *P. anserina* (monomorphic), and we chose this locus to investigate in *A. niger*. The *A. niger het-c* homolog was searched and found in the recently sequenced genome

(Pel *et al.* 2007). Based on the sequence data, primers were designed to reveal whether polymorphism was present or not in nearly 100 black *Aspergillus* isolates. A functional test was performed with the *N. crassa het-c* alleles to find out if *het-c* in *A. niger* is functional or not.

Chapter 4: The *P. anserina hch* and *het-c* loci were subjects of studies on vegetative incompatibility in earlier studies. The *hch* is the homolog of the *N. crassa het-c* gene, but in contrast to *N. crassa* no polymorphism was found in 11 *P. anserina* isolates (Saupe *et al.* 2000). In this study *P. anserina* samples were collected in France. We wanted to see whether geographical location has any influence on the polymorphism of the *hch*, therefore we performed CAPS analysis on all of our Dutch isolates. Next to that we partially sequenced the *hch* from 11 random isolates. Our results support the previous study in which no polymorphism was found with sequencing or with CAPS analysis.
The *het-c* gene forms a non-allelic recognition system with the *het-d* and *het-e* loci in *P. asnerina*. A recent study suggests that these loci are under positive selection and form a dynamic self/non-self recognition system. We sequenced 15 isolates from our collection and found 8 new alleles, which is strong evidence in favour of the hypothesis.

Chapter 5: In this chapter we performed a comparative analysis on eight ascomycete fungi genomes, including sexual and asexual species on mating type and mating related genes. Our aim was to find out whether there are sex related genes present in *presumed* asexual species or not and whether these could be functional. We compared the gene sets and proteins of two *A. niger* in detail. In the same chapter we tested 160 natural black *Aspergillus* strains for the presence/absence of *mat-1* and *mat-2* genes.

Chapter 6: In the general discussion and summary the most important findings of this thesis are reviewed, and suggestions for further research are given.

References

Accensi F, Cano J, Figuera L, Abarca ML, Cabañes FJ, 1999. New PCR method to differentiate species in the *Aspergillus niger* aggregate. *FEMS Microbiology Letters* 180, 191-196.

Agrawal AF, 2006. Evolution of sex: why do organisms shuffle their genotypes? *Current Biology* 16, 696-704.

Anwar, MM, Croft JH, Dales RB, 1993. Analysis of heterokaryon incompatibility between heterokaryon-compatibility (h-c) groups R and GL provides evidence that at least eight *het* loci control somatic incompatibility in *Aspergillus nidulans*. *Journal of General Microbiology* 139, 1599-1603.

Barton NH, Charlesworth B, 1998. Why sex and recombination? *Science* 281, 1986-1990.

Brown JL, Eklund A, 1994. Kin recognition and the major histocompatibility complex: an integrative review. *American Naturalist* 143, 5435-461.

Casselton L, Zolan M, 2002. The art and design of genetic screens: filamentous fungi. *Nature Reviews Genetics* 3, 683-697.

Caten CE, 1972. Vegetative incompatibility and cytoplasmic infection in fungi. *Journal of General Microbiology* 72, 221-229.

Cavalier-Smith T, 2002. Origins of the machinery of recombination and sex. *Heredity* 88, 125-141.

Coenen A, Debets F, Hoekstra RF, 1994. Additive action of partial heterokaryon incompatibility (*partial-het*) genes in *Aspergillus nidulans*. *Current Genetics* 26, 233-237.

Coenen A, 1997. The transmission of cytoplasmic genes in *Aspergillus nidulans*. PhD Thesis, Chapter 1. Wageningen University, The Netherlands.

Cortesi P, Milgroom MG, 1998. Genetics of vegetative incompatibility in *Cryphonectria parasitica*. *Applied and Environmental Microbiology* 64, 2988–2994.

de Visser JGM, Elena SF, 2007. The evolution of sex: empirical insights into the roles of epistasis and drift. *Nature Reviews Genetics* 8, 139-149.

de Vries RP, Frisvad JC, van de Vondervoort PJL, Burgers K, Kuijpers AFK, Samson RA, Visser J, 2005. *Aspergillus vadense*, a new species of the group of black Aspergilli. *Antonie van Leeuwenhoek* 87, 195–203.

Debets AJM, Griffiths AJF, 1998. Polymorphism of *het*-genes prevents resource plundering in *Neurospora crassa*. *Mycological Research* 102, 1343-1349.

Doctor Fungus: http://www.doctorfungus.org/index.htm

Dyer PS, Paoletti M, 2005. Reproduction in *Aspergillus fumigatus*: sexuality in a supposedly asexual species? *Medical Mycology Supplement* 1, 43, S7_/S14)

Fraser AJ, Heitman J, 2004. Evolution of fungal sex chromosomes. *Molecular Microbiology* 51, 299-306.

Fungal Cell Biology Group, Edinburgh University: http://129.215.156.68/

Geiser DM, Timberlake WE, Arnold ML, 1996. Loss of meiosis in *Aspergillus*. *Molecular Biology and Evolution* 13, 809-817.

Glass NL, Kuldau GA, 1992. Mating type and vegetative incompatibility in filamentous ascomycetes. *Annual Review of Phytopathology* 30, 201–224.

Glass NL, Jacobson DJ, Shiu PKT, 2000. The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycete fungi. *Annual Review of Genetics* 34, 165-186.

Götze D, 1977. The major histocompatibility system in man and animals. Springer-Verlag, Berlin.

Hamari Z, Kevei F, Kovács É, Varga J, Kozakiewicz Z, Croft JH, 1997. Molecular and phenotypic characterization of *Aspergillus japonicus* strains with special regard to their mitochondrial DNA polymorphisms. *Antonie van Leeuwenhoek* 72, 337-347.

Häyry P, Isoniemi H, Yilmaz S, Mennander A, Lemström K, Räisänen-Sokolowski A, Koskinen P, Ustinov J, Lautenschlager I, Taskinen E, Krogerus L, Aho P, Paavonen T, 1993. Chronic allograft rejection. *Immunological Reviews* 134, 33-81.

Hurst LD, Peck JR, 1996. Recent advances in understanding of the evolution and maintenance of sex. *Trends in Ecology and Evolution* 11, 46-52.

Kevei F, Hamari Z, Varga J, Kozakiewicz Z, Croft JH, 1996. Molecular polymorphism and phenotypic variation in *Aspergillus carbonarius*. *Antonie van Leeuwenhoek* 70, 59-66.

Knox RB, Clarke AE, 1980. Discrimination of self and non-self in plants. In: *Contemporary topics in immunobiology* Volume 9, J.J. Marchalonis & N. Cohen (eds.) Plenum press, New York, 1-36.

Kovacs M, 2004. Nutritional health effects of mycotoxins. *Orvosi hetilap 145*, 1739-1746.

Kusters-van Someren MA, Samson RA, Visser J, 1991. The use of RFLP analysis in classification of the black *Aspergilli*: reinterpretation of *Aspergillus niger* aggregate. *Current Genetics* 19, 21-26.

Leslie JF, 1993. Fungal vegetative incompatibility. *Annual Review of Phytopathology* 31, 127–150.

Leslie JF, Zeller KA, 1996. Heterokaryon incompatibility in fungi – more than just another way to die. *Journal of Genetics* 75, 415-424.

Mégnégneau B, Debets F, Hoekstra RF, 1993. Genetic variability and relatedness in the complex Group of black *Aspergilli* based on random amplification of polymorphic DNA. *Current Genetics* 23, 323-329.

Milgroom MG, Fry WE, 1997. Contributions of population genetics to plant disease epidemiology and management. *Advances in Botanical Research* 24, 1-30.

MycoBank: http://www.mycobank.org/DefaultPage.aspx

Nauta MJ, Hoekstra RF, 1993. Evolutionary dynamics of spore killers. *Genetics* 135, 923-930.

Nettancourt D, 1977. Incompatibility in angiosperms. Springer-Verlag, Berlin.

Paoletti M, Rydholm C, Schwier EU, Anderson MJ, Szakacs G, Lutzoni F, Debeaupuis JP, Latgé JP, Denning DW, Dyer PS, 2005. Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Current Biology* 15, 1242–1248.

Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, Andersen MR, Bendtsen JD, Benen JAE, van den Berg M, Breestraat S, Caddick MX, Contreras R, Cornell M, Coutinho PM, Danchin EGJ, Debets AJM, Dekker P, van Dijck PWM, van Dijk A, Dijkhuizen L, Driessen AJM, d'Enfert C, Geysens S, Goosen C, Groot GSP, de Groot PWJ, Guillemette T, Henrissat B, Herwijer M, van den Hombergh JPTW, van den Hondel CAMJJ, van der Heijden RTJM, van der Kaaij RM, Klis FM, Kools HJ, Kubicek CP, van Kuyk PA, Lauber J, Lu X, van der Maarel MJEC, Meulenberg R, Menke H, Mortimer AM, Nielsen J, Oliver SG, Olsthoorn M, Pal K, van Peij NNME, Ram AFJ, Rinas U, Roubos JA, Sagt CMJ, Schmoll M, Sun J, Ussery D, Varga J, Vervecken W, van den Vondervoort PJI, Wedler H, Wösten HAB, Zeng A-P, van Ooyen AJJ, Visser J, Hein S, 2007. Genome sequence of *Aspergillus niger* strain CBS 513.88: a versatile cell factory. *Nature Biotechology* 25, 221-231.

Podospora anserina Genome Project: http://podospora.igmors.u-psud.fr/

Pontecorvo G, 1956. The parasexual cycle in fungi. *Annual Review of Microbiology* 10, 393-400.

Pöggeler S, 2002. Genomic evidence for mating abilities in the asexual pathogen *Aspergillus fumigatus*. *Current Genetics* 42, 153-160.

Raper KB and Fennel DI, 1965. The genus *Aspergillus*. The Williams & Wilkins Company, Baltimore.

Sarkar S, Iyer G, Wu J, Glass NL, 2002. Nonself recognition is mediated by HET-C heterocomplex formation during vegetative incompatibility. *EMBO Journal* 21, 4841-4850.

Saupe S, Descamps C, Turcq B, Bégueret J, 1994. Inactivation of the *Podospora anserina* vegetative incompatibility locus *het-c*, whose product resembles a glycolipid transfer protein, drastically impairs ascospore production. *Proceedings of the National Academy of Sciences of the USA* 91, 5927-5931.

Saupe SJ, 2000. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews* 64, 489-502.

Saupe SJ, Clavé C, Sabourin M, Bégueret J, 2000. Characterization of *hch*, the *Podospora anserina* homolog of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa*. *Current Genetics* 38, 39-47.

Université de Paris-Sud XI / CNRS: http://podospora.igmors.u-psud.fr/

Taylor JW, Jacobson DJ, Fisher MC, 1999. The evolution of asexual fungi: reproduction, speciation and classification. *Annual Review of Phytopathology* 37, 197-246.

van Diepeningen AD, Debets AJM, Hoekstra RF, 1997. Heterokaryon incompatibility blocks virus transfer among natural isolates of black *Aspergilli*. *Current Genetics* 32, 209-217.

van Diepeningen AD, 1999. Horizontal transfer of genetic elements in the black *Aspergilli*. PhD Thesis. Wageningen University, The Netherlands.

van der Gaag M, 2005. Genomic conflicts in *Podospora anserina*. PhD Thesis, Wageningen University, The Netherlands.

Varga J, Kevei F, Fekete C, Coenen A, Kozakiewicz Z, Croft JH, 1993. Restriction fragment length polymorphisms in the mitochondrial DNAs of the *Aspergillus niger* aggregate. *Mycological Research* 97, 1207-1212.

Varga J, Kevei F, Vagvolgyi C, Vriesema A, Croft JH, 1994a. Double-stranded RNA mycoviruses in section Nigri of the *Aspergillus* genus. *Canadian Journal of Microbiology* 49, 325-329.

Varga J, Kevei F, Vriesema A, Debets F, Kozakiewicz Z, Croft JH, 1994b. Mitochondrial DNA restriction fragment length polymorphisms in field isolates of the *Aspergillus niger* aggregate. *Canadian Journal of Microbiology* 40, 612-621.

Varga J, 2003. Mating type gene homologues in *Aspergillus fumigatus*. *Microbiology* 149, 816-819.

Wu J, Saupe SJ, Glass NL, 1998. Evidence for balancing selection operating at the *het-c* heterokaryon incompatibility locus in a group of filamentous fungi. *Proceedings of the National Academy of Sciences of the USA* 95, 12398-12403.

CHAPTER 2

Heterokaryon incompatibility genes in *Aspergillus niger*

Authors:

Károly Pál[1, 2], Anne D. van Diepeningen[1], János Varga[2, 3], Rolf F. Hoekstra[1]

and Alfons J.M. Debets[1]

[1]) Laboratory of Genetics, Plant Sciences, Wageningen University, Wageningen, The Netherlands
[2]) University of Szeged, Faculty of Science, Dept. of Microbiology, Szeged, H-6701 Hungary [3])
CBS Fungal Biodiversity Centre, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands.

Abstract

In this study we present the results of data mining for heterokaryon incompatibility (*het*) and programmed cell death (PCD) related genes in the genomes of two heterokaryon incompatible isolates of the asexual species *Aspergillus niger*. *Het*-genes regulate the formation of anastomoses and heterokaryons, may protect resources and prevent the spread of infectious genetic elements. Depending on the *het* locus involved, hetero-allelism is not tolerated and fusion of genetically different individuals leads to growth inhibition or cell death. The high natural level of heterokaryon incompatibility in *A. niger* blocks parasexual analysis of the *het*-genes involved, but *in silico* experiments in the sequenced genomes allow us to identify putative *het*-genes. Homologous sequences to known *het*- and PCD-genes were compared between different sexual and asexual species including different *Aspergillus* species, *Sordariales* and the yeast *Saccharomyces cerevisiae*. Both *het*- and PCD-genes were well conserved in *A. niger*. However some point mutations and other small differences between the *het*-genes in the two *A. niger* isolates examined may hint to functions in heterokaryon incompatibility reactions.

Keywords: self/non-self recognition, heterokaryon incompatibility genes, apoptosis, ascomycetes, *Aspergillus niger*, *Aspergillus fumigatus*, *Aspergillus nidulans*, *Podospora anserina*, *Neurospora crassa*, *Saccharomyces cerevisiae*

1. Introduction

For filamentous fungi the establishment of hyphal anastomoses, both within and between individuals of the same species, is considered to be of high importance. A limitation to intermycelial fusions is heterokaryon incompatibility which is widespread among fungi and prevents the coexistence of genetically dissimilar nuclei within a common cytoplasm. The adaptive significance of heterokaryon incompatibility is unclear but it may serve to limit the spread of detrimental cytoplasmic or nuclear elements (Caten 1972) or prevent resource plundering (Debets and Griffiths 1998). The genes that control heterokaryon (or vegetative) incompatibility are termed *het* (for heterokaryon incompatibility; Glass and Kuldau 1992) or *vic* (for vegetative incompatibility; Leslie 1993) genes. A *het* locus can be any locus, at which heteroallelism is not tolerated in a heterokaryon (Saupe 2000). When heterokaryon-incompatible strains fuse, the resulting heterokaryotic hyphae are rapidly compartmentalized and destroyed (often with surrounding cells) or seriously inhibited in their growth, depending on the involved *het* locus. Heterokaryotic cells are often destroyed within 30 minutes after hyphal fusion. The process of destruction of hyphal compartments shows similarity at the microscopic level in different fungi, and some steps have common features even with multicellular metazoan programmed cell death (PCD) (Glass and Kaneko 2003).

Since heterokaryon incompatible strains can be sexually compatible, the number of *het*-genes that segregate from a cross can be deduced from the progeny. Genetic analysis of heterokaryon incompatibility was performed in a few sexual fungi. The number of identified *het*-genes varied with species. There are at least eight and maximum 18 *het* loci in *A. nidulans* (Anwar *et al.* 1993), six in *Cryphonectria parasitica* (Cortesi and Milgroom 1998), at least 11 in *N. crassa* (Glass *et al.* 2000) and nine in *P. anserina* (Saupe 2000). The majority of the incompatibility reactions is regulated by allelic systems, where two (e.g. *mat-a*/*mat-A* in *N. crassa*, Saupe 2000) or more alleles (e.g. *het-c* Groveland, Oakridge and Panama alleles in *N. crassa*, Sarkar *et al.* 2002) of the same locus interact. In other cases, two distinct loci trigger non-allelic incompatibility (e.g. *het-c*/*het-e* in *P. anserina*, Saupe 2000). In non-allelic systems, incompatible alleles can be present in the same haploid nucleus in the progeny and thus vegetative incompatibily may occur also in homokaryotic cells, like in the *het-r*/*het-v* incompatibility in *P. anserina*. Such homokaryotic strains can be obtained for each non-allelic system, and are named self-incompatible (SI) strains (Bourges *et al.* 1998).

Unlike the ascomycetes mentioned above, which have a sexual and usually also an asexual reproduction cycle, for asexual fungi like the black mould *A. niger*, the only way to achieve (mitotic) recombination is via the parasexual cycle. Mitotic recombination has been used for genetic analysis of mutants in an isogenic background of *A. niger* and the construction of a genetic map (Debets *et al.* 1993). But, natural isolates of black *Aspergilli* are highly incompatible with each other, efficiently blocking virus transfer as well as the formation of heterokaryons (van Diepeningen *et al.* 1997). As a result, mitotic recombination between genetically dissimilar

isolates is also blocked, so genetic analysis cannot reveal the genetic basis of heterokaryon incompatibility in *A. niger*.
In this study we present the results of data mining in the genomes of two heterokaryon incompatible isolates of the asexual *A. niger* species. We compared incompatibility and cell death related proteins of the two *A. niger* isolates with each other and similar proteins of related sexual (*A. nidulans*) and asexual Aspergilli (*A. fumigatus*[1], *A. oryzae, A. terreus*), two members of the *Sordariales* (*P. anserina*, *N. crassa)* and the yeast *Saccharomyces cerevisiae*. Our analyses identified the major putative *het*-genes in the genome of *A. niger* and related *Aspergilli*. These findings can be used for further functional analysis of candidate *het*-genes.

2. Materials and methods

2.1 Heterokaryon compatibility tests

DSM Research BV's Strain CBS 513.88, DOE Joint Genome Institute's culture collection strain ATCC1015 and our lab strain N400 (ATCC 9029; CBS 120.49) were tested for heterokaryon (in)compatibility. Therefore, pairs of strains distinguishable in colour (fawn or black coloured conidiospores) and with complementing deficiency mutations (e.g. arginine, lysine and others) or dominant resistancies like oligomycine resistance were inoculated both individually and together on Minimal Medium (Pontercorvo 1953) without supplementing for deficiencies and with antibiotics if necessary. Only in compatible combinations of strains the mutations can be complemented in the formed heterokaryon, the single strains don't grow (van Diepeningen *et al.* 1997).

2.2 Het- and apoptosis genes

A list of known genes involved in either programmed cell death from *S. cerevisiae* or involved in heterokayon incompatibility and/or programmed cell death in *N. crassa* or *P. anserina* was constructed based on the literature (Table 1). Most of these genes were also listed in the article of Glass and Kaneko (2003). All these listed genes were tested in this study. The presumed functions and references for these genes are given in Table 1.

[1] Traces of cryptic sexual cycle were identified in the *A. fumigatus* genome (Varga 2003) and further molecular evidences were found (Paoletti *et al.* 2005), but functional analysis has not yet been performed.

2.3 Identification of HET homolog sequences in the A. niger databases

We searched two *A. niger* databases: the genome of the culture collection strain CBS513.88 from DSM Research BV (prepared by shotgun sequencing, 7.5 times coverage) and the genome of the culture collection strain ATCC1015 from DOE Joint Genome Institute (assembly release version 1.0 of whole genome shotgun reads was constructed with the JGI assembler, Jazz, using paired end sequencing reads at a coverage of 8.9X) (Table 2). These strains were tested in our lab and found to be heterokaryon incompatible with one another.

For blastp searches we used the apoptosis-like PCD proteins of *S. cerevisiae* and the HET, modifier and suppressor protein sequences of *N. crassa* and *P. anserina* (Table 1). For validation of the identified *A. niger* sequences, a bi-directional best hit analysis was performed, using the polypeptide sequence of the identified *A. niger* ORFs as a query for a blastp search at the *N. crassa*, *P. anserina S. cerevisiae* and GenBank database (http://ncbi.nih.gov/BLAST; Altschul *et al.* 1990). As criterion for homologs we used an accepted E-value of $<e^{-10}$.

2.4 Identification of HET homolog sequences in other Aspergilli

In turn the *S. cerevisiae* PCD- and *N. crassa* and *P. anserina* HET-, modifier and suppressor protein sequences were used to search the genomes of *A. fumigatus* Af 293, *A. nidulans* FGSC A4, *A. oryzae* RIB40, and *A. terreus* NIH2624 for homologs (Table 2). Again bi-directional best hit analyses were performed with as criterion for homologs an accepted E-value of $<e^{-10}$.

2.5 Genome databases used for BLAST search

A. fumigatus preliminary sequence data was obtained from The Institute for Genomic Research website. The *A. nidulans*, *N. crassa* (release 7) and *A. terreus* sequence data were from the *Aspergillus nidulans*, *Neurospora crassa* and *Aspergillus terreus* Sequencing Projects, Broad Institute of MIT and Harvard. The *A. oryzae* sequences were available on the server of National Institute of Technology and Evaluation (NITE). The *P. anserina* genome was published by the Institut de Génétique et Microbiologie (Université de Paris-Sud XI / CNRS). *A. niger* ATCC1015 sequence data were produced by the US Department of Energy Joint Genome Institute and *A. niger* CBS513.88 sequence data by the DSM Research BV (Table 2).

2.6 Comparison of homologs of different origin

Homologous polypeptides were aligned with ClustalX (Thompson *et al.* 1997) and edited with Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html). Neighbour-joining phylogenetic trees were also made with ClustalX and displayed with TreeView 1.6.6 (Page, 1996).

Table 1. Genes used in this study involved in heterokaryon incompatibility in *N. crassa* and *P. anserina* and in Programmed Cell Death (PCD) in *S. cerevisiae*. All genes are given with their presumed function and references. ID numbers for *N. crassa* and *S. cerevisiae* proteins refer to the numbers given in their respective sequencing projects, the ID numbers for the *P. anserina* proteins were taken from Genbank. The table is an expanded version of the table in Glass and Kaneko (2003).

Species	Class	Gene	Function	References
N. crassa				
	Heterokaryon incompatibility genes			
		het-6^{OR} (NCU03533.2)	allelic *het*-gene, TOL/HET-6/HET-E domain	Saupe 2000
		het-c (NCU3125.2)	allelic *het*-gene, signal peptide	Sarkar *et al.* 2002 Saupe *et al.* 2000
		un-24 (NCU03539.2)	allelic *het*-gene, ribonucleotide reductase large subunit	Saupe 2000 Smith *et al.* 2000
	Suppressor genes			
		tol (NCU04453.2)	TOL/HET-6/HET-E domain	Shiu and Glass 1999
		vib-1 (NCU03725.2)	regulation of conidiation and maybe of nrAPase	Xiang and Glass 2002
	Incompatibility related genes			
		ham-2 (NCU03727.2)	hyphal fusion, putative transmembrane protein	Xiang *et al.* 2002
		pin-c (NCU03494.2)	allelic gene with HET-domain, linked to *het-c*	Kaneko et al 2006
		rnr-a (NCU07887.2)	suppresses *un-24* temperature sensitive mutation, ribonucleotide reductase small subunit	Kotierk,M. and Smith,M.L 2001
P. anserina				
	Heterokaryon incompatibility genes			
		het-c2 (AAA20542)	nonallelic *het*-gene interacts with *het-d* and *het-e*, glycolipid transfer protein	Saupe *et al.* 1994
		het-d2y (AAL37301)	nonallelic *het*-gene against *het-c2*, GTP-binding, WD repeat, TOL/HET-6/HET-E domain	Espagne *et al.* 2002
		het-e4s (AAL37297)	nonallelic *het*-gene against *het-c2*, GTP-binding, WD repeat, TOL/HET-6/HET-E domain	Espagne *et al.* 2002
		het-S (AAB88771)	allelic *het*-gene, prion analog	Coustou *et al.* 1997
	Incompatibility related genes			
		idi-1 (AAC24119)	induced by *het-c/e* and *r/v* incompatibility, signal peptide	Dementhon *et al.* 2003 Bourges *et al.* 1998
		idi-2 (AAC24120)	induced by *het-r/v* incompatibility, signal peptide	Bourges *et al.* 1998
		idi-3 (AAC24121)	induced by *het-c/e* and *r/v* incompatibility, signal peptide	Bourges *et al.* 1998
		idi-4 (*jlb-a*) (AAT40415)	bZIP motif, putative trans-activation domain	Dementhon *et al.* 2004, 2005
		idi-6 (*pspA*) (AAC03564)	induced by *het-c/e* and *r/v* incompatibility, subtilisin-like serine protease	Paoletti *et al.* 2001 Reichard *et al.* 2000
		idi-7 (AAN41258)	ortholog of the *S. cerevisiae aut7p*	Pinan-Lucarre *et al.* 2003
	Modifier genes			
		mod-A (AAC25496)	modifier of *het-c*/*e*, *c*/*d* and *r*/*v* incompatibility, SH3-binding motif	Barreau *et al.* 1998 Bourges *et al.* 1998
		mod-D (AAC24766)	modifier of *het-c*/*e* incompatibility, G protein α subunit	Loubradou *et al.* 1999
		mod-E (AAB97626)	modifier of *het-r*/*v* incompatibility, HSP90	Loubradou *et al.* 1997
S. cerevisiae				
	Programmed Cell Death genes			
		atp4 (YPL078C)	F_0F_1-ATPase	Matsuyama *et al.* 1998
		cdc48 (YDL126C)	cell division cycle, AAA ATPase, fusion of ER-derived vesicles	Madeo *et al.* 1997
		hel10 (YNL208W)	unknown	Ligr *et al.* 2001
		hel13 (YOR309C)	unknown	Ligr *et al.* 2001
		mca1/*yca1* (YOR197W)	metacaspase	Madeo *et al.* 2002
		nsr1 (YGR159C)	rRNA processing	Ligr *et al.* 2001
		ppa1 (YHR026W)	vacuolar H^+-ATPase	Ligr *et al.* 2001
		rsm23 (YGL129C)	mitochondrial small ribosomal unit	Madeo *et al.* 2002
		sar1 (YPL218W)	ER to Golgi transport	Ligr *et al.* 2001
		stm1 (YLR150W)	suppressor of *pop2* and *tom2*	Ligr *et al.* 2001
		tor1 (YJR066W)	regulation of cell death, phosphatidylinositol 3-kinase	Rohde *et al.* 2001 Dementhon *et al.* 2003 Fitzgibbon *et al.* 2005

Table 2. Genome databases and their websites used in this research.

Species	Strain	Website	References
A. fumigatus	Af293	http://tigrblast.tigr.org/er-blast/index.cgi?project=afu1	Nierman *et al.*, 2005
A. nidulans	FGSC A4	http://www.broad.mit.edu/annotation/genome/aspergillus_nidulans/Blast.html	Galagan *et al.*, 2005
A. niger	CBS513.88	http://www.dsm.com/en_US/html/dfs/genomics_aniger.htm	DSM Company
A. niger	ATCC1015	http://genome.jgi-psf.org/cgi-bin/runAlignment?db=Aspni1&advanced=1	DOE -JGI
A. oryzae	RIB40	http://www.bio.nite.go.jp/dogan/MicroTop?GENOME_ID=ao	Machida *et al.*, 2005
A. terreus	NIH2624	http://www.broad.mit.edu/annotation/genome/aspergillus_terreus/Blast.html	Broad Institute of Harvard and MIT
N. crassa	OR74A	http://www.broad.mit.edu/annotation/genome/neurospora/Blast.html	Galagan *et al.*, 2003
P. anserina	S	http://podospora.igmors.u-psud.fr/blast_ol.html	*P. anserina* genome project
S. cerevisiae	S288C	http://seq.yeastgenome.org/cgi-bin/nph-blast2sgd	*Saccharomyces* Genome Database

3. Results

3.1 Heterokaryon (in)compatibility

Using strains with different spore colours and different auxotrophic mutations or dominant resistances, one can test for the formation of heterokaryotic mycelium on media on which the single partners are unable to grow. Different mutant lines were isolated from DSM Research BV's Strain CBS 513.88, DOE Joint Genome Institute's culture collection strain ATCC1015 and our lab strain N400 (ATCC 9029; CBS 120.49), making it possible to test for heterokaryon incompatibility between these strains. Strains ATCC1015 and our laboratory strain N400 (ATCC 9029; CBS 120.49) proved heterokaryon compatible with one another. That they thus belong to the same vegetative compatibility group suggests that they share a common clonal ancestor. However, DSM Research BV's Strain CBS 513.88 proved incompatible with the two other strains. Thus the two genomes sequenced by DSM and DOE Joint Genome Institute respectively are from heterokaryon incompatible strains.

3.2 In silico comparison of yeast and filamentous fungi

The *in silico* experiments performed with the different sequenced and available genomes show that the majority of the PCD genes from *S. cerevisiae* have homologs in the filamentous fungi, only *hel10* and *hel 13*, whose functions are unknown, are mostly missing (Table 3). In contrast the majority of vegetative incompatibility and cell death related genes from *N. crassa* and *P. anserina* can be found in the other filamentous fungi, but many of them, like *het-6*, *het-c*, *tol*, *vib-1*, *pin-c*, *het-c2*, *het-s*, *idi-1*, *idi-2*, *idi-3*, *idi-4* and *mod-A,* are missing in the baker's yeast.

Some genes like the un-24 gene, known from *N. crassa* to be involved in heterokaryon incompatibility and coding for the ribonucleotide reductase large subunit, are well conserved in the filamentous fungi and have a highly similar homolog in *S. cerevisiae* (Table 4). Whereas the putative transmembrane protein HAM-2 is very well conserved within the filamentous fungi, it differs considerably from the *S. cerevisiae* homolog.

These results show that most of the genes involved in programmed cell death are well conserved among both the filamentous fungi and the yeast *S. cerevisiae*, but that many genes involved in heterokaryon incompatibility are not.

3.3 In silico comparison of the Sordariales and the Eurotiales

N. crassa and *P. anserina* belong to the *Sordariales*, whereas the Aspergilli belong to the Eurotiales. Comparing the incompatibility/apoptosis gene sets between these two groups of filamentous fungi (see Table 3 and 4), the most remarkable difference is shown by *P. anserina*, bearing many more *idi* (induced during incompatibility) gene homologs than other fungi and missing the suppressor gene *stm1*. *P. anserina* also contains the highest number of homologs of HET-6 (35), TOL (48), HET-D (94) and IDI-6 (8), probably due to many genes with HET-domain sequences (Table 4). With the set criteria of acceptable E-values smaller than e^{-10} homologs of the IDI-1, IDI-2 and IDI-3 proteins are absent from the other filamentous fungi besides *P. anserina*. Only the IDI-2 and IDI-3 blast searches resulted in hits with very weak similarity (both e^{-5}) in *A. oryzae* and *A. terreus*.

Comparing the phylogenies of different proteins in the different species (see e.g. Figure 1 with the HET-6 gene homologs from *A. niger* and *N. crassa*) one can see that the homologs of these genes show old polymorphisms. The sequences found in the *A. niger* strains are often very similar. The most similar homologs in *N. crassa* however, can be quite different from the *A. niger* sequences. Old duplications of the ancestor genes with possible new functions are also visible in the phylogenies.

Within the tested set of proteins, there are no exclusive proteins for members of the *Sordariales*. Only the MOD-A protein has but one hit in the Aspergilli: the protein blast in *A. oryzae* resulted in one hit with only low similarity (e^{-11}) (Table 4).

3.4 In silico comparison of the different Aspergilli.

A. nidulans is a sexual *Aspergillus* species with the ability of forming both sexual and asexual spores. Also *A. fumigatus* is now supposed to be sexual (Varga 2003, Paoletti *et al.* 2005), but the other tested Aspergilli- *A. niger, A. oryzae*, and *A. terreus* – are all presumed exclusively asexual species. Comparing the different *Aspergillus* species for their putative heterokaryon incompatibility and cell death related proteins, there are no large differences between the species in the presence of certain proteins (Table 3). But, in the number of homologs of HET-6^{OR}, TOL and TOL-related PIN-C proteins there is a surprisingly big difference between the asexual and (supposedly) sexual lines. Whereas there are at least 10 HET-6, 9 TOL and 11 PIN-C homologs in the asexual strains, in the *A. nidulans* and *A. fumigatus* we found only 2 and 3 HET-6, 0 and 1 TOL, and 0 and 1

PIN-C homologs respectively (Table 4).

3.5 In silico comparison of the two A. niger *genomes*

The two sequenced *A. niger* strains differ in their sets of heterokaryon incompatibility and apoptosis related genes. Strain CBS513.88 lacks a HET-S homologue and strain ATCC1015 lacks the IDI-4/JLB-A and HEL10 homologs (Table 3). In nearly all of their putative HET proteins the two *A. niger* isolates show little to no variation in sequence. The two strains are heterokaryon incompatible and the differences –sometimes only simple substitutions, sometimes small stretches of amino acids - in the known indel regions of some of the putative *het*-genes may explain the observed incompatibility reaction between the two strains (Table 5). For the *idi-6* (*psp*) genes two alleles are present, both very similar with very few substitutions. However, for the MCA1/YCA1 (a meta-caspase) one pair is completely identical, whereas the second pair shows more
differences and is only 77% identical.

Table 3. HET and incompatibility related proteins in filamentous fungi and yeast. ID numbers for *N. crassa* and *S. cerevisiae* proteins refer to the numbers given in their respective sequencing projects, the ID numbers for the *P. anserina* proteins were taken from Genbank.

Species	Class	Protein	Species (Strains)								
			A. fumigatus (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC*1015*)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (S)	*S. cerevisiae* (S288C)
N. crassa											
	Heterokaryon incompatibility genes										
		HET-6 (NCU03533.2)	●	●	●	●	●	●	●	●	
		HET-C (NCU03125.2)	●	●	●	●	●	●	●	●	
		UN-24 (NCU03539.2)	●	●	●	●	●	●	●	●	●
	Suppressor genes										
		TOL (NCU04453.2)	●		●	●	●	●	●	●	
		VIB-1 (NCU03725.2)	●	●	●	●	●	●	●	●	
	Incompatibility related genes										
		HAM-2 (NCU03727.2)	●	●	●	●	●	●	●	●	●
		PIN-C (NCU03494.2)	●		●	●	●	●	●	●	
		RNR-A (NCU07887.2)	●	●	●	●	●	●	●	●	●
P. anserina											
	Heterokaryon incompatibility genes										
		HET-C2 (AAA20542)	●	●	●	●	●	●	●	●	
		HET-D2Y (AAL37301)	●	●	●	●	●	●	●	●	●
		HET-E4S (AAL37297)	●	●	●	●	●	●	●	●	●
		HET-S (AAB88771)		●		●		●			
	Incompatibility related genes										
		IDI-1 (AAC24119)								●	
		IDI-2 (AAC24120)								●	
		IDI-3 (AAC24121)								●	
		IDI-4/ JLB-A (AAT40415)			●			●	●	●	
		IDI-6/ PSP (AAC03564)	●	●	●	●	●	●	●	●	●
		IDI-7 (AAN41258)	●	●	●	●	●	●	●	●	●
	Modifier genes										
		MOD-A (AAC25496)					●		●	●	
		MOD-D (AAC24766)	●	●	●	●	●	●	●	●	●
		MOD-E (AAB97626)	●	●	●	●	●	●	●	●	●
S.cerevisiae											
	Programmed Cell Death genes										
		ATP4 (YPL078C)	●	●	●	●	●	●	●	●	●
		CDC48 (YDL126C)	●	●	●	●	●	●	●	●	●
		HEL10 (YNL208W)	●		●						●
		HEL13 (YOR309C)			●						●
		MCA1/YCA1 (YOR197W)	●	●	●	●	●	●	●	●	●
		NSR1 (YGR159C)	●	●	●	●	●		●	●	●
		PPA1 (YHR026W)	●	●	●	●	●	●	●	●	●
		RSM23 (YGL129C)							●	●	●
		SAR1 (YPL218W)	●	●	●	●	●	●	●	●	●
		STM1 (YLR150W)	●	●	●	●	●	●	●		●
		TOR1 (YJR066W)	●	●	●	●	●	●	●	●	●

Table 4. Number and best hits of homologs (ID numbers of the respective genome databases or Genbank in case of *P. anserina*) of *N. crassa* proteins involved in vegetative incompatibility and PCD. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

Species	Class	Protein	Species (Strains)								
			A. fumigatus (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* [1] (CBS513.88)	*A. niger* (ATCC*1015*)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* [2] (OR74A)	*P. anserina* [2] (S)	*S. cerevisiae* [2] (S288C)
N. crassa											
	Heterokaryon incompatibility genes										
		HET-6 (NCU03533.2)	3 Afu8g01020 (9e-27)	2 AN5504.3 (2e-22)	11 An02g12950 (1e-27)	10 37491 (4e-30)	17 AO090026000320 (5e-22)	13 ATEG_00782.1 (2e-30)	13 NCU04589.2 (0.0)	35 Pa_4_1050 (6e-74)	0
		HET-C (NCU03125.2)	2 Afu7g02570 (0.0)	2 AN9067.3 (0.0)	1 An15g06140 (0.0)	1 182955 (0.0)	2 AO090038000633 (0.0)	2 ATEG_01884.1 (0.0)	2 NCU03493.2 (0.0)	2 Pa_1_11060 (0.0)	0
		UN-24 (NCU03539.2)	1 Afu4g06690 (0.0)	1 AN4380.3 (0.0)	1 An04g01080 (0.0)	1 57102 (0.0)	1 AO090023000916 (0.0)	1 ATEG_05560.1 (0.0)	1	1 Pa_6_4460 (0.0)	2 YER070W (0.0)
	Suppressor genes										
		TOL (NCU04453.2)	1 Afu2g12280 (2e-13)	0	10 An15g04930 (3e-47)	9 40668 (4e-33)	12 AO090701000899 (2e-37)	17 ATEG_06478.1 (9e-41)	31 NCU03378.2 (0.0)	48 Pa_1_1070 (2e-52)	0
		VIB-1 (NCU03725.2)	2 Afu8g04050 (5e-62)	2 AN1414.3 (0.0)	1 An16g09130 (1e-57)	1 193642 (1e-60)	2 AO090103000018 (4e-61)	2 ATEG_00017.1 (0.0)	2 NCU04729.2 (1e-39)	2 Pa_2_9470 (0.0)	0
	Incompatibility related genes										
		HAM-2 (NCU03727.2)	1 Afu6g04250 (0.0)	1 AN6611.3 (0.0)	1 An15g01470 (0.0)	1 181774 (0.0)	1 AO090701000144 (0.0)	1 ATEG_07050.1 (0.0)	1	1 Pa_2_9440 (0.0)	1 YNL127W (2e-25)
		PIN-C1 (NCU03494.2) (DQ309556)[3]	1 Afu2g12280 (1e-19)	0	11 An09g05610 (1e-46)	13 43178 (2e-38)	12 AO090023000291 (1e-39)	17 ATEG_08482.1 (2e-39)	35 NCU03484.2 (0.0)	51 Pa_7_3460 (7e-49)	0
		RNR-A (NCU07887.2)	1 Afu5g12350 (0.0)	1 AN0067.3 (0.0)	1 An14g06870 (0.0)	1 56635 (0.0)	1 AO090120000352 (0.0)	1 ATEG_02041.1 (0.0)	1	1 Pa_7_11090 (0.0)	2 YJL026W (0.0)

[1] Maximum number of hits was limited in 25 hits per blast search. [2] The number of homologs includes the protein with which the blast search was done with. If there is at least one additional homolog, the E-value of the best homolog is given. If there is no more homolog only the blasting protein, E-value is not given for the blasting protein itself. [3] To distinguish the three alleles of PIN-C (which are under the same ID number in the *N. crassa* database) the Genbank accession number is also given.

Table 4. (Continued) Number and best hits of homologs (ID numbers of the respective genome databases or Genbank in case of *P. anserina*) of *P. anserina* proteins involved in vegetative incompatibility and PCD. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

Species	Class	Protein	Species (Strains)								
			A. fumigatus (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* [1] (CBS513.88)	*A. niger* (ATCC*1015*)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* [2] (OR74A)	*P. anserina* [2] (S)	*S. cerevisiae* [2] (S288C)
P. anserina											
	Heterokaryon incompatibility genes										
		HET-C2 (AAA20542)	1 Afu3g13820 (5e-62)	1 AN3147.3 (0.0)	1 An02g09550 (2e-66)	1 52371 (6e-73)	1 AO090012000776 (4e-59)	1 ATEG_04047.1 (0.0)	1 NCU07947.2 (0.0)	1	0
		HET-D2Y (AAL37301)	56 Afu7g07100 (0.0)	62 AN8468.3 (0.0)	25 An11g03460 (0.0)	60 129126 (0.0)	81 AO090005000041 (0.0)	69 ATEG_09923.1 (0.0)	67 NCU00794.2 (0.0)	94 Pa_4_1190 (0.0)	32 YCR072C (2e-54)
		HET-E4S (AAL37297)	39 Afu7g07100 (0.0)	48 AN6803.3 (0.0)	25 An11g03460 (0.0)	52 118837 (0.0)	65 AO090005000041 (0.0)	55 ATEG_09957.1 (0.0)	45 NCU05919.2 (0.0)	76 Pa_7_3110 (0.0)	18 YCR072C (2e-24)
		HET-S (AAB88771)	0	1 AN2022.3 (1e-20)	0	1 53015 (5e-12)	0	1 ATEG_01784.1 (2e-17)	0	0	0
	Incompatibility related genes										
		IDI-1 (AAC24119)	0	0	0	0	0	0	0	1	0
		IDI-2 (AAC24120)	0	0	0	0	0	0	0	1	0
		IDI-3 (AAC24121)	0	0	0	0	0	0	0	1	0
		IDI-4/ JLB-A (AAT40415)	0	0	1 An04g04510 (4e-12)	0	0	1 ATEG_08748.1 (4e-12)	1 NCU08055.2 (1e-11)	1	0
		IDI-6/ PSP (AAC03564)	3 Afu5g09210 (0.0)	2 AN10030.3 (0.0)	2 An07g03880 (0.0)	2 200187 (0.0)	2 AO090020000517 (0.0)	2 ATEG_06546.1 (0.0)	4 NCU00673.2 (0.0)	8 Pa_2_13590 (2e-71)	3 YEL060C (0.0)
		IDI-7 (AAN41258)	1 Afu1g07470 (2e-59)	1 AN5131.3 (0.0)	1 An07g10020 (8e-63)	1 209252 (3e-75)	1 AO090012000997 (7e-63)	1 ATEG_10180.1 (0.0)	1 NCU01545.2 (0.0)	1	1 YBL078C (1e-48)

[1] Maximum number of hits was limited in 25 hits per blast search. [2] The number of homologs includes the protein with which the blast search was done with. If there is at least one additional homolog, the E-value of the best homolog is given. If there is no more homolog only the blasting protein, E-value is not given for the blasting protein itself.

Table 4. (Continued) Number and best hits of homologs (ID numbers of the respective genome databases or Genbank in case of *P. anserina*) of *P. anserina* and *S. cerevisiae* proteins involved in vegetative incompatibility and PCD. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

Species	Class	Protein	Species (Strains)								
			A. fumigatus (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* [1] (CBS513.88)	*A. niger* (ATCC*1015*)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* [2] (OR74A)	*P. anserina* [2] (S)	*S. cerevisiae* [2] (S288C)
P. anserina											
	Modifier genes										
		MOD-A (AAC25496)	0	0	0	0	1 AO090103000290 (5e-11)	0	1 NCU07121.2 (0.0)	2 Pa_3_500 (0.0)	0
		MOD-D (AAC24766)	3 Afu1g12930 (0.0)	3 AN1016.3 (0.0)	4 An08g05820 (0.0)	4 55775 (0.0)	4 AO090012000600 (0.0)	3 ATEG_00488.1 (0.0)	3 NCU05206.2 (0.0)	4 Pa_7_7970 (7e-95)	2 YER020W (2e-99)
		MOD-E (AAB97626)	1 Afu5g04170 (0.0)	1 AN8269.3 (0.0)	1 An09g06590 (0.0)	1 202811 (0.0)	1 AO090102000620 (0.0)	1 ATEG_07996.1 (0.0)	1 NCU04142.2 (0.0)	1	2 YMR186W (0.0)
S.cerevisiae											
	Programmed Cell Death genes										
		ATP4 (YPL078C)	1 Afu8g05440 (2e-51)	1 AN1534.3 (0.0)	1 An16g07290 (2e-55)	1 214565 (4e-64)	1 AO090005000604 (7e-53)	1 ATEG_04777.1 (0.0)	1 NCU00502.2 (0.0)	1 Pa_1_17630 (1e-53)	1
		CDC48 (YDL126C)	20 Afu2g17110 (0.0)	19 AN7254.3 (0.0)	21 An04g09170 (0.0)	19 205183 (0.0)	20 AO090102000107 (0.0)	19 ATEG_10033.1 (0.0)	21 NCU00018.2 (0.0)	23 Pa_7_5590 (0.0)	21 YLR397C (0.0)
		HEL10 (YNL208W)	1 Afu1g06580 (9e-12)	0	3 An18g05640 (3e-15)	0	0	0	0	0	1
		HEL13 YOR309C)	0	0	1 An16g08250 (2e-11)	0	0	0	0	0	1
		MCA1/YCA1 (YOR197W)	2 Afu3g14140 (7e-89)	2 AN5712.3 (0.0)	2 An09g04470 (0.0)	2 122116 (0.0)	2 AO090012000660 (9e-95)	2 ATEG_04015.1 (0.0)	2 NCU09882.2 (0.0)	2 Pa_5_5940 (5e-92)	1

[1] Maximum number of hits was limited in 25 hits per blast search. [2] The number of homologs includes the protein with which the blast search was done with. If there is at least one additional homolog, the E-value of the best homolog is given. If there is no more homolog only the blasting protein, E-value is not given for the blasting protein itself.

Table 4. (Continued) Number and best hits of homologs (ID numbers of the respective genome databases or Genbank in case of *P. anserina*) of *S. cerevisiae* proteins involved in vegetative incompatibility and PCD. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

Species	Class	Protein	Species (Strains)								
			A. fumigatus (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* [1] (CBS513.88)	*A. niger* (ATCC*1015*)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* [2] (OR74A)	*P. anserina* [2] (S)	*S. cerevisiae* [2] (S288C)
S.cerevisiae											
	Programmed Cell Death genes										
		NSR1 (YGR159C)	13 Afu3g07710 (3e-51)	12 AN4865.3 (0.0)	25 An02g12640 (3e-97)	11 173997 (2e-33)	11 AO090020000411 (9e-52)	0	10 NCU03092.2 (3e-41)	9 Pa_1_11830 (1e-43)	8 YNL016W (1e-17)
		PPA1 YHR026W)	3 Afu2g15560 (2e-47)	3 AN7603.3 (0.0)	4 An15g05730 (1e-72)	4 210268 (7e-57)	3 AO090012000274 (8e-47)	3 ATEG_02422.1 (0.0)	3 NCU09747.2 (0.0)	4 Pa_6_1870 (1e-47)	2 YEL027W (5e-77)
		RSM23 (YGL129C)	0	0	0	0	0	0	1 NCU08120.2 (3e-18)	1 Pa_1_6720 (2e-20)	1
		SAR1 (YPL218W)	7 Afu1g04940 (8e-73)	5 AN0411.3 (0.0)	7 An01g04040 (2e-75)	7 206151 (2e-88)	9 AO090003000842 (4e-76)	7 ATEG_02597.1 (0.0)	7 NCU00381.2 (0.0)	7 Pa_7_2150 (7e-77)	7 YDL192W (1e-24)
		STM1 (YLR150W)	1 Afu3g10920 (1e-14)	1 AN10614.3 (2e-20)	1 An02g06710 (2e-26)	1 47011 (3e-15)	1 AO090003000636 (1e-17)	1 ATEG_04616.1 (2e-14)	1 NCU00225.2 (2e-13)	0	1
		TOR1 (YJR066W)	6 Afu2g10270 (0.0)	4 AN5982.3 (0.0)	3 An16g04720 (0.0)	3 53581 (0.0)	6 AO090011000608 (0.0)	4 ATEG_04462.1 (0.0)	3 NCU05608.2 (0.0)	4 Pa_4_9630 (0.0)	5 YKL203C (0.0)

[1] Maximum number of hits was limited in 25 hits per blast search. [2] The number of homologs includes the protein with which the blast search was done with. If there is at least one additional homolog, the E-value of the best homolog is given. If there is no more homolog only the blasting protein, E-value is not given for the blasting protein itself.

Table 5. A comparison between the different heterokaryon incompatibility and programmed cell death related genes in the two sequenced *A. niger* genomes (CBS513.88 and ATCC1015). If two proteins differ in size, the longer one is the basis for counting percentage of identities, similarities and gaps. Gaps are counted only in the homologous region.

Species	Function	Protein	Type and size of difference		Identities	Gaps
N. crassa						
	Heterokaryon incompatibility genes					
		HET-6	only HET domain motifs are slightly conserved		–	–
		HET-C	indel: 622 substitution: 196		791/793 (99%)	1/793 (0%)
		UN-24	no difference		(100%	(0%)
	Suppressor genes					
		TOL	very diverse proteins, with conserved HET domain motifs		–	–
		VIB-1	substitution: 195		585/586 (99%)	(0%)
	Incompatibility related genes					
		HAM-2	substitution: 756		1066/1067 (99%)	(0%)
		PIN-C	very diverse proteins, with conserved HET domain motifs		–	–
		RNR-A	no difference		(100%)	(0%)
P. anserina						
	Heterokaryon incompatibility genes					
		HET-C2	no difference		(100%)	(0%)
		HET-D / HET-E	diverse proteins with WD40 repeats, but no remarkable similarity		–	–
	Incompatibility related genes					
		IDI-6 / PSP (2 alleles)	1st pair: indel: 534 substitutions: 398, 525		531/535 (99%)	1/534 (0%)
			2nd pair: substitutions: 398, 436-443, 450-459, 471-514		413/416 (99%)	(0%)
		IDI-7	no difference		(100%)	(0%)
	Modifier genes					
		MOD-D (2 alleles)	Members of the two allele pairs are 100% identical, between the pairs there are some differences: 1-167 variable part, 1-60 and 168-360 more conserved region.			
		MOD-E	substitution: 244		672/702 (96%)	(0%)
S. cerevisiae						
	Programmed Cell Death genes					
		ATP4	no difference		100%	0%
		CDC48	no difference		100%	0%
		HEL13	present only in CBS 513.88		–	–
		MCA1 / YCA1		1st pair	438/438 (100%)	(0%)
		(2 alleles)	2nd allele: :indel region: 1-56, 75-104, 237, 443-447 subtitutions: 57-75, 105-106. 236, 440-442	2nd pair	341/441 (77%)	1/333 (0%)
		NSR1	diverse proteins, with short conserved motifs		–	–
		PPA1	no difference		100%	0%
		SAR1	no difference		100%	0%
		STM1	indel region: 8-13		297/303 (98%)	6/303 (1%)
		TOR1	substitution: 16		2389/2390 (99%)	(0%)

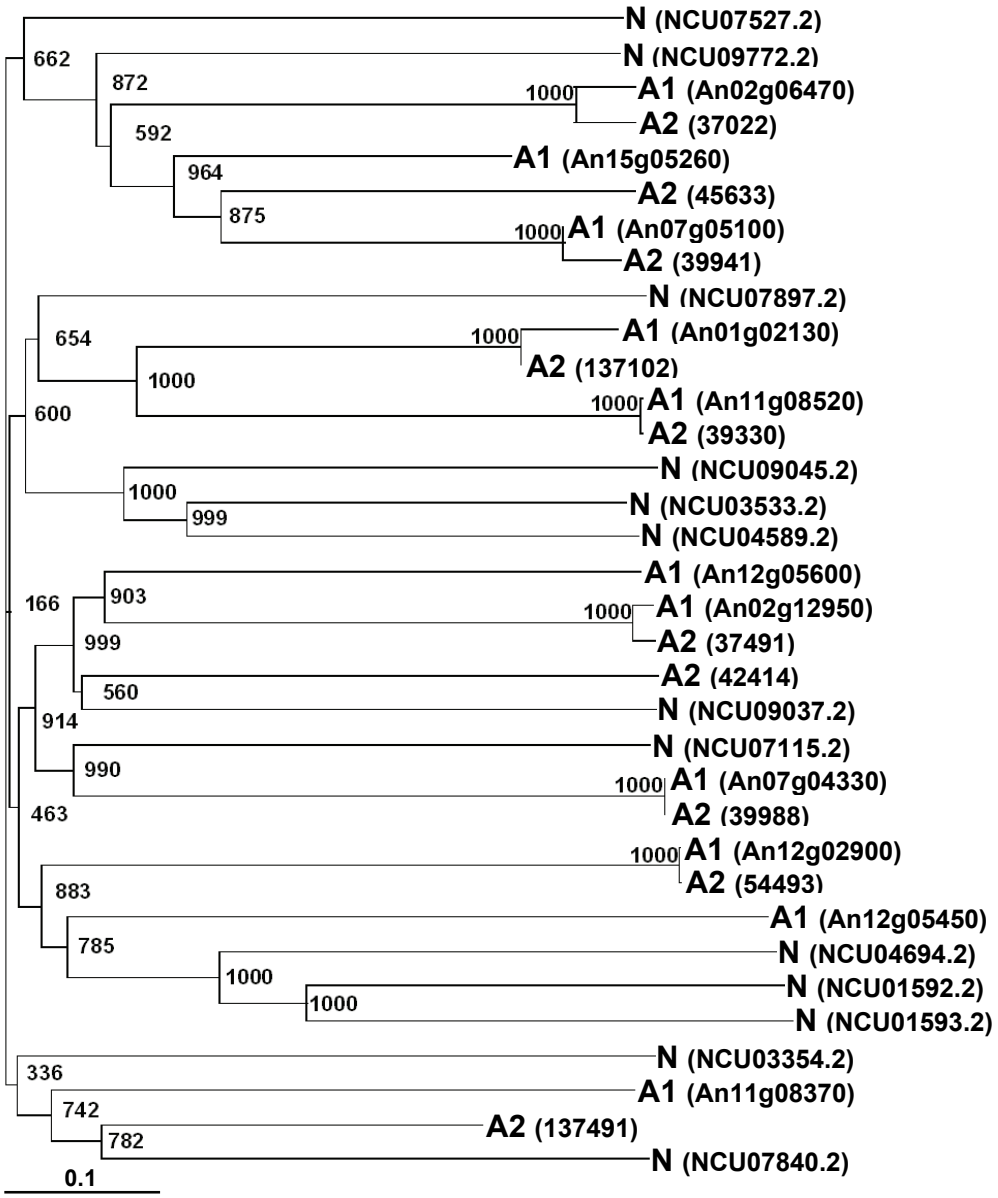


Figure 1. Neighbour-joining tree of HET-6 proteins from *N. crassa* and two *A. nigers*. Strains on the tree: A1: *A. niger* CBS513.88; A2: *A. niger* ATCC1015; N: *N. crassa* OR74A. Accession numbers of the different homologs are given from the respective databases. Sequence alignment and bootstrapping was performed with ClustalX.

4. Discussion

Very little is known about the genetic basis of heterokaryon incompatibility in the (presumably) asexual *A. niger*. For this mainly *in silico* study we used two different genome sequences of two heterokaryon incompatible *A. niger* strains that we show to be heterokaryon incompatible. We compared these genomes and the genome sequences of four closely related sexual and asexual species. We searched these databases with genes known in *P. anserina*, *N. crassa* and *S. cerevisiae* to be involved in heterokaryon incompatibility reactions or apoptosis. Our aim was to find out whether the same genes may be involved in the incompatibility reactions between different *A. niger* isolates as the ones found to interact in *N. crassa* and *P. anserina*, fungi that have different sets of active *het*-genes. Few differences were found between the two sequenced *A. niger* genomes, but many of the known heterokaryon incompatibility genes were indeed present in the *A. niger* genomes. Some of the examined *het*-genes were even found to have many homologs.

The majority of programmed cell death related genes from *S. cerevisiae* had also homologs in the filamentous fungi. But, as could be expected, we found that very few of the genes that in filamentous fungi are involved in (blocking) hyphal fusions had homologs in baker's yeast. Yeast and filamentous fungi diverged approximately 1.1 billion years ago (Cai *et al.* 2006) and yeast have a different, mainly unicellular, life style. Only the HET-D and HET-E proteins which have GTP binding capacity and the so-called WD (tryptophan-aspartate) repeats (Espagne *et al.* 2002) yielded many homologous sequences in *S. cerevisiae*. Of course these genes may have pleiotropic functions and may have a different function in *S. cerevisiae* than controlling anastomoses formation. The *idi-7* gene of *P. anserina* is classified as being an ortholog of the *S. cerevisiae aut7p* gene coding for a protein whose binding to the membrane represents an early step in vesicle formation (Lang *et al.* 1998). Not surprisingly this protein is quite well preserved between both filamentous fungi and yeast as are the genes coding for ribonucleotide reductases and serine proteases.

For filamentous fungi, the remarkable difference is in the number of protein homologs, not in the presence/absence of these proteins. The numbers of genes with a *het*-domain varies largely in different filamentous fungi: between 6 and 150 *het*- domain containing genes have been found in sequenced fungal genomes (Sven Saupe, pers. comm.). Within our investigated group of filamentous fungi *P. anserina* bears the most different gene set and the largest set of genes with *het*-domains: HET-6 has 35, HET-D 94, PIN-C 51 and TOL 48 homologs below the threshold of e^{-10}. Some of these homologous sequences overlap between the different genes due to their conserved *het*-domains. A possible explanation for this can be in its life cycle. *P. anserina* is a saprophytic fungus, which feeds on partially digested materials in the dung of herbivorous animals. As a coprophilic fungus it grows in synchrony and under rather high density with

competitors for the same ephemeral and limited substrate. Therefore, the risk of exploitation or genetic infection by others may be relatively high in comparison to most other fungi. An efficient way to limit exchange of genetic materials is heterokaryon incompatibility, which process is governed by the so called *het*-genes. In *P. anserina* the majority of the incompatibility reactions is due to non-allelic interactions (exception is het *het-s*/*het-S* reaction). *Het-d* and *het-e* trigger incompatibility with the *het-c2*. Both *het-d* and *het-e* encode HET domain proteins, and due to the presence of this domain, these proteins seem to be responsible for nonself recognition in filamentous fungi, including *P. anserina* (Kaneko *et al.* 2006). Whether the huge number of HET domain protein homologs thus reflects the relative importance of fusion-rejection systems in the life history of *P. anserina* is unclear, though this has been suggested for coprophiles (Buss 1982). Of course, het-domain containing genes also may have other functions than just heterokaryon incompatibility reactions and there certainly seems to be a large family of het-domain genes.

Aspergillus species are similar in many components of heterokaryon incompatibility in agreement with previous findings (Table 3; Fedorova *et al.* 2005). Interestingly, *A. nidulans* and *A. fumigatus* are the two species, which have the least HET domain protein homologs. Maybe this finding can be explained by the capability of sex: *A. nidulans* is able to reproduce sexually, and there are indirect proofs for the presence of a sexual life cycle in *A. fumigatus* (Varga 2003, Paoletti *et al.* 2005). The other Aspergilli (*A. niger*, *A. oryzae*, *A. terreus*) are known as asexual species. The question in this case is why do asexual species bear much more HET domain genes? As it was mentioned above these genes are the main components of non-allelic incompatibility, therefore there could be a disadvantage for fungi with a sexual cycle to have such genes.
Another reason could be the proposed function for heterokaryon incompatibility in limiting the spread of detrimental cytoplasmic or nuclear elements (Caten 1972). In *A. niger* and related black *Aspergillus* species dsRNA mycoviruses occur in approximately 10% of the natural isolates. These mycoviruses are effectively transferred to all asexual spores (van Diepeningen *et al.* 2006). Tests showed that heterokaryon incompatibility indeed efficiently blocks the transfer of mycoviruses in these black Aspergilli (Van Diepeningen *et al.* 1997). In sexual *A. nidulans* no dsRNA viruses were found in nature and also here artificially introduced viruses efficiently find their way to the asexual spores. However, when sexual spores are produced the mycoviruses are excluded from the offspring (Coenen *et al.* 1997). Therefore *A. nidulans* has an extra option to get rid of parasitic elements through its sexual cycle and thus heterokaryon incompatibility may be less important between *A. nidulans* strains.
Also *A. fumigatus* has presumably a sexual reproduction mode though never observed in lab or field, and may eliminate parasites during this sexual cycle. The species does contain less *het*-genes when compared to asexual relatives. Although Varga *et al.* (1998) could not detect dsRNA elements in any of the examined 61 *A. fumigatus* isolates, mycovirus-infected *A. fumigatus* strains have been found recently (Warn *et al.* 2006). However, it does have an infection rate of approximately 3.5 % in nature (Varga *et al.* 1998). So far, several sexual *Aspergillus* species were

reported to have a hampered sexual virus transmission (Coenen *et al.* 1997; Varga *et al.*, 1998), however, *Neosartorya hiratsukae* from *Aspergillus* section *Fumigati* transmits the viruses efficiently through the ascospores (Varga *et al.* 1998). Of asexual *A. oryzae* and *A. terreus* nothing is known about mycovirus infections and the need to block intermycelial transfer.

We compared the putative HET proteins of two sequenced *A. niger* strains. As can be expected, they largely possess the same gene set, except for one gene, *idi-4*, which was not found in the ATCC1015 strain. Similarly to the other asexual Aspergilli, we found a high number of HET domain proteins (PIN-C, HET-6, HET-D and TOL homologs). Nearly all *het*-genes were highly similar between both *A. niger* strains, differences were limited to a few substitutions but are potentially crucial for incompatibility reactions. The regions of these genes involved in self/non-self recognition may be under positive selection and single-amino-acid differences can be sufficient to trigger incompatibility (Saupe 2000). Thus, the observed small differences may be an explanation for the observed heterokaryon incompatibility between the two *A. niger* strains. However, as both *P. anserina* and *N. crassa* seem to have selected different sets of heterokaryon incompatibility genes to block intermycelial transfer, Aspergilli may use a completely different set of genes as well.

In this article we investigated nine fungal databases to compare the genes involved in heterokaryon incompatibility and programmed cell death. We found that the majority of the genes are present in filamentous fungi, but many of the incompatibility related genes are missing from the yeast *S. cerevisiae*. We found that the number of homologs can vary from species to species, and there is a remarkable difference between *Sordaria* and *Aspergillus* species in the number of HET domain genes. *A. fumigatus* and *A. nidulans* seem to be the most different ones among the examined *Aspergillus* species as these species have much less HET domain genes than the asexual species. The annotation in the two *A. niger* genomes was not always perfect, but close examination of the two *A. niger* strains showed very high similarities at protein level. Further practical research is needed to find a satisfying explanation for the high level of incompatibility in the natural populations of black Aspergilli and to pinpoint functional *het*-genes in the species. As a result of our data mining, the sequences of the known incompatibility genes are available for functional analysis, to uncover the secrets of incompatibility between the natural isolates of black Aspergilli.

Acknowledgments

We want to thank all institutes, companies and individuals involved in the genome projects of the species used in this study: *A. fumigatus* preliminary sequence data were obtained from The Institute for Genomic Research website at http://www.tigr.org. *A. niger* CBS513.88 sequence data were obtained from the DSM Research B.V., the *A. niger* ATCC1015 sequence data from the DOE Joint Genome Institute. The *A. nidulans* genome data are from the *Aspergillus* Sequencing Project, the *Aspergillus terreus* genome data from the *Aspergillus terreus* Sequencing Project, and the release 7 data of the *N. crassa* genome were all from the Broad Institute of Harvard and MIT (http://www.broad.mit.edu). *A. oryzae* data were obtained from the NITE Biological Resource Center. Data from the complete sequence of *P. anserina* were obtained from the *P. anserina* genome project and the *S. cerevisiae* genome data from the *Saccharomyces* Genome Database.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, 1990. Basic Local Alignment Search Tool. *Journal of Molecular Biology* 215, 403–410.

Anwar MM, Croft JH, Dales RB, 1993. Analysis of heterokaryon incompatibility between heterokaryon-compatibility (h-c) groups R and GL provides evidence that at least eight *het* loci control somatic incompatibility in *Aspergillus nidulans*. *Journal of General Microbiology* 139, 1599-1603.

Barreau C, Iskandar M, Loubradou G, Levallois V, Bégueret J, 1998. The mod-A suppressor of nonallelic heterokaryon incompatibility in *Podospora anserina* encodes a proline-rich polypeptide involved in female organ formation. *Genetics* 149, 915-926.

Bourges N, Groppi A, Barreau C, Clavé C, Bégueret J, 1998. Regulation of gene expression during the vegetative incompatibility reaction in Podospora anserina: characterization of three induced genes. *Genetics* 150, 633-641.

Buss LW, 1982. Somatic-cell parasitism and the evolution of somatic tissue compatibility. *Proceedings of the National Academy of Sciences of the USA* 79, 5337-5341.

Cai JJ, Woo PCY, Lau SKP, Smith DK,Yuen KY, 2006. Accelerated evolutionary rate may be responsible for the emergence of lineage-specific genes in Ascomycota. *Journal of Molecular Evolution* 63, 1-11.

Caten CE, 1972. Vegetative incompatibility and cytoplasmic infection in fungi. *Journal of General Microbiology* 72, 221-229.

Coenen A, Kevei F, Hoekstra RF, 1997. Factors affecting the spread of double stranded RNA mycoviruses in *Aspergillus nidulans*. *Genetical Research* 69, 1–10.

Cortesi P, Milgroom MG, 1998. Genetics of vegetative incompatibility in *Cryphonectria parasitica*. *Applied and Environmental Microbiology* 64, 2988–2994.

Coustou V, Deleu C, Saupe S, Bégueret J, 1997. The protein product of the *het-s* heterokaryon incompatibility gene of the fungus *Podospora anserina* behaves as a prion analog. *Proceedings of the National Academy of Sciences of the USA* 94, 9773-9778.

Debets AJM, Swart K, Bos CJ, 1990. Genetic analysis of the *Aspergillus niger*: isolation of chlorate resistant mutants, their use in mitotic mapping and evidence for an eighth linkage group. *Molecular and General Genetics* 221, 453-458.

Debets AJM, Swart K, Hoekstra RF, Bos CJ, 1993. Genetic maps of eight linkage groups of *Aspergillus niger* based on mitotic mapping. *Current Genetics* 23, 47-53.

Debets AJM, Griffiths AJF, 1998. Polymorphism of *het*-genes prevents resource plundering in *Neurospora crassa*. *Mycological Research* 102, 1343-1349.

Dementhon K, Paoletti M, Pinan-Lucarré B, Loubradou-Bourges N, Sabourin M, Saupe SJ, Clavé C, 2003. Rapamycin mimics the incompatibility reaction in the fungus *Podospora anserina*. *Eukaryotic Cell* 2, 238-246.

Dementhon K, Saupe SJ, Clavé C, 2004. Characterization of IDI-4, a bZIP transcription factor inducing autophagy and cell death in the fungus *Podospora anserina*. *Molecular Microbiology* 53, 1625-1640.

Dementhon K, Saupe SJ, 2005. DNA-Binding specificity of the IDI-4 basic leucine zipper factor of *Podospora anserina* defined by systematic evolution of ligands by exponential enrichment (SELEX). *Eukaryotic Cell* 4, 476-483.

Espagne E, Balhadère P, Penin M-L, Barreau C, Turcq B, 2002. HET-E and HET-D belong to a new subfamily of WD40 proteins involved in vegetative incompatibility specificity in the fungus *Podospora anserina*. *Genetics* 161, 71-81.

Fedorova ND, Badger JH, Robson GD, Wortman JR, Nierman WC, 2005. Comparative analysis of programmed cell death pathways in filamentous fungi. *BMC Genomics* 6, 177-190.

Fitzgibbon GJ, Morozov IY, Jones MG, Caddick MX, 2005. Genetic analysis of the TOR pathway in *Aspergillus nidulans*. *Eukaryotic Cell* 4, 1595-1598.

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CPC, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B, 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 438, 859-868.

Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, Batzoglou S, Lee SI, Bastürkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scazzocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Draht O, Busch S, D'Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Peñalva MA, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Nierman WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, Birren BW, 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438, 1105-1115.

Glass NL, Kuldau GA, 1992. Mating type and vegetative incompatibility in filamentous ascomycetes. *Annual Review of Phytopathology* 30, 201–224.

Glass NL, Jacobson DJ, Shiu PKT, 2000. The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycete fungi. *Annual Review of Genetics* 34, 165-186.

Glass NL, Kaneko I, 2003. Fatal attraction: nonself recognition and heterokaryon incompatibility in filamentous fungi. *Eukaryotic Cell* 2, 1-8.

Kaneko I, Dementhon K, Xiang Q, Glass NL, 2006. Nonallelic interactions between *het-c* and a polymorphic locus, *pin-c*, are essential for nonself recognition and programmed cell death in *Neurospora crassa*. *Genetics* 172, 1545-1555.

Kotierk M, Smith ML, 2001. Unpublished sequence, direct submission Submitted (22-FEB-2001) Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ont K1S 5B6, Canada

Lang T, Schaeffeler E, Bernreuther D, Bredschneider M, Wolf DH, Thumm M, 1998. Aut2p and Aut7p, two novel microtubule-associated proteins are essential for delivery of autophagic vesicles to the vacuole. *EMBO Journal* 17, 3597-3607.

Leslie JF, 1993. Fungal vegetative incompatibility. *Annual Review of Phytopathology* 31, 127–150.

Ligr M, Velten I, Froehlich E, Madeo F, Ledig M, Froehlich K-U, Wolf DH, Hilt W, 2001. The proteosomal substrate Stm1 participates in apoptosis-like cell death in yeast. *Molecular Biology of the Cell* 12, 2422-2432.

Loubradou G, Bégueret J, Turcq B, 1997. A mutation in an HSP90 gene affects the sexual cycle and suppresses vegetative incompatibility in the fungus *Podospora anserina*. *Genetics* 147, 581-588.

Loubradou G, Bégueret J, Turcq B, 1999. MOD-D, a G-alpha subunit of the fungus *Podospora anserina*, is involved in both regulation of development and vegetative incompatibility. *Genetics* 152, 519-528.

Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto KI, Arima T, Akita O, Kashiwagi Y, Abe K, Gomi K, Horiuchi H, Kitamoto K, Kobayashi T, Takeuchi M, Denning DW, Galagan JE, Nierman WC, Yu J, Archer DB, Bennett JW, Bhatnagar D, Cleveland TE, Fedorova ND, Gotoh O, Horikawa H, Hosoyama A, Ichinomiya M, Igarashi R, Iwashita K, Rao JP, 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 JR, 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.

Madeo F, Froehlich E, Froehlich KU, 1997. A yeast mutant showing diagnostic markers of early and late apoptosis. *Journal of Cell Biology* 139, 729-734.

Madeo F, Herker E, Maldener C, Wissing S, Laechelt S, Herlan M, Fehr M, Lauber K, Sigrist SJ, Wesselborg S, Froehlich KU, 2002. A caspase-related protease regulates apoptosis in yeast. *Molecular Cell* 9, 911-917.

Matsuyama S, Xu Q, Velours J, Reed JC, 1998. The mitochondrial F0F1-ATPase proton pump is required for function of the proapoptotic protein Bax in yeast and mammalian cells. *Molecular Cell* 1, 327–336.

Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulsen R, Davies R, Dyer PS, Farman M, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, Garcia JL, Garcia MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, Gwilliam R, Haas B, Haas H, Harris D, Horiuchi H, Huang J, Humphray S, Jimenez J, Keller N, Khouri H, Kitamoto K, Kobayashi T, Konzack S, Kulkarni R, Kumagai T, Lafon A, Latge JP, Li W, Lord A, Lu C, Majoros WH, May GS, Miller BL, Mohamoud Y, Molina M, Monod M, Mouyna I, Mulligan S, Murphy L, O'Neil S, Paulsen I, Penalva MA, Pertea M, Price C, Pritchard BL, Quail M, Rabbinowitsch E, Rawlins N, Rajandream MA, Reichard U, Renauld H, Robson GD, Rodriguez de Cordoba S, Rodriguez-Pena JM, Ronning CM, Rutter S, Salzberg SL, Sanchez M, Sanchez-Ferrero JC, Saunders D, Seeger K, Squares R, Squares S, Takeuchi M, Tekaia F, Turner G, Vazquez de Aldana CR, Weidman J, White O, Woodward J, Yu JH, Fraser C, Galagan JE, Asai K, Machida M, Hall N, Barrell B, Denning DW, 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* 438, 1151-1156.

Page RDM, 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12, 357-358.

Paoletti M, Castroviejo M, Bégueret J, Clavé C, 2001. Identification and characterization of a gene encoding a subtilisin-like serine protease induced during the vegetative incompatibility reaction in *Podospora anserina*. *Current Genetics* 39, 244-252.

Paoletti M, Rydholm C, Schwier EU, Anderson MJ, Szakacs G, Lutzoni F, Debeaupuis JP, Latgé JP, Denning DW, Dyer PS, 2005. Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Current Biology* 15, 1242-1248.

Pinan-Lucarre B, Paoletti M, Dementhon K, Coulary-Salin B, Clavé C, 2003. Autophagy is induced during cell death by incompatibility and is essential for differentiation in the filamentous fungus *Podospora anserina*. *Molecular Microbiology* 47, 321-333.

Reichard U, Cole GT, Hill TW, Ruchel R, Monod M, 2000. Molecular characterization and influence on fungal development of ALP2, a novel serine proteinase from *Aspergillus fumigatus*. *International Journal of Medical Microbiology* 290, 549-558.

Rohde J, Heitman J, Cardenas ME, 2001. The TOR kinases link nutrient sensing to cell growth. *Journal of Biological Chemistry* 276, 9583-9586.

Sarkar S, Iyer G, Wu J, Glass NL, 2002. Nonself recognition is mediated by HET-C heterocomplex formation during vegetative incompatibility. *EMBO Journal* 21, 4841-4850.

Saupe SJ, Descamps C, Turcq B, Bégueret J, 1994. Inactivation of the *Podospora anserina* vegetative incompatibility locus *het-c*, whose product resembles a glycolipid transfer protein, drastically impairs ascospore production. *Proceedings of the National Academy of Sciences of the USA* 91, 5927–5931.

Saupe SJ, Clavé C, Sabourin M, Bégueret J, 2000. Characterization of *hch*, the *Podospora anserina* homolog of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa*. *Current Genetics* 38, 39-47.

Saupe SJ, 2000. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews* 64, 489-502.

Shiu PK, Glass NL, 1999. Molecular characterization of *to*l, a mediator of mating-type-associated vegetative incompatibility in *Neurospora crassa*. *Genetics* 151, 545-555.

Smith ML, Micali OC, Hubbard SP, Mir-Rashed N, Jacobson DJ, Glass NL, 2000. Vegetative incompatibility in the *het-6* region of *Neurospora crassa* is mediated by two linked genes. *Genetics* 155, 1095-1104.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG, 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25, 4876–4882.

van Diepeningen AD, Debets AJM, Hoekstra RF, 1997. Heterokaryon incompatibility blocks virus transfer among natural isolates of black Aspergilli. *Current Genetics* 32, 209-17.

van Diepeningen AD, Debets AJM, Hoekstra RF, 2006. Dynamics of dsRNA mycoviruses in the black *Aspergillus* population. *Fungal Genetics and Biology* 43, 446-452.

Varga J, 2003. Mating type gene homologues in *Aspergillus fumigatus*. *Microbiology* 149, 816-819.

Varga J, Rinyu E, Kevei F, Toth B, Kozakiewicz Z, 1998. Double stranded RNA mycoviruses in species of *Aspergillus* sections *Circumdati* and *Fumigati*. *Canadian Journal of Microbiology* 44, 569–574.

Warn PA, Sharp A, Anderson MJ, Denning DW 2006. Single-spore cultures of the sequences strain of *Aspergillus fumigatus* (Af293) have significantly different mortality rates in a murine model of IPA. 2nd Advances Against Aspergillosis Meeting Syllabus P031.

Xiang Q, Glass NL, 2002. Identification of *vib-1*, a locus involved in vegetative incompatibility mediated by *het-c* in *Neurospora crassa*. *Genetics* 162, 89-101.

Xiang Q, Rasmussen C, Glass NL, 2002. The *ham-2* locus, encoding a putative transmembrane protein, is required for hyphal fusion in *Neurospora crassa*. *Genetics*
160, 169-180.

CHAPTER 3

The *het-c* gene in *Aspergillus niger*

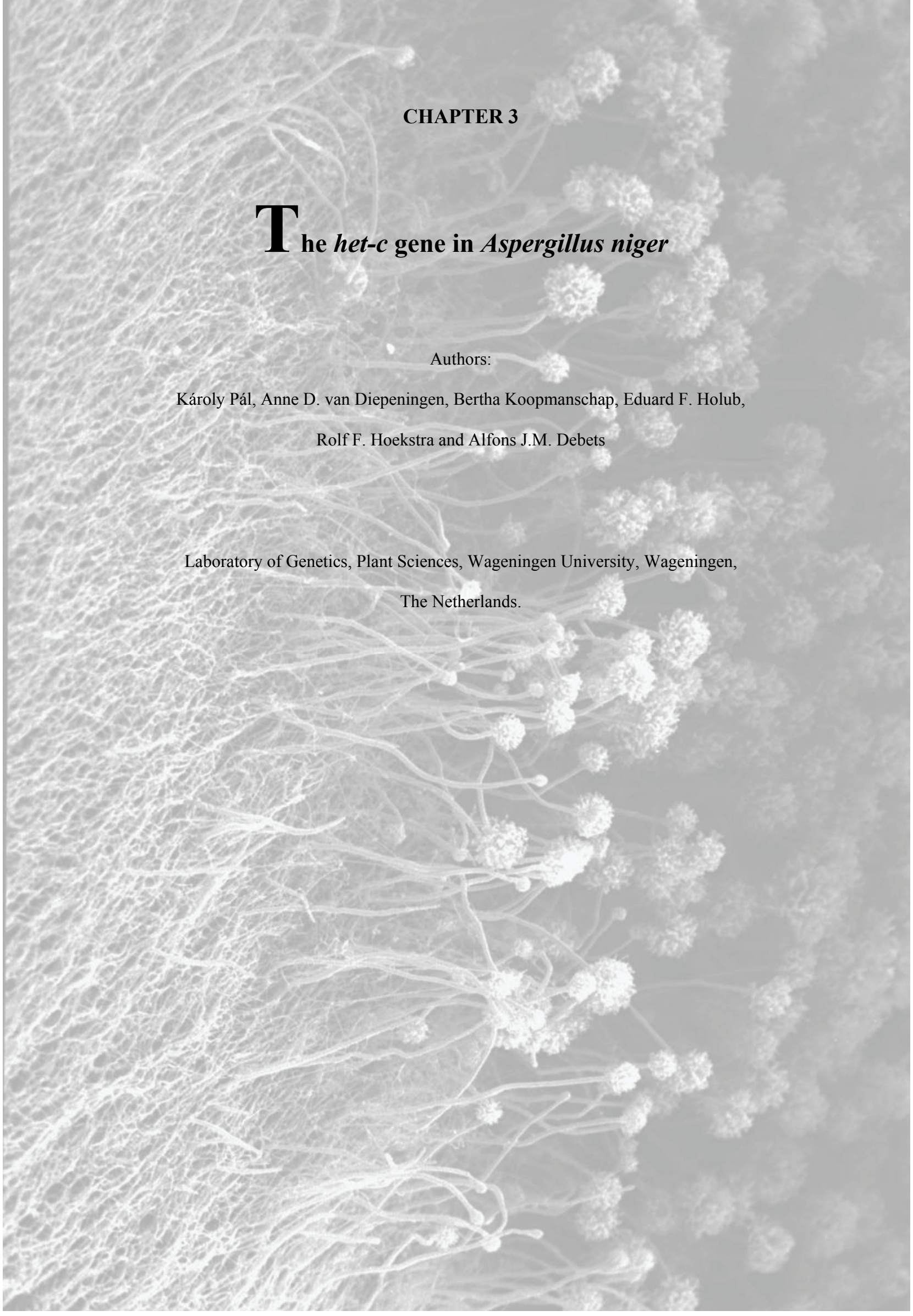
Authors:

Károly Pál, Anne D. van Diepeningen, Bertha Koopmanschap, Eduard F. Holub,

Rolf F. Hoekstra and Alfons J.M. Debets

Laboratory of Genetics, Plant Sciences, Wageningen University, Wageningen,

The Netherlands.



Abstract

Heterokaryon incompatibility among *Aspergillus niger* strains is a widespread phenomenon that is observed as the inability to form stable heterokaryons and to exchange viruses and mitochondrial markers. The genetic basis of heterokaryon incompatibility reactions is well established in some sexual filamentous fungi but is unknown in the (presumably) asexual *A. niger*. The *het-c* gene known from *Neurospora crassa* with at least three different alleles, and two different suppressor genes influencing the *het-c* heterokaryon incompatibility reaction, all have homologs in the *A. niger* genome. We used the sequenced *A. niger het-c* gene to develop degenerate primers to screen the natural population of black Aspergilli for polymorphisms. No polymorphism was found in the known variable *N. crassa het-c* indel site in the 99 strains tested, but some variation was found in a different part of the *het-c* gene. Sequencing of this variable region showed that this size difference is caused by different numbers of glutamine coding triplets possibly due to polymerase slippage, presumably without a heterokaryon incompatibility function. Introduction in *A. niger* by transformation of one of the three *N. crassa* alleles *(het-c^{PA})* showed an abortive phenotype. This suggests that the indigenous *het-c* can trigger heterokaryon incompatibility reactions in *A. niger* but probably does not function as a *het*-gene.

Keywords: heterokaryon incompatibility, *Neurospora crassa het-c*, *Aspergillus niger het-c*, protoplast transformation, cotransformation, pCB1004 plasmid, pAB4-ARp1 plasmid, glutamine repeat

1. Introduction

For filamentous fungi the formation of heterokaryotic cytoplasm and the exchange of genetic material between individuals of the same species may largely depend on the ability of two mycelia to establish hyphal anastomoses. A limitation to such heterokaryon or vegetative compatibility may restrict the exchange of detrimental cytoplasmic or nuclear elements and thus have an adaptive significance (Caten 1972). The genes that control this heterokaryon (or vegetative) incompatibility are called *het* (heterokaryon incompatibility; Glass and Kuldau 1992) or *vic* (for vegetative incompatibility; Leslie 1993) genes.

Two genetic systems, one allelic and involving a single *het* locus, the other non-allelic and involving alleles at two different *het* loci, regulate the heterokaryon incompatibility in filamentous fungi (Saupe 2000; Glass and Kaneko 2003). In *Neurospora crassa* the *het-c* locus is one of 11 identified *het* loci (Perkins 1988). Three different *het-c* alleles (*het-c*^{PA} (Panama), *het-c*^{OR} (Oak Ridge), and *het-c*^{GR} (Groveland)) have been identified in which the specificity is dependent on an indel encoding 30-48 amino acids (Mylyk 1976; Howlett *et al.* 1993; Saupe and Glass 1997; Wu and Glass 2001). The *het-c* gene has characteristics of a glycine rich cell-wall protein (Saupe *et al.* 1996). Also, several different suppressor genes that interact with the *het-c* locus have been identified: the *vib-1* gene (vegetative incompatibility blocked; Xiang and Glass 2002) and the *pin-c* gene (partner for incompatibility with *het-c*; Kaneko *et al.* 2006). Homologs of *het-c* and both modifying genes have been identified in several other filamentous fungi, e.g. in *P. anserina* where the *het-c* homologous gene is called *hch* (Saupe *et al.* 2000; Pál *et al.* in preparation, Chapter 2). However, in *P. anserina* no polymorphism was detected and thus the *hch* gene does not seem to function as a *het*-gene. Heterologous expression of *N. crassa het-c* alleles in *P. anserina* leads to a growth defective phenotype reminiscent of that of co-expression of incompatible *het-c* alleles in *Neurospora crassa* (Saupe *et al.* 2000).

To assess heterokaryon incompatibility in different filamentous fungi four methods have been used: heterokaryon tests, partial diploid analysis, transformation and confrontation (barrage) tests (Xiang and Glass 2004). For the asexual black mould *Aspergillus niger*, the only way to achieve (mitotic) recombination is via the parasexual cycle. Heterokaryon tests have been used to check for heterokaryon incompatibility in natural isolates that appeared to have a high grade of incompatibility in the population (van Diepeningen *et al.* 1997). However, little is known about the genes involved in heterokaryon incompatibility mechanisms in *A. niger*.

In this study we first did *in silico* experiments with the genome sequences of two *A. niger* strains, showing that *A. niger* does contain *het-c*, as well as homologs of both suppressor genes *vib-1* and *pin-c* and we compared the *het-c* sequences to the *N. crassa* known *het-c* alleles. Secondly, we checked nearly 100 isolates of our black *Aspergillus* culture collection, belonging to the species of *A. niger*, *A. tubingensis*, *A. carbonarius,* and *A. japonicus* for polymorphism at the *het-c* locus. No polymorphism was detected in the indel site homologous to the *N. crassa* alleles, though some variation was found in a glutamine rich region. Since we did not find any presumed

functional polymorphism in the natural isolates, we could not test the functionality of the *het-c* gene in the black Aspergilli directly. But, as transformation is an alternative way to create partial diploids for *het* loci that have been cloned (Glass *et al.* 1988, Glass *et al.* 1990, Saupe *et al.* 1996 and Smith *et al.* 2000), we could introduce the three *N. crassa* alleles *het-c*^{PA}, *het-c*^{OR}, and *het-c*^{GR} in *A. niger*. Co-expression of *het-c*^{PA} in *A. niger* lead to an abortive phenotype, suggesting that the indigenous *het-c* is functional and may cause heterokaryon incompatibility reactions in *A. niger*, though it probably does not function as a *het*-gene in this fungus.

2. Materials and methods

2.1 A. niger genome database searches for het-c, *vib-1*, and *pin-c*

Two *A. niger* genome databases were searched for the presence of the *N. crassa het-c* (Mylyk 1976, Howlett *et al.* 1993, Saupe and Glass 1997, Wu and Glass 2001), *vib-1* (Xiang and Glass 2002), and *pin-c* (Kaneko *et al.* 2006) homologs: (1) the genome of culture collection strain CBS513.88 from DSM (prepared by shotgun sequencing, 7.5 times coverage) and (2) the genome of strain ATCC1015 from DOE Joint Genome Institute (assembly release version 1.0 of whole genome shotgun reads was constructed with the JGI assembler, Jazz, using paired end sequencing reads at a coverage of 8.9 times).

2.2 Strains, characterization and culture conditions

In total 98 wild-type black *Aspergillus* strains and the laboratory strain N402 were tested for polymorphisms in the *het-c* gene. The strains were originally isolated exclusively on complete medium with 20 % tannic acid added (van Diepeningen *et al.* 2004) and originated from all over the world. They were classified based on their mitochondrial RFLP patterns as *A. niger* (1a-1d), *A. tubingensis* (2a-2f) (Varga *et al.* 1993, 1994), *A. japonicus* and close relatives (J) (Hamari *et al.* 1997) or *A. carbonarius* and close relatives (C) (Kevei *et al.* 1996) or based on ITS typing (Medina *et al.* 2005) (Table 1). Strains were grown on complete medium (CM, according to Pontecorvo *et al.* 1953) supplemented with 1 mg /l $ZnSO_4$, $FeSO_4$, $MnCl_2$ and $CuSO_4$. All incubations took place at 30 ºC.

Table 1. The wild-type strains used in the experiments ordered per black *Aspergillus* species. In bold the strains of which the *het-c4* region (Figure 1) was sequenced.

Species	Haplotype[a]	Strains[b]			
A. carbonarius	C	Isr1.1.1			
A. japonicus	J	F2.1	G1.5	Ind1.3.2	M1.3
		F2.2	G2.1	Kam2.1	
		F4.1	G2.6	Kam7.4	
A. niger	1a	Bra2.4	F3.2	G5.7	**N402**
		F3.1	F3.3		
	1b	**Bra2.5**	G3.9	Ind1.8.10	Nep1.6
		Cam6.6	G4.2	Nep1.1	NZ2.1
		Eg1.13	G5.1	Nep1.2	NZ2.2
		Eg2.5	**Gu1.1**	Nep1.3	NZ3.3
		G1.1	**Ind1.6.17**	Nep1.4	NZ3.4
		G3.3	Ind1.8.1	Nep1.5	
	1c	**Bra2.2**	Can2.5	G2.9	**Ind1.7.3**
		Bra2.6	Can2.7	G3.6	NZ3.1
		Bra2.34	Eg1.1	Gu1.6	NZ3.5
		Can2.2	Eg1.6	Ind1.4.24	
		Can2.4	F1.3	Ind1.5.3	
	n.d.[c]	Ind3.1.10	Ind8.1.1		
A. tubingensis	2a	Bra2.23	Eg2.1	Isr4.1.1	Ned 1.1
		Can2.1	Eg2.4	M1.1	Ned13.1
		Can2.3	G2.12	M1.2	Ned14.2
		Can2.6	Ind1.8.6	M1.5	Ned15.3
		Can2.8	Isr1.1.2	M2.1	**Ned16.3**
		Can2.9	Isr2.1.2	M2.3	**Ned23.2**
		Eg1.8	Isr3.1.1	M2.5	NZ3.2
		Eg1.16			
	2b	Bra2.1	F1.1	G4.4	M2.2
		Bra2.3	F1.2	Ind1.1.1	M2.4
		Bra3.1			
	n.d.[c]	Ind3.3.5			

[a] haplotype based on mitochondrial haplotype (Varga *et al.*, 1993, 1994; Kevei *et al.*, 1996; Hamari *et al.*, 1997). [b] Origin of the strains: Bra = Brazil; Cam = Cameroon; Can = Canada; Eg = Egypt; F = France; G = Gabon, Gu = Equatorial Guinea; Ind = Indonesia; Isr = Israel; M = Maroc; N = Niger culture collection; Ned = Netherlands; Nep = Nepal; and NZ = New Zealand. [c] Strains typed only as *A. niger* or *A. tubingensis*
based on ITS typing.

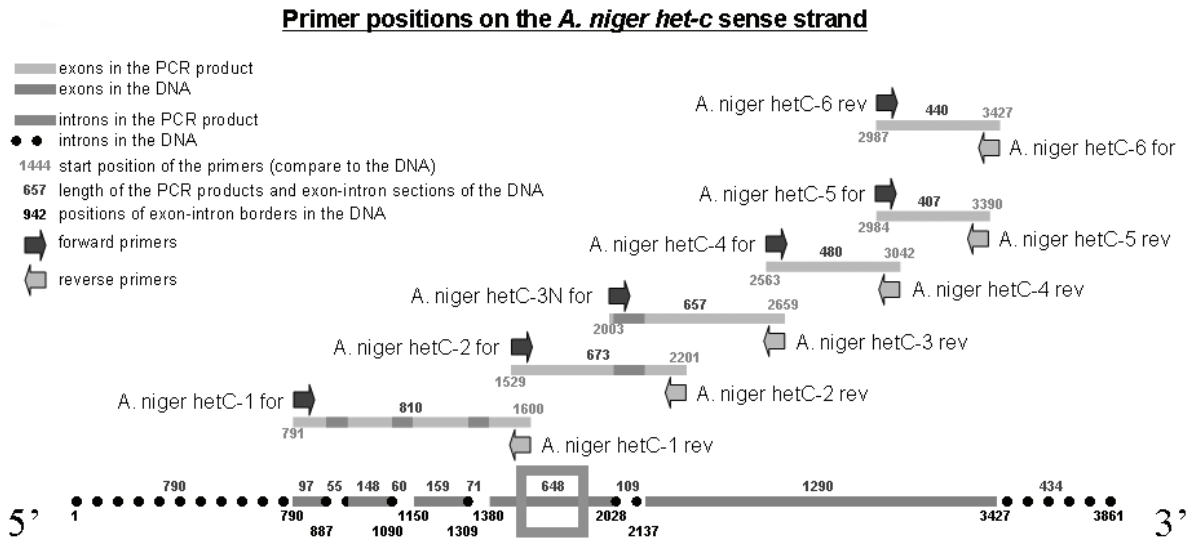


Figure 1. The primer positions and resulting PCR fragments of the six primer pairs used to screen the whole *het-c* gene for polymorphisms.

2.3 DNA isolation

An 0.1 g mycelial mat, overnight grown on liquid CM, was frozen with liquid nitrogen in a 1.5 ml Eppendorf tube and disrupted with a bead beater. Total nucleic acids were isolated using a phenol/chloroform extraction adapted after Maniatis *et al.* (1982): Preheated extraction buffer (0.5 M NaCl, 10 mM Tris-HCl, 10 mM EDTA, 1% SDS, pH 7.5) was added and the suspension was incubated at 65 ºC for 30 min. Nucleic acids were then extracted once with 1:1 phenol:chloroform and once with 24:1 chloroform:isoamyl alcohol. Isopropanol was used to precipitate the nucleic acids, they were washed with 70% ethanol and finally resuspended in demi water or TE buffer (10 mM Tris/ 10 mM EDTA, pH 7.5).

2.4 PCR and sequencing for het-c polymorphism

Based on the *het-c* sequences of the DSM *A. niger* strain (CBS513.88) five sets of primers were developed that covered the complete gene (Table 2). PCR amplifications were performed in 25 µl volume, with concentrations of the chemicals of the following: 0,35 units SuperTaq DNA Polymerase (HT Biotechnology, 5 units/µl), 200µM dNTPs (Promega), 0.2 µM primers (MWG Biotech), 2.5 µl PCR buffer (10 mM Tris-HCl pH=9,0, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1 % (w/v) Triton X-100, 0.01 % (w/v) Gelatin). The amplifications were done using a Biorad iCycler thermal cycler, and an initial temperature of 95°C for 1 min, 35 cycles at 95°C for 1 min, 56°C-68°C for 1 min, and 72°C for 2 min were carried out with a final step at 72°C for 5 min.

Products were separated by 1% TBE (tris-borate-EDTA) agarose electrophoresis

.

Table 2. Primer sets used for amplifying the *het-c* gene in the *A. niger* and other black *Aspergillus* strains.

Primer pairs	Sequence (5' → 3')	Tm (°C)	Amplified region (see Fig.1)	Length of PCR product
A. niger hetC-1for	ATG GCT CCG AGA ATG AGT TT	56	791-1600	810 bp
A. niger hetC-1rev	TTT CTT GCT TCC ATG GGT ATA T			
A. niger hetC-2for	TCT GCT TAC ATC AAG TAT AGC	56	1529-2201	673 bp
A. niger hetC-2rev	GAT CGA AGT CCT GTA GAG C			
A. niger hetC-3for	TCC GAA GCT CAG GAG TCT GCG	66	2003-2659	657 bp
A. niger hetC-3rev	CAG CGC TGG GTG GTG GAA CA			
A. niger hetC-4for	CGT TGC GCC TCG GGT TCT GTA	68	2563-3042	480 bp
A. niger hetC-4rev	GCT TCT TCG GTC TCG CCT CTC G			
A. niger hetC-5for	AAC AGC GCC TTC TCC AAT CT	65	2984-3390	407 bp
A. niger hetC-5rev	TTC CCG CGT CCA TAA CTC G			
A. niger hetC-6for	TCA ATA GTT CCC TCC ATA ATA G	59	2987-3427	440 bp
A. niger hetC-6rev	AGC GCC TTC TCC AAT CTG			

2.5 Introducing N. crassa het-c^{PA}, het-c^{OR}, and *het-c^{GR} in A. niger*

pCB1004 plasmids containing the *N crassa het-c^{PA}, het-c^{OR}*, or *het-c^{GR}* alleles and the hygromycin resistance marker *hygB,* were obtained from Dr. Sven Saupe's laboratory in Bordeaux (Saupe *et al.* 2000). Transformations were performed on the *A. niger* N814 laboratory strain, which bears auxotrophy for uridine and sensitivity to hygromycin. Protoplasts were co-transformed with two plasmids: pCB1004 plasmids contained one of the *N. crassa het-c* alleles and hygromycin-resistance, and an autonomously replicating helper plasmid, pAB4-ARp1, which contained the *pyrG* gene for uridine auxotrophy, as described by Verdoes *et al.* (1994). Transformants were plated on selective media, in which hygromycin concentration was 85 μg/ml. When upon introduction of a *het-c* allele no viable transformants are formed, heterokaryon incompatibility is concluded. Control protoplasts were not supplied with plasmids.

2.6 Comparison of homologs of different origin

Homologous polypeptides were aligned with ClustalX (Thompson *et al.* 1997) and edited with Boxshade 3.21 (written by Kay Hofmann and Michael D. Baron and available at http://www.ch.embnet.org/software/BOX_form.html). The degree of identity of proteins was described with blasting the proteins against each other by BLOSUM62 matrix (Tatusova and Madden 1999, at http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi,). Neighbour-joining phylogenetic trees were also made with Clustal X and displayed with TreeView 1.6.6 (Page 1996). Comparison of proteins to get the degree of identity and similarity was performed on the NCBI homepage in the 'Blast' menu with the 'Align two sequences (bl2seq)' option (http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi). Amino acid composition of different proteins (in percent) was calculated by ProtParam (http://www.expasy.org/tools/protparam.html).

2.7 Genome databases used for BLAST search

A. fumigatus preliminary sequence data was obtained from The Institute for Genomic Research website at http://www.tigr.org. *A. nidulans*, *N. crassa* (release 7) and *A. terreus* sequence data were from the *Aspergillus nidulans*, *Neurospora crassa* and *Aspergillus terreus* Sequencing Project, Broad Institute of MIT and Harvard (http://www.broad.mit.edu). *A. oryzae* sequences were available on the server of National Institute of Technology and Evaluation (NITE). *P. anserina* genome was published by the Institut de Génétique et Microbiologie (Université de Paris-Sud XI / CNRS). Origin of the two investigated *A. niger* databases: the genome of the culture collection strain CBS513.88 from DSM Research BV (prepared by shotgun sequencing, 7.5 times coverage) and the genome of the culture collection strain ATCC1015 from DOE Joint Genome Institute (assembly release version 1.0 of whole genome shotgun reads was constructed with the JGI assembler, Jazz, using paired end sequencing reads at a coverage of 8.9X).

3. Results

3.1 Comparing the A. niger and N. crassa genomes for het-c genes and related sequences

Blasting the two *A. niger* genomes with the *N. crassa het-c* resulted in only one homologous sequence, whereas two were found in the examined other *Aspergillus* and *Sordaria* species (see Chapter 2, Table 4.). The CBS513.88 HET-C protein is 793 amino acids long and the length of the ATCC1015 HET-C protein is 792 amino acids. The two proteins are almost 100% identical. Two minor differences exist: an extra glutamine unit in the CBS513.88 HET-C protein (at position of 622 AA on the consensus) and a serine-phenylalanine substitution (at position of 380 AA on the consensus). A silent polymorphism was found in the *het-c* gene at position of 1815 nucleotide, a serine coding TCC triplet in the CBS513.88 is changed to the also serine coding TCT triplet in the ATCC1015 (Figure 2).

Comparison of the two *A. niger* HET-C alleles with the *N. crassa* HET-C^{OR} (GenBank accession number is AAB48349) showed that conservation is strong at the N-terminal part of the proteins (first 530 amino acids of the *A. niger* proteins) and decreases toward the C-terminal end. The same fact was found when the *N. crassa* HET-C^{OR} protein was compared with the *P. anserina* HCH (HET-C homolog, Pa_6_4390) (Saupe *et al.* 2000, Figure 2.). Identity of the two *A. niger* HET-C proteins to the *N. crassa* HET-C^{OR} allele is 56%, while the identity between *A. niger* HET-C proteins and the *P. anserina* HCH is 59%.

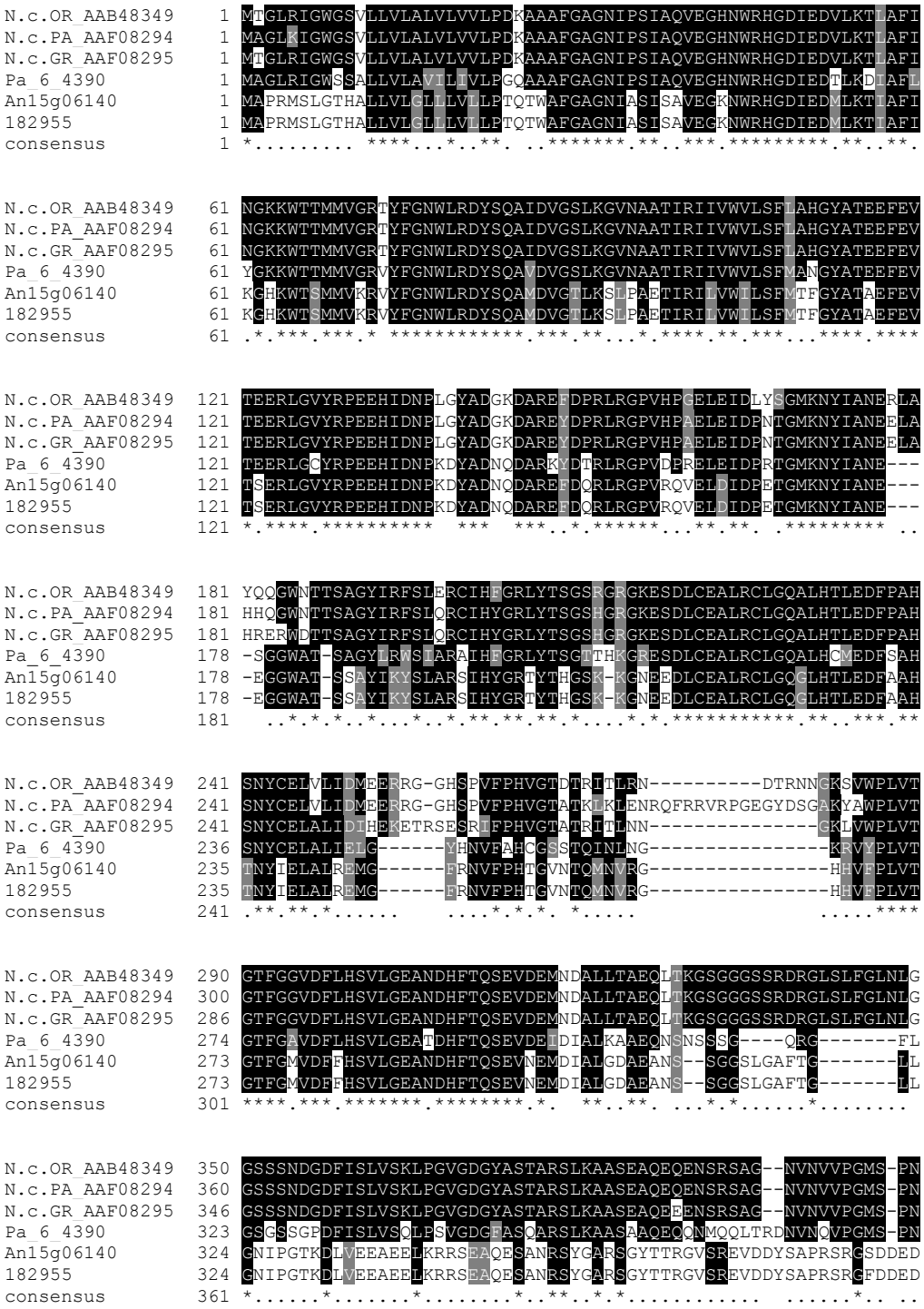


Figure 2. The three *N. crassa het-c* alleles Oakridge (OR), Panama (PA), and Groveland (GR) compared to the *het-c* genes detected in the sequenced genome databases of *A. niger* strains CBS513.88 (An15g06140) and ATCC1015 (Aspni_182955).

.

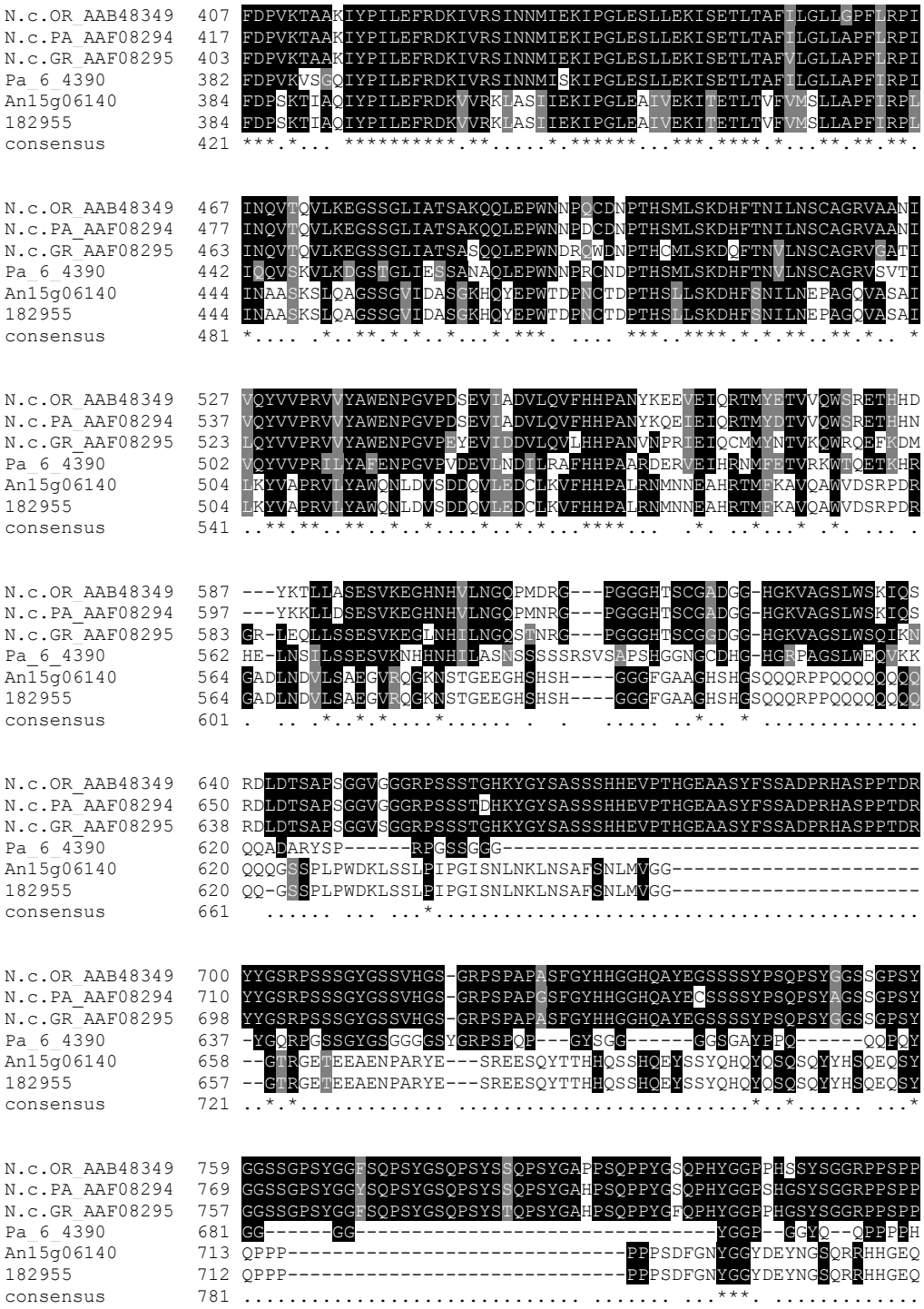


Figure 2. (Continued) The three *N. crassa het-c* alleles Oakridge (OR), Panama (PA), and Groveland (GR) compared to the *het-c* genes detected in the sequenced genome databases of *A. niger* strains CBS513.88 (An15g06140) and ATCC1015 (Aspni_182955).

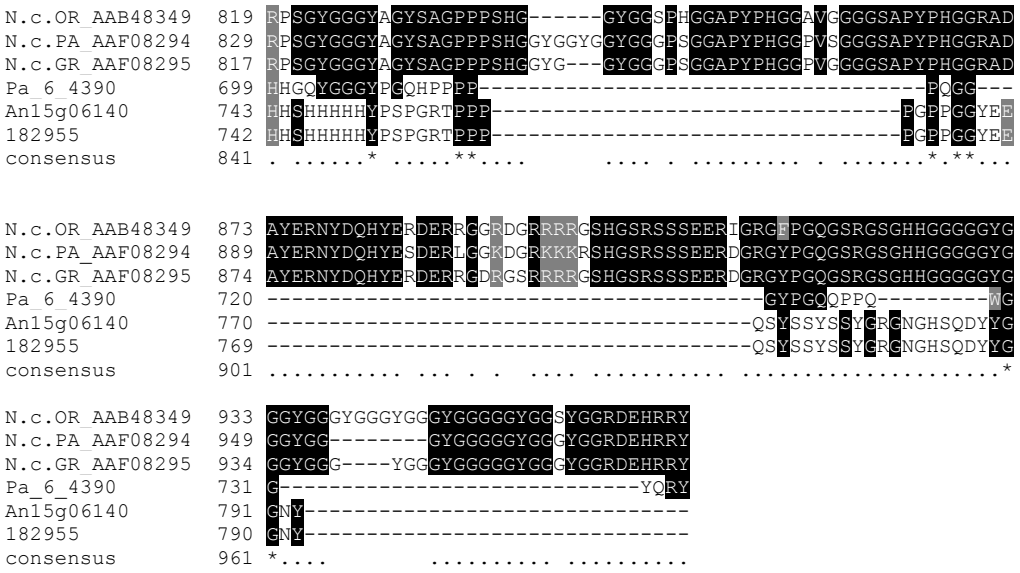


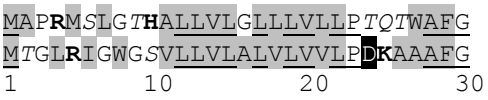
Figure 2. (Continued) The three *N. crassa het-c* alleles Oakridge (OR), Panama (PA), and Groveland (GR) compared to the *het-c* genes detected in the sequenced genome databases of *A. niger* strains CBS513.88 (An15g06140) and ATCC1015 (Aspni_182955).

The first 30 amino acids at the N-terminal part of the *A. niger* HET-C display similarity to the supposed signal peptide sequence found in the *N. crassa* HET-C (Saupe *et al.* 1996, Figure 3A). However, the requirements to enter the secretory pathway are not totally fulfilled: The hydrophobic block is interrupted by glycine and the C-terminal end of the proposed signal sequence does not contain charged amino acids. Therefore it is not clear whether this 30 amino acids stretch serves as a signal sequence or not.

Similarly to the *P. anserina* HCH, the *A. niger* HET-C protein does not bear the indel region responsible for allelism in the *N. crassa* HET-C (in position at 277-292 on consensus, Figure 2.). PCR analysis of this region in 99 black *Aspergillus* isolates also did not show any size polymorphism (see below).

A similarity between the *A. niger* and *N. crassa* HET-C proteins is the presence of a heptad repeat structure (in position at 440-472 on consensus), with hydrophobic residues at positions **a** and **d** of the heptad (Figure 3B and 3C). This motif was found in protein domains with a coiled-coil structure, such as several transcriptional regulators or myosin. In position **a** and **d** only hydrophobic residues are grouped in the heptads (Saupe *et al.* 1996).

A

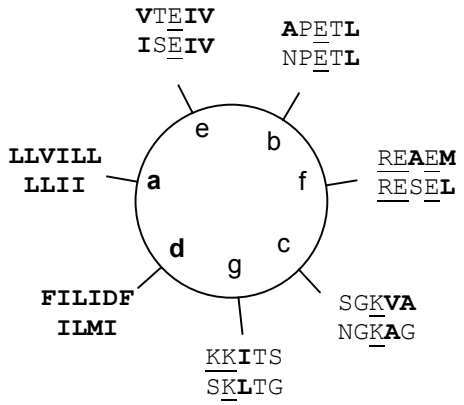


A. niger CBS513.88 HET-C
N. crassa HET-C^{OR}

B

A. niger CBS513.88 HET-C
N. crassa HET-C^{OR}

C



Top: *A. niger*
Bottom: *N. crassa*

Legends: bold: hydrophobic residues
underlined: charged residues

Figure 3. Comparison of (**A**) signal peptide region and (**B**) coiled-coil heptade repeat structure between *A. niger* and *N. crassa*. Legends for (**A**) and (**B**): grey box: hydrophobic residues (A, F, L, M, V, W); black box: negatively charged amino acid (D); bold: positively charged amino acids (H, K, R); italicized: polar, but uncharged residues (Q, S, T); underlined: identical amino acids in the same position. (**C**) is the helical wheel representation of the coiled-coil heptades.

The carboxy-terminal end of the *A. niger* HET-C protein (the last 170 amino acids) does not resemble the C-terminal end of the *N. crassa* HET-C (Figure 2). Long deletions in this region are the reason why the *A. niger* HET-C is shorter than that of *N. crassa*. Comparing this region to the *N. crassa* HET-C^{OR} the *A. niger* HET-C is 122 amino acids shorter. Therefore, it does not contain the glycine-rich domain found in *N. crassa* (Saupe *et al.* 1996) and its amino acid composition is different, as well. Major components of the C-terminal part of the *A. niger* HET-C (final 170 amino acids) are 15.8 % serine (27), 11.7 % glycine (20), 11.1 % proline (19) and tyrosine (19), 8.8 % glutamine (15) and histidine (15) and 7.6 % glutamic acid (13). The numbers in brackets show the total number of different residues in the 170 amino acids long C-terminal end. Interestingly, there is no cysteine in the C-terminal part of the protein. There are two proline blocks (of 6 and 4 proline residues, respectively) and a histidine block (of 5 histidine units) downstream of the glutamine-repeat.

The *A. niger het-c* gene contains 4 putative introns (of 56-, 61-, 72- and 70 bp, respectively). The *N. crassa het-c* contains only two introns (of 66- and 65 bp, respectively) and the *P. anserina hch* gene contains three (of 52-, 55- and 59 bp, respectively). The first and second intron in the *A. niger het-c* seem to be more related to the corresponding ones in the *P. anserina hch* than to the *N. crassa* introns.

Pin-c and *vib-1* are two loci involved in *het-c* related incompatibility in *N. crassa*. The first is linked to *het-c* in *N. crassa* and these proteins together interact with unlike *het-c*/*pin-c* alleles and trigger incompatibility. The incompatibility process is governed by the *vib-1*, as a central part of the incompatibility (Kaneko *et al.* 2006). Homologs of both PIN-C and VIB-1 were found in both *A. niger* genomes. Blasting with PIN-C resulted in many hits (due to its HET-domain), but we found only one VIB-1 homolog in both *A. niger*s (two were found in the investigated *Sordariales* and other *Aspergillus* species, see Chapter 2.). Unlike in *N. crassa*, the *A. niger het-c* and *pin-c* genes are localized in different contigs, the locations are (CBS513.88 / ATCC1015, respectively): *pin-c* on An09c0170 / scaffold11 and *het-c* on An15c0200 / scaffold6.

3.2 Screening a population of wild-type black Aspergilli for het-c polymorphisms

We tested 99 black *Aspergillus* isolates (Table 1) from our strain collection to find out whether a similar length polymorphism exists between the different natural black *Aspergillus* isolates in the *het-c* gene as that was found in *N. crassa*. Six primer pairs were developed (Table 2), that overlap each other and cover the *het-c* gene (Figure 1). We did not find any size difference in the *N. crassa* allelic indel region, but in another region (amplified with the *het-c4* primer pair) we found size differences (480-550 bp) in the PCR products, (Figure 4). PCR products of this variable sized region were sequenced from nine randomly chosen strains, and the size polymorphism proved due to a variable stretch of glutamine repeats (Figure 5). The number of glutamine units is independent of the species; there is no correlation to *A. niger* and *A. tubingensis*. The role of glutamine repeats is not clear. Polymerase slippage during DNA amplification in the different strains may have led to the observed size differences (Zoghbi and Orr 2000).

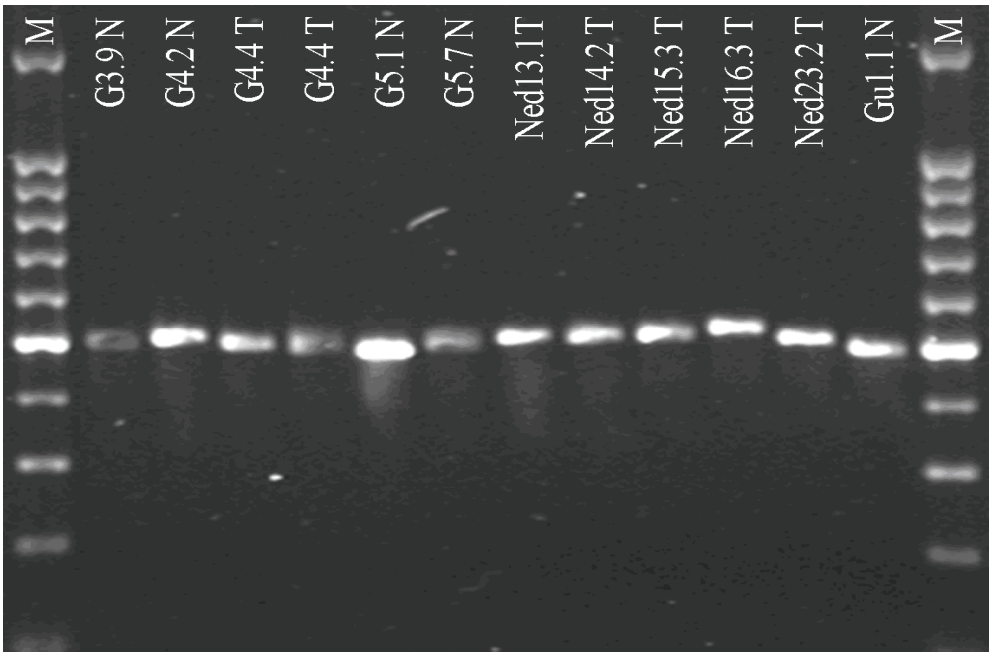


Figure 4. Photograph of a gel with different black *Aspergillus* strains showing polymorphisms in the *het-c* locus with primer pair hetC-4for and hetC-4rev. Expected product size 480 bp, found range approximately 480 bp-550 bp. N = *A.* niger, T = *A. tubingensis.* Markers: 100bp ladder with extra bright band at 500 bp.

```
CBS513.88     1 FGAAGHSHGSQQQRPPQQQ----QQQQQQQQ---GSSPLPWDKLSSLPIP
Ned16.3       1 FGAAGHSHGSQQQRPPQQQ----QQQQQQQQQQQGSSPLPWDKLSSLPIP
Ind1.7.3      1 FGAAGHSHGSQQQRPPQQQ----QQQQQQ-----GSSPLPWDKLSSLPIP
N402          1 FGAAGHSHGSQQQRPPQQQ----QQQQQQQ----GSSPLPWDKLSSLPIP
Ned23.2       1 FGAAGHSHGSQQQRPPQQQ----QQQQQQQ----GSSPLPWDKLSSLPIP
Bra3.1        1 FGAAGHSHVSQQQRPPQQQ----QQQQ-------GSSPLPWDKLSSLPIP
Gu1.1         1 FSAAGHSHGSQQQRPPQQQ----QQQQQQ-----GSSPLPWDKLSSLPIP
Bra2.2        1 FSAAGHSHGSQQQRPPQQQRPPQQQQQQ-----GSSPLPWDKLSSLPIP
Ind1.6.17     1 FSAAGHSHGSQQQRPPQQQ----QRSQQQ-----GSSPLPWDKLSSLPIP
Bra2.5        1 FSAAGHSHGSQQQRPPQQQ----Q----------GSSPLPWDKLSSLPIP
consensus     1 *.******.**********    *.....     ****************
```

Figure 5. Protein sequences of the variable glutamine rich region, probably caused by polymerase slippage, in ten randomly sequenced strains.

3.3 Testing the functionality of the A. niger het-c gene

The absence of significant polymorphism for *het-c* in *A. niger* suggests that this gene is not involved in heterokaryon incompatibility among the isolates tested. In order to find out whether a heterokaryon incompatibility reaction can be triggered by one of the known *het-c* alleles from *N. crassa*, we transformed protoplasts of *A. niger* with plasmids containing either of the three known alleles of *N. crassa het-c*. Transformants containing the *het-c*^{PA} allele showed the most severe growth inhibition: only abortive transformants were obtained. Protoplasts transformed with *het-c*^{GR} or *het-c*^{OR} showed normal growth (Figure 6). Since media contained hygromycin and no uridine, only the successfully co-transformed protoplast could grow, which contained both the pCB1004 plasmid (with hygromycin resistency) and the pAB4-ARp1 plasmid (with the *pyrG* gene).

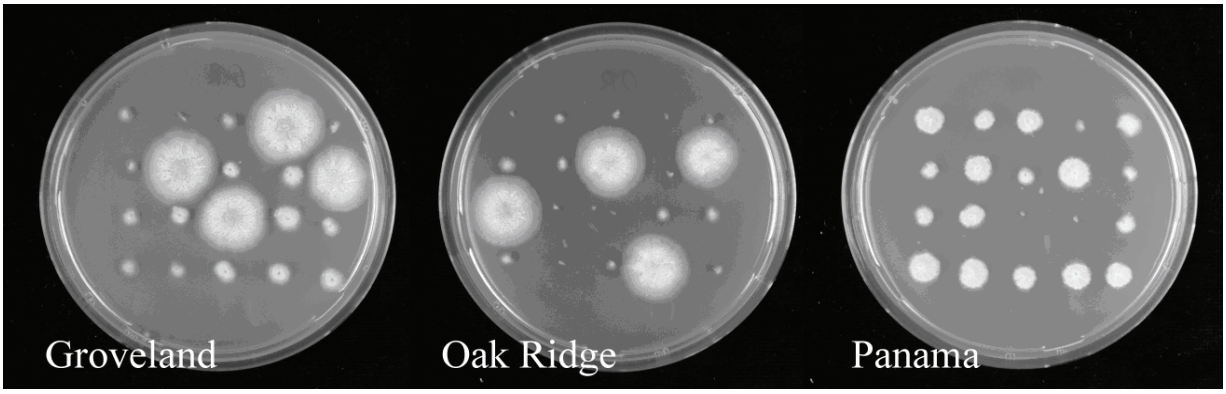


Figure 6. Photograph of Petri dished with each twenty lines of *A.* niger, transformed with either the Groveland, Oakridge or Panama *het-c* alleles of *N.* crassa. The Groveland and Oakridge alleles yield viable transformants among the non-transformed strains. The Panama allele did not yield any well-growing transformants.

4. Discussion

Heterokaryon incompatibility among *A. niger* strains is a widespread phenomenon that is observed as the inability to form stable heterokaryons and to exchange viruses and mitochondrial markers. Heterokaryon incompatibility is shown to be a genetically regulated process in several sexual filamentous fungi like *N. crassa* and *P. anserina,* when crosses between heterokaryon incompatible strains are analysed for the number of segregating *het*-loci. But, for the (presumably) asexual *A. niger* the genetic basis of heterokaryon incompatibility is unknown as genetic data from crosses cannot be obtained. Nowadays with the availability of two whole sequenced genomes of incompatible *A. niger* strains one can *in silico* screen for polymorphisms in putative *het*-genes, homologue to known *het*-genes in sexual species and than use the sequence data to screen for polymorphisms within a population. We screened the *A. niger* population for polymorphism in the known *het*-gene *het-c* from *N. crassa* but did not find any, though

introduction of the alleles from *N. crassa* by transformation shows that the gene can be functional in heterokaryon incompatibility.

In *N. crassa* sofar, three different *het-c* alleles have been identified in which the specificity is dependent on an indel encoding 30-48 amino acids (Mylyk 1976, Howlett *et al.* 1993, Saupe and Glass 1997, Wu and Glass 2001). Saupe *et al.* (2000) tested 11 natural *P. anserina* strains without finding evidence for polymorphism in the *het-c* homologue *hch*. We tested 99 black *Aspergillus* strains belonging to the species *A. carbonarius*, *A. japonicus*, *A. niger* and *A. tubingensis*, that all proved to contain a *het-c* homologous gene, but we did not find any polymorphisms in the active indel region homologous to the corresponding *N. crassa* region. We found some variation in a different part of the *het-c* gene. Sequencing of this variable region showed that this size difference is caused by different numbers of glutamine coding triplets in a glutamine rich region, presumably without a heterokaryon incompatibility function. Glutamine repeats with variable number of glutamine residues might originate from polymerase slippage (Zoghbi and Orr 2000). Glutamine repeats play a role in neurodegenerative human diseases, like Huntington-disease. Poly-glutamine repeats may function as polar zippers and cause protein aggregation (Perutz *et al.* 1994, Rega *et al.* 2001, Masino and Pastore 2002), but its function for *A. niger* is unknown. Nevertheless, although both huntingtin and the *A. niger* HET-C bear glutamine-repeats and proline blocks, differences are quite remarkable in size and position of these units.

The *A. niger* HET-C protein shows strong similarity to the *N. crassa* HET-C and the *P. anserina* HCH at the N-terminal part, and weaker similarity at the C-terminal end. When HET-C proteins from two different isolates of *A. niger* (CBS513.88 and ATCC1015) were compared, only two amino acids were found as difference. The *A. niger* HET-C protein shows some similar features to the *Fusarium proliferatum* HCH protein. The latter protein was described in a recent article (Kerényi *et al.* 2006), and it is 770 amino acids long and like *A. niger* the open reading frame contains four introns.

Only heterologous expression of the *N. crassa het-c*^{PA} and not of the other two alleles in *A. niger* by transformation showed a totally abortive phenotype. This strongly suggests that the indigenous *het-c* is functional and may cause heterokaryon incompatibility reactions in the black Aspergilli. Formally, however, there are two explanations for this phenomenon: either the *N. crassa het-c*^{PA} allele itself is toxic when expressed in *A. niger*, or this allele interacts with the *A. niger het-c*, influencing directly growth and viability. This finding is similar to that found in *P. anserina* (Saupe *et al.* 2000). But, in contrast, transformation of *F. proliferum* with the *N. crassa het-c* did not result in heterokaryon incompatibility. This finding is interesting, since the *F. proliferum* HCH shows a higher level of homology to the *N. crassa* HET-C^{GR} (63% identity and 73% similarity) than the *A. niger* HET-C (55% identity and 67% similarity).

Homologs to the different suppressor genes that interact with the *het-c* locus in *N. crassa*, the *vib-1* gene (Xiang and Glass 2002) and the *pin-c* gene (Kaneko *et al.* 2006) were also found in *P. anserina* and in *A. niger* (Pál *et al.* in preparation; Chapter 2). In *N. crassa pin-c* and *het-c* are linked, but these loci are located in different contigs in *A. niger*.

Based on our findings it seems that *het-c* does not act as a *het*-gene in *A. niger*. Comparison of HET-C proteins from two different *A. niger* strains revealed only negligible differences (two amino acids were different) and PCR analysis of 99 black *Aspergillus* strains taken from a global population, only showed small variations in the number of glutamine. This suggests that allelic interaction, similar to that found between *N. crassa het-c* alleles, does not exist in *A. niger,* though, we cannot exclude extremely unbalanced allele distributions as found in *Cryphonectria* (Milgroom and Cortesi 1999).

Acknowledgments

We are grateful to dr. Sven Saupe for providing us with the plasmids with the different *N. crassa het-c* genes to transform the *A. niger* with. We thank to the DSM Research BV for the genome of the *A. niger* CBS513.88 strain and to the DOE Joint Genome Institute for the genome of the *A. niger* ATCC1015 strain.

References

Caten CE, 1972. Vegetative incompatibility and cytoplasmic infection in fungi. *Journal of General Microbiology* 72, 221-229.

Glass NL, Grotelueschen J, Metzenberg RL, 1990. *Neurospora crassa* A mating-type region. *Proceedings of the National Academy of Sciences of the USA* 87, 4912-4916.

Glass NL, Kaneko I, 2003. Fatal attraction, nonself recognition and heterokaryon incompatibility in filamentous fungi. *Eukaryotic Cell* 2003, 1-8.

Glass NL, Kuldau GA, 1992. Mating type and vegetative incompatibility in filamentous ascomycetes. *Annual Reviews in Phytopathology* 30, 201–224.

Glass NL, Vollmer SJ, Staben C, Metzenberg RL, Yanofsky C, 1988. DNAs of the two mating type alleles of *Neurospora crassa* are highly dissimilar, *Science* 241, 570–573.

Hamari Z, Kevei F, Kovács É, Varga J, Kozakiewicz Z, Croft JH, 1997.. Molecular and phenotypic characterization of *Aspergillus japonicus* strains with special regard to their mitochondrial DNA polymorphisms. *Antonie van Leeuwenhoek* 72, 337-347.

Howlett BJ, Leslie JF, Perkins DD, 1993. Putative multiple alleles at the vegetative , heterokaryon. incompatibility loci *het-c* and *het-8* in *Neurospora crassa*. *Fungal Genetics Newsletter* 40, 40-42.

Kaneko I, Dementhon K, Xiang QJ, Glass NL, 2006. Nonallelic interactions between het-c and a polymorphic locus, pin-c, are essential for nonself recognition and programmed cell death in *Neurospora crassa*. *Genetics* 172, 1545-1555.

Kerényi Z, Oláh B, Jeney A, Hornok L, Leslie JF, 2006. The homologue of *het-c* of *Neurospora crassa* lacks vegetative compatibility function in *Fusarium proliferatum*. *Applied and Environmental Microbiology* 72, 6527-6532.

Kevei F, Hamari Z, Varga J, Kozakiewicz Z, Croft JH, 1996. Molecular polymorphism and phenotypic variation in *Aspergillus carbonarius*. *Antonie van Leeuwenhoek* 70, 59-66.

Leslie JF, 1993. Fungal vegetative incompatibility. *Annual Review of Phytopathology* 31,127–150.

Leslie JF, Zeller KA, 1996. Heterokaryon incompatibility in fungi – more than just another way to die. *Journal of Genetics* 75, 415-424.

Maniatis T, Fritsch EF, Sambrook J, 1982. *Molecular cloning, a laboratory manual*. Cold Spring Harbor laboratories, New York.

Masino L, Pastore A, 2002. Glutamine repeats: structural hypotheses and neurodegeneration. *Biochemical Society Transactions* 30, 548-551.

Milgroom MG, Cortesi P, 1999. Analysis of population structure of the chestnut blight fungus based on vegetative incompatibility genotypes. *Proceedings of the National Academy of Sciences of the USA* 96, 10518-10523.

Mylyk MO, 1976. Heteromorphism for heterokaryon incompatibility genes in natural-populations of *Neurospora crassa*. *Genetics* 83, 275-284.

Perkins DD, 1988. Main features of vegetative incompatibility in *Neurospora crassa*. *Fungal Genetics Newsletters* 35, 44-46.

Perutz MF, Johnson T, Suzuki M, Finch JT, 1994. Glutamine repeats as polar zippers: their role in inherited neurodegenerative diseases. *Proceedings of the National Academy of Sciences of the USA* 91, 5355-5358.

Pontecorvo G, Roper JA, Forbes E, 1953. Genetic recombination without sexual reproduction in *Aspergillus niger*. *Journal of General Microbiology* 8, 198–210.

Saupe SJ, 2000. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews* 64, 489-502.

Saupe SJ, Clave C, Sabourin M, Begueret J, 2000. Characterization of *hch*, the *Podospora anserina* homolog of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa*. *Current Genetics* 38, 39-47.

Saupe SJ, Glass NL, 1997. Allelic specificity at the *het-c* heterokaryon incompatibility locus of *Neurospora crassa* is determined by a highly variable domain. *Genetics* 146, 1299-1309.

Saupe SJ, Turcq B, Begueret J, 1995. Sequence diversity and unusual variability at the het-c locus involved in vegetative incompatibility in the fungus *Podospora anserina*. *Current Genetics* 27, 466-471.

Saupe SJ, Kuldau GA, Smith ML, Glass NL, 1996. The product of the *het-C* heterokaryon incompatibility gene of *Neurospora crassa* has characteristics of a glycine rich cell-wall protein. *Genetics* 143, 1589-1600.

Saupe SJ, Clave C, Sabourin M, Begueret J, 2000. Characterization of *hch*, the *Podospora anserina* homolog of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa*. *Current Genetics* 38, 39-47.

Smith ML, Micali OC, Hubbard SP, Mir-Rashed N, Jacobson DJ, 2000. Vegetative incompatibility in the *het-6* region of *Neurospora crassa* is mediated by two linked genes. *Genetics* 155, 1095–1104.

van Diepeningen AD, Debets AJM, Hoekstra RF, 1997. Heterokaryon incompatibility blocks virus transfer among natural isolates of black *Aspergilli*. *Current Genetics* 32, 209-217.

van Diepeningen AD, Debets AJM, Varga J, van der Gaag M, Hoekstra RF, 2004. Efficient degradation of tannic acid by black *Aspergillus* species. *Mycological Research* 108, 919-925.

Varga J, Kevei F, Fekete C, Coenen A, Kozakiewicz Z, Croft JH, 1993. Restriction fragment length polymorphisms in the mitochondrial DNAs of the *Aspergillus niger* aggregate. *Mycological Research* 97, 1207-1212.

Varga J, Kevei F, Vriesema A, Debets F, Kozakiewicz Z, Croft JH, 1994. Mitochondrial DNA restriction fragment length polymorphisms in field isolates of the *Aspergillus niger* aggregate. *Canadian Journal of Microbiology* 40, 612-621.

Verdoes J, Punt PJ, van der Berg P, Debets F, Stouthamer AH, van den Hondel CA, 1994. Characterization of an efficient gene cloning strategy for *Aspergillus niger* based on an autonomously replicating plasmid: cloning of the *nicB* gene of *A. niger*. *Gene* 146, 159-165.

Wu J, Glass NL, 2001. Identification of specificity determinants and generation of alleles with novel specificity at the *het-c* heterokaryon incompatibility locus of *Neurospora crassa. Molecular and Cellular Biology* 21, 1045-1057.

Xiang QJ, Glass NL, 2002. Identification of *vib-1*, a locus involved in vegetative incompatibility mediated by *het-c* in *Neurospora crassa*. *Genetics* 162, 89-101.

Xiang QJ, Glass NL, 2004. The control of mating type heterokaryon incompatibility by *vib-1*, a locus involved in *het-c* heterokaryon incompatibility in *Neurospora crassa*. *Fungal Genetics and Biology* 41, 1063-1076.

Zoghbi HY, Orr HT, 2000. Glutamine repeats and neurodegeneration. *Annual Review of Neuroscience* 23, 217-247.

CHAPTER 4

Analysis of the *hch* and *het-c* genes in natural isolates of *Podospora anserina*

Authors:

Károly Pál, Anne D. van Diepeningen, A. Bertha Koopmanschap, Rolf F. Hoekstra

and Alfons J.M. Debets

Laboratory of Genetics, Plant Sciences, Wageningen University, Wageningen,

The Netherlands.

Abstract

Heterokaryon (or vegetative) incompatibility in filamentous fungi is very common between genetically unlike individuals. It is governed by the so called *het*-genes, which can interact both in allelic or non-allelic reactions in ascomycetes. In this research we screened our population of wild-type *P. anserina* strains for two different (potential) *het*-genes. The first is the homologue of the allelic *Neurospora crassa* heterokaryon incompatibility gene *het-c*, called *hch* in the *Podospora anserina* genome. This *hch* gene did not show any polymorphism in the *P. anserina* population when we screened our natural population of 113 isolates with CAPS analysis and partly by sequencing and thus does not seem to be functional as *het*-gene. The second gene is the *P.anserina het-c* locus that is involved in non-allelic incompatibility and interacts with the *het-d* and *het-e* genes. Fifteen random natural *P. anserina* strains were sequenced for their *het-c* gene; three of the four known *het-c* alleles were found in six isolates, but in the other nine wild-type strains five novel alleles were detected. The mostly non-synonymous substitutions and distribution of the changed amino acids along the HET-C protein point to functionality and even positive selection for new alleles at this locus.

Keywords: Podospora anserina, allelic/non-allelic incompatibility, *het-c*, *hch*, CAPS analysis, sequence analysis.

1. Introduction

Filamentous fungi are able to form anastomoses, which are a prerequisite for the formation of heterokaryotic cells in which two genetically different nuclei coexist. The viability of these cells is controlled by the *het* or *vic* loci, which are responsible for the heterokaryon or vegetative incompatibility. If the parental strains are isogenic in their *het*-genes, a stable heterokaryon can be formed, but differences between *het* alleles and genes trigger incompatibility, which is manifested as growth inhibition or even rapid destruction of heterokaryotic (and often the surrounding) cells (Glass and Kaneko 2003). Any locus at which heteroallelism leads to cell death, can behave as a *het* locus (Saupe 2000).

Two systems - one allelic the other non-allelic – govern heterokaryon incompatibility. In the first case two or more alleles of the same locus interact, whereas in the latter one two distinct loci trigger the incompatibility reaction (Saupe 2000). The heterothallic ascomycete fungus *Podospora anserina* is a model system for both allelic and non-allelic incompatibility, bearing at least nine *het* loci consisting of both types of incompatibility genes (Saupe 2000).

The *hch* (*het-c* homolog) gene is a homolog of the *Neurospora crassa het-c* gene where it takes part in allelic incompatibility reactions. Three different alleles of *het-c* are known in *N. crassa.* The Oak Ridge allele of *het-c* (*het-c*^{OR}) in *N. crassa* encodes a 966-amino acid polypeptide with a putative signal peptide, a coiled-coil motif and a C-terminal glycine-rich domain, similar to glycine-rich domains detected in various extracellular and structural cell envelope proteins. Both the coiled-coil and one-third of the glycine-rich carboxyl terminal domains were required for full activity in *N. crassa* (Saupe *et al.* 1996). However, in *P. anserina* there was no multiple allelism detected in 11 *P. anserina* isolates (collected in France) in the *hch* locus in that part of the gene which is homologous to the *N. crassa* allelic region (Saupe *et al.* 2000). But, contrary to the lack of polymorphisms, the *hch* was able to show incompatibility features when different alleles of *N. crassa het-c* were introduced into *P. anserina* (Saupe *et al.* 2000).

The *P. anserina het-c* (which is not identical to that of in *N. crassa*) is an interesting example of a non-allelic *het*-gene, for it interacts with both the *het-d* and *het-e* genes and triggers incompatibility. There are four known alleles of *het-c*, four of *het-e* and three of *het-d*, but only certain alleles interact with each other (Saupe *et al.* 1994). In addition, several other genes were found which are possible *het-d*/*het-e* paralogs. Whether these paralogs interact with the *het-c* and even with new alleles of *het-c* is not known yet (Sven Saupe, personal comm.).

The *het-c* resembles a glycolipid transfer protein (GTP) isolated from the pig brain (Saupe *et al.* 1994) while the antagonistic *het-d* and *het-e* genes are paralogs of each other and contain TOL, NACHT (named after the NAIP, CIITA, HET-E and TP1 domains) and WD (named after Trp-Asp, the W-D repeats) domains. In addition, the *het-e* contains a specific P-loop with GTP-ase activity located in the NACHT domain, which is required together with at least 10 WD domains to interact with the *het-c* (Espagne *et al.* 1997). Next to the *het-d* and *het-e* there are other genes which contain variable numbers of WD repeats and have an allorecognition function (called

HNWD or NWD genes depending on the presence/absence of a HET domain). The HET domain is composed of three blocks of high amino acid similarity within an 200-aa region and is conserved among proteins involved in heterokaryon incompatibility in both *N. crassa* and *P. anserina* (Espagne *et al.* 2002), but can also be found in other genes. The WD repeats are of key importance in the generation of new *het-d* and *het-e* alleles, due to reshuffling of WD repeat variants between each other and other members of the NWD group (Paoletti *et al.* unpublished).
In this study we have examined a population of 113 natural isolates of *P. anserina* collected in Wageningen, the Netherlands, focussing on two *het* loci, the *het-c* and *hch.* The questions we address in this study are: 1) can we find polymorphism in the region homologous to the *N. crassa* allelic region of the *hch* gene in 113 isolates of the natural population of *P. anserina*? For this we performed CAPS (*Cleaved Amplified Polymorphic Sequence*) analysis on the *hch* locus in all strains and sequenced a subset of ten random strains for the known polymorphic region of the *hch* gene. 2) Can we find polymorphism in the *het-c* gene in the Wageningen natural population of *P. anserina*? For this we randomly selected 15 strains in which the whole *het-c* gene was sequenced. With the obtained results we speculate on the functionality of the uniform *hch* gene and the polymorphic *het-c* gene in the natural population.

2. Materials and methods

2.1 Strains and culture collections

In this study we used 113 wild-type *P. anserina* strains, numbered from Wa01 till Wa119, of our wild-type collection. These wild-type strains of the pseudo-homothallic ascomycete *P. anserina* were dikaryotic and self-fertile. The strains were isolated between 1991 and 2001 from herbivore dung collected in the surroundings of Wageningen, The Netherlands (van der Gaag *et al.* 1998, van der Gaag 2005, Maas 2005).

2.2 DNA isolation

DNA was isolated from 113 *P. anserina* samples. Mycelium for DNA extraction was grown on CornMeal agar covered with cellophane in a Petri dish for 48-72 hours at 27 ºC in the dark (Esser 1974). After 2-3 days of incubation the mycelium was harvested in 2 ml Eppendorf tubes, avoiding the heavily pigmented parts. After freezing with liquid nitrogen the mycelium was ground with glass beads of 2-3 mm in diameter with bead beater (Silamat S5, Ivoclar Vivadent AG, Austria). The freezing and grinding was repeated twice. To the ground mycelium 1 ml of LETS buffer (0.1 M LiCl, 10 mM EDTA pH=8.0, 10 mM Tris pH=8.0, 0.5 % SDS) was added, vortexed and than centrifuged for 5 min at 13.000 rpm. 700 µl of supernatant was transferred into a fresh 2 ml Eppendorf tube, 5 µl of proteinase K was added and incubated for 1 hour on 37 ºC. After this treatment a phenol/chloroform extraction was performed (Sambrook *et al.* 1989). 500 µl of the supernatant of the choloroform extraction was transferred into a new 1.5 ml Eppendorf

tube, 0.6 volume isopropanol was added and this mixture was stored on -20 ºC for 15 min or 4 ºC overnight. The mix was centrifuged for 15 min at 13.000 rpm, the supernatant was discarded and the pellet was washed with 100 µl, -20 ºC, 70 % ethanol. After centrifuging for 1 min at 13.000 rpm, the supernatant was discarded, the pellet was dried in vacuum and dissolved in 50 µl Milli-Q water.

2.3 CAPS (Cleaved Amplified Polymorphic Sequence) analysis of the hch fragment

113 *P. anserina* isolates (numbered from Wa01 till Wa113) were investigated by PCR and restriction analysis of PCR fragments to find out whether there was polymorphism in the *hch* locus. To detect possible length polymorphisms in the amplified fragments (see Figure 2), we used EcoR I (Invitrogen), Hpa II (New England Biolabs) and Rsa I (New England Biolabs) enzymes for restriction analysis. Cutting positions of enzymes and size of expected fragments are shown in Table 2. and Figure 2. If polymorphism in *hch* of *P. anserina* is present as long insertions/deletions similarly to those found in the *N. crassa het-c* (Saupe and Glass 1997), size difference between fragments should be visible on gel. To support the results of enzymatic digestions, whether length differences between fragments are present or not, we sequenced 11 randomly chosen isolates. Conditions for digestions (temperature, buffer) were set following the manufacturers' instructions. The reactions were performed in 1.5 ml Eppendorf tubes at 37 ºC, in an Eppendorf-mixer (Eppendorf Thermomixer comfort, Eppendorf AG. Germany). Restriction fragments were separated on 2% agarose gel in 0.5x TBE (tris-borate-EDTA) buffer. Products were stained with 0.1 µg/ml Ethidium-Bromide (Sigma-Aldrich) and visualized by UV transillumination.

2.4 Primer design and PCR amplification of hch and het-c genes

The sequence of the *P. anserina hch* gene was collected from the GenBank (accession number AF169793). The *hch* primers (Table 1) were designed to cover the region homologous to the known polymorphic region of *N. crassa het-c* (Figure 1). Ten random *P. anserina* wild-type isolates were sequenced for this region. Sequences of *P. anserina het-c* alleles were also collected from the GenBank and had the following accession numbers: *het-c1* - L36207, *het-c2* - U05236, *het-c3* - L36209, *het-c4* - L36210. Three primer pairs (Table 1) were designed to produce overlapping PCR fragments that cover both the coding sequence of the *het-c* gene and flanking regions (Figure 2). Randomly 15 *P. anserina* strains were selected for analysis of the complete *het-c* gene. PCR amplifications were performed in 25 µl volume, with concentrations of the following components per reaction: 0,35 units SuperTaq DNA Polymerase (HT Biotechnology, 5 units/µl), 200µM dNTPs (Promega), 0.2 µM primers (MWG Biotech), and 2.5 µl PCR buffer (10 mM Tris-HCl pH=9,0, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1 % (w/v) Triton X-100, and 0.01 % (w/v) Gelatin). All amplifications were done with a Biorad iCycler thermal cycler. For the *het-c* fragments an initial temperature of 95°C for 5 min, 35 cycles at 95°C for 30 s, 46°C (for *het-c* II) or 57°C (for *het-c* I and *het-c* III) for 1 min, and 72°C for 1 min were used with a final step at

72°C for 5 min. For the *hch* fragment the amplifications were done with an initial temperature of 95°C for 5 min, 35 cycles at 95°C for 30 s, 46°C for 1 min, and 72°C for 2 min and a final step at 72°C for 5 min.
PCR products were separated on a 1% agarose gel in 0.5x TBE (tris-borate-EDTA) buffer. The fragments were stained with 0.1 µg/ml Ethidium-Bromide (Sigma-Aldrich) and visualized by UV transillumination.

Table 1. Primers to amplify *hch* and *het-c* genes from *P. anserina* and length of PCR products

Primer pairs	Sequence (5' → 3')	Melting temperature (°C)	Length of PCR product (bp)
hch forward *hch* reverse	TTT GGT GCT GGG AAC ATT TCG GAT TGG GTG AAG TGG	46	913
het-c I forward *het-c* I reverse	AAA ACG AGA CGA ATC AGC TGG GAC ATC AAC GAA GGA	57	585
het-c II forward *het-c* II reverse	TAA GAC CTT CAA GAA GTC CTT CG GCT GTG GTG TCG CAT GAG	46	513
het-c III forward *het-c* III reverse	CCT TCC GCG GGT CTT ACA G GGT GGG GGA TGG TAC CAT G	57	455

Table 2. Restriction enzymes used in *hch* CAPS analysis

Restriction enzyme	cutting positions in the amplified fragment	Length of restriction fragments
EcoR I	316	316, 597
Hpa II	785	128, 785
Rsa I	62, 468, 582	62, 406, 114, 331

2.5 Sequencing
Sequencing of *hch* and *het-c* fragments was done with DETT 2 chemicals (Amerham): 4 µl in 10 µl reaction mixture with 200 ng DNA and 5 ng primers. The sequencing reaction was performed in a 16 capillaries ABI PRISM Genetic Analyzer (Applied Biosystems).

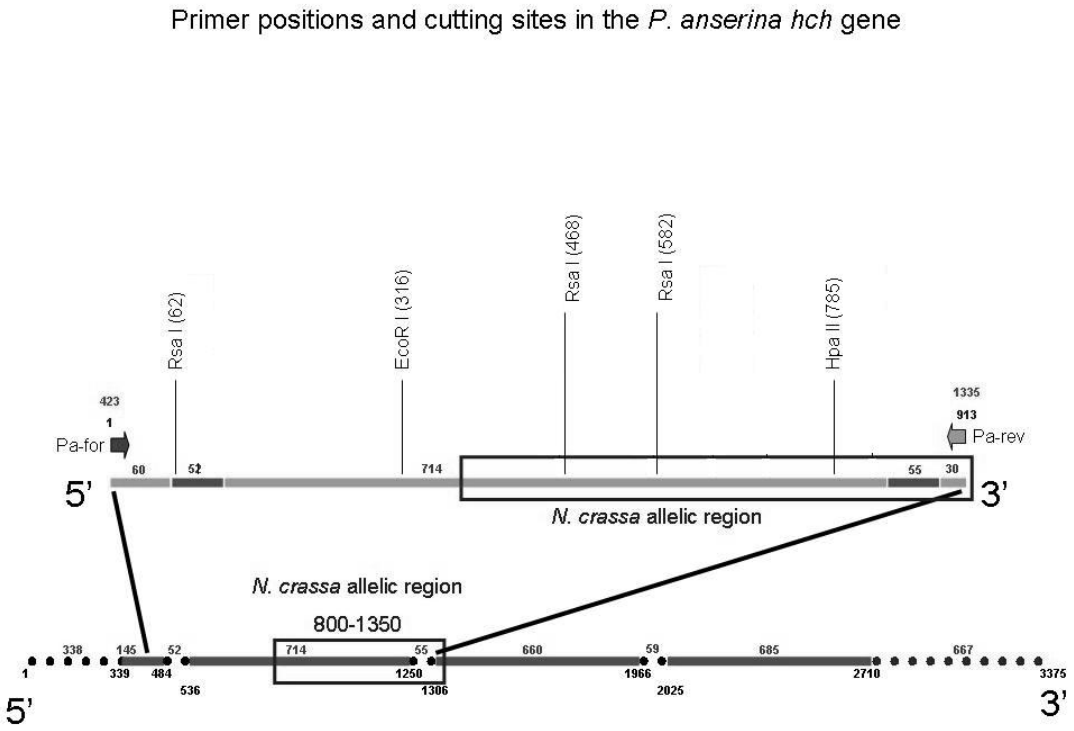


Figure 1. Primer positions and restriction sites for CAPS analysis of the *hch* gene

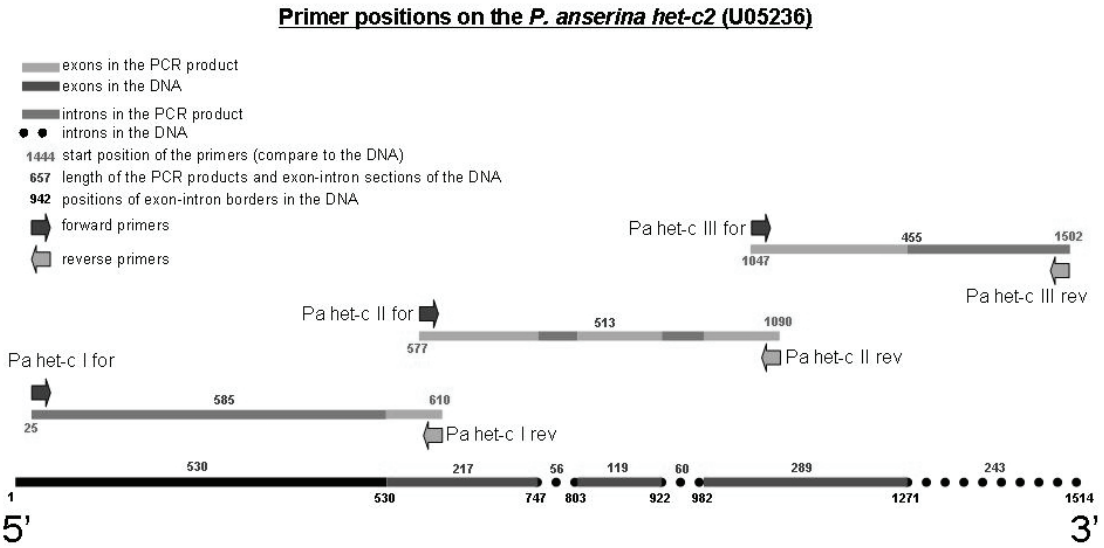


Figure 2. Primer positions for amplifying *het-c* in *P. anserina* isolates

3. Results

3.1 Absence of polymorphism in the hch gene

All the 113 *P. anserina* isolates of our Wageningen collection of wild-type strains were investigated with restriction analysis. Furthermore, 11 of them (with the following strain numbers: Wa08, Wa25, Wa27, Wa61, Wa63, Wa87, Wa92, Wa100, Wa106, Wa116 and Wa118) were sequenced to find out whether there is polymorphism in the *hch* locus in this population. Neither the CAPS analysis (Figure 3) nor the sequencing (not shown) resulted in any differences between the *hch* fragments of different isolates.

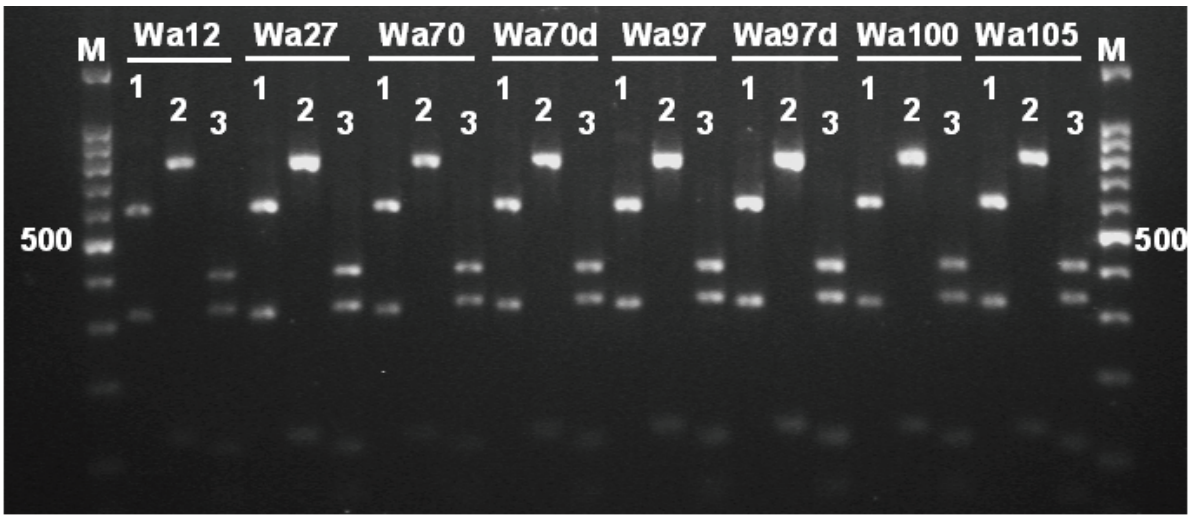


Figure 3. Restriction pattern of *P. anserina hch* fragments on 2 % gel. Legend: M: 100 bp ladder (5 µl); 500: 500 bp sign; numbers 1, 2 and 3: digestion product of EcoR I, Hpa II and Rsa I, respectively. The names of the isolates are indicated as Wageningen + number.

3.2 New polymorphisms in the het-c gene

Randomly 15 strains from the collection of 113 isolates were selected for analysis of the total *het-c* gene and flanking regions. The PCR fragments were sequenced and translated to proteins, which were aligned by ClustalX and a neighbour-joining tree was constructed from this alignment (Figure 4). Three isolates (Wa01, Wa24 and Wa61) had an identical gene sequence to that of the known HET-C3 protein, two isolates (Wa91 and Wa106) proved identical to the HET-C1 and isolate Wa08 showed a strong similarity to the HET-C2 protein. The described HET-C4 protein was not found in the sequenced isolates.

Nine of our natural isolates did not fit with any of the known *het-c* alleles, but formed separate clusters. Wa43, Wa63, Wa100, Wa108 and Wa116 seem to be transitions between the HET-C1 and HET-C4 groups, whereas Wa27, Wa87, W92 and Wa118 are
closer related to the HET-C2.

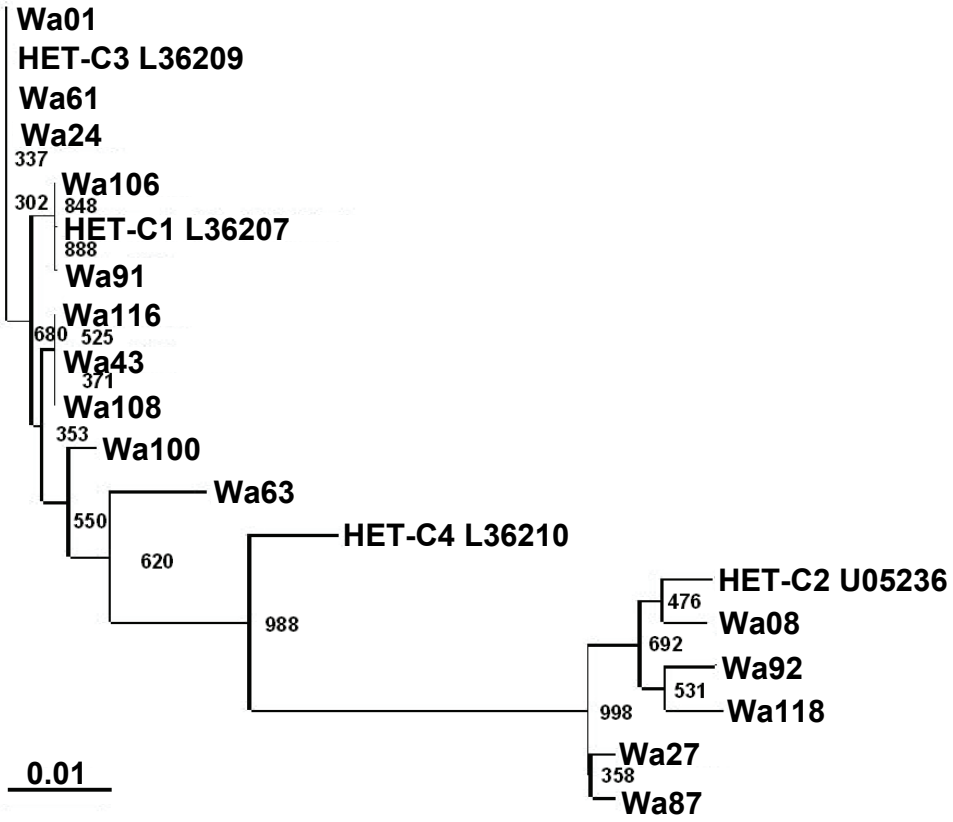


Figure 4. Neighbour-joining tree of HET-C proteins from *P. anserina*. Strains on the tree: Wa: isolates from Wageningen, HET-C1, -C2, -C3 and –C4 are *P. anserina* reference sequences from the GenBank. Sequence alignment and bootstrapping was done with ClustalX.

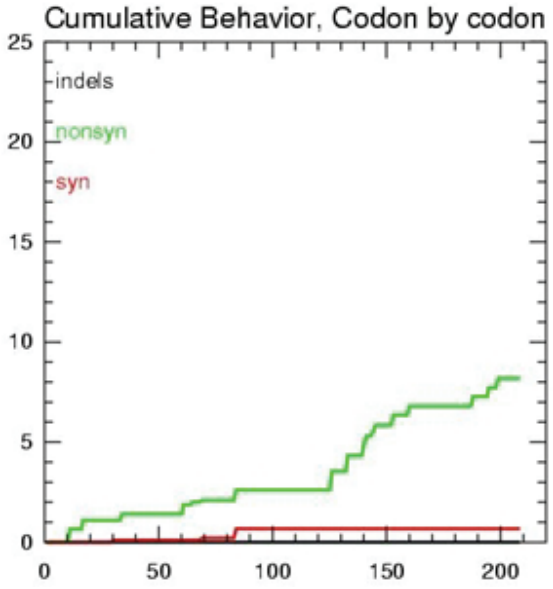


Figure 5. Ratio of non-synonymous and synonymous mutations between the 15 HET-C alleles. Courtesy of Mathieu Paoletti.

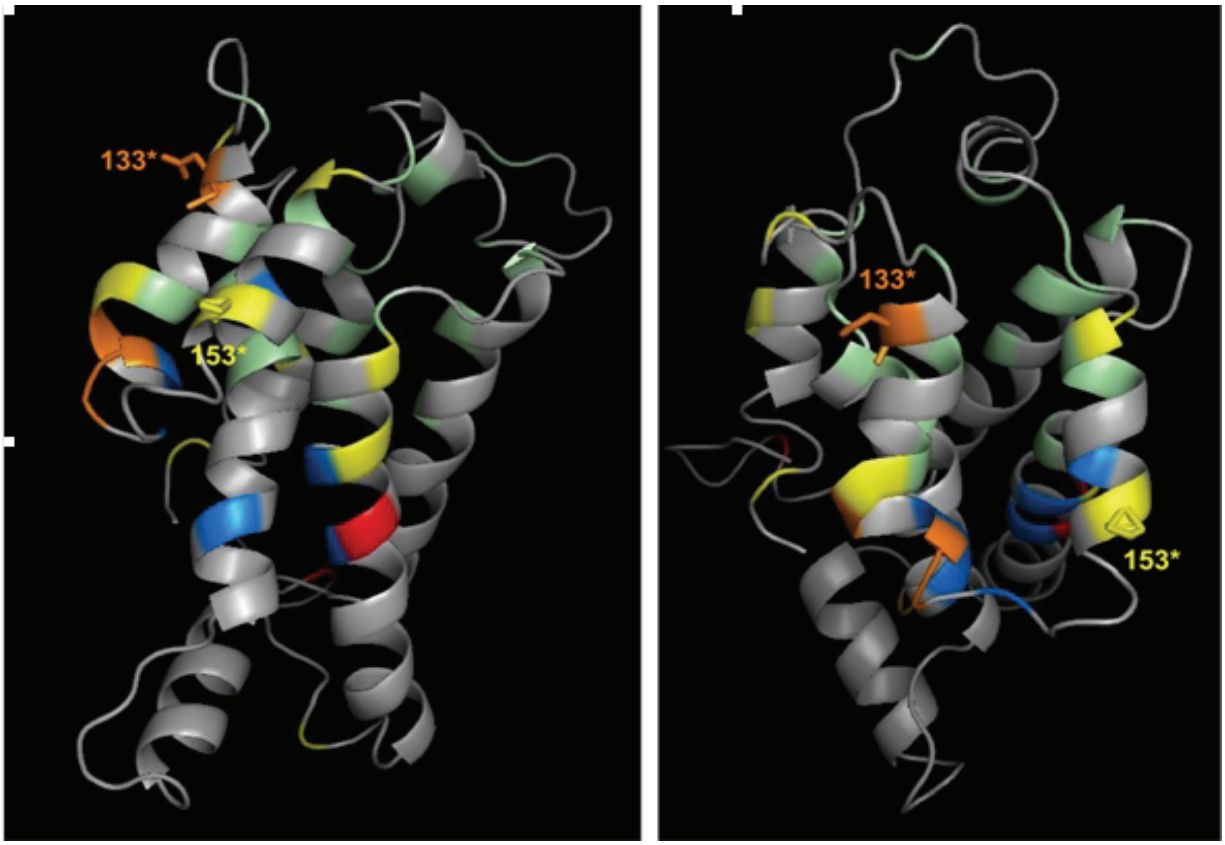


Figure 6. Polymorphisms mapped onto the 3-dimensional structure model of HET-C. Two different views of the structure are shown. All new polymorphisms mapped at external positions and are distinct from the positions predicted to be involved in lipid and sugar binding ensuring the glycolipid transfer activity. Legend of colours: blue: sugar binding site; green: lipid binding site; yellow: known polymorphism in *het-c1* to *het-c4*; orange: position where polymorphism was already known from *het-c1* and *het-c-4* and we found additional variation like for instance at position 133 (Wa63); red: new polymorphism described in this study. Note that all the polymorphic sites mapped at external positions on the protein. The 3D model was constructed by Mathieu Paoletti.

Analysis of the DNA sequences showed that most of the differences between the newly discovered alleles are due to non-synonymous substitutions and much less to synonymous substitutions (Figure 5). The 3D model of the HET-C protein shows that the newly found polymorphic sites are located at the external side of the protein (Figure 6), where protein-protein interactions occur in both the non-allelic HET-C/HET-D and HET-C/HET-E reactions. Amino acid position 133 was already known as a polymorphic site comparing the *het-c1* and *het-c4* alleles, but in the Wa63 sample we found a new variation. This position 133 is known to be involved in recognition specificity. The new polymorphic sites are distinct from the positions predicted to be involved in lipid and sugar binding ensuring the glycolipid transfer activity of the *het-c* gene.

4. Discussion

Heterokaryon incompatibility is an extensive occurring phenomenon in filamentous fungi. Heterokaryon incompatibility is regulated by so-called het-genes that can function by either allelic (with two or more alleles at one locus) or by non-allelic (with two distinct loci involved) systems. In different filamentous fungi various *het*-genes have been identified involved in heterokaryon incompatibility in these species. *P. anserina* is a model system for both allelic and non-allelic incompatibility and at least nine *het*-loci have been identified. Here we investigated a natural population of *P. anserina* strains for polymorphisms at two putative *het*-genes. The allelic *hch* locus is known in many *Sordariales* to have at least three different alleles, but in our population of 113 *P. anserina* strains we found no evidence for polymorphism by either CAPS or sequence analyses in our collection of natural isolates. The non-allelic *het-c* gene interacts with the *het-d* and *het-e* genes and of all three genes different alleles are known. In our wild-type collection, we found five isolates, which harbour *het-c* genes with slight differences compared to the already known alleles, and may be new alleles. All of these new *het-c* forms are possibly involved in new incompatible combinations with *het-d* and *het-e*.

The *P. anserina hch* locus is the homolog of the *N. crassa het-c* locus (Saupe *et al.* 2000). In *N. crassa* and all other heterothallic *Sordariales* so far analysed, three alleles of the *hch/het-c* were found which show trans-species polymorphism and differ in a short indel motif of 56 amino acids (Wu *et al.* 1998). In a previous study to find out whether the *P. anserina hch* is allelic or not, 11 strains (mostly laboratory strains, isolated in France) were sequenced and analysed in the region homologus to the region responsible for polymorphism in the *N. crassa het-c*. Only one contained a silent polymorphism in a nucleotide (Saupe *et al.* 2000). We investigated the same region in a much larger set of 113 natural isolates collected in Wageningen, the Netherlands, to see if we could find any polymorphism. We focused on the region of *hch* homologous to the variable region of *N. crassa het-c* gene, and amplified a 913 bp long fragment by PCR from the isolates. This fragment was digested by three different enzymes to visualize possible indel motifs by CAPS analysis. Furthermore ten of these fragments were sequenced. We did not find any polymorphism in our isolates, which correlates with the results in the literature (Saupe *et al.* 2000). Although the *P. anserina hch* does not show polymorphism as does the *N. crassa het-c* locus, it was found to be able to function as *het*-gene when *N. crassa* alleles were introduced (Saupe *et al.* 2000).
The *hch/het-c* seems to be highly conserved amongst the filamentous fungi. Similar to the *P. anserina* population no polymorphism of *hch/het-c* was found in the region responsible for polymorphism in the *N. crassa het-c* in a population of wild-type *Aspergillus niger* strains (Pál *et al.*, unpublished results; Chapter 3). However, like in *P. anserina*, incompatibility could be triggered in *A. niger* by introduction of alleles of the *N. crassa het-c*. Interestingly, some polymorphism was found in the *A. niger het-c*, but this can be the reason of polymerase slippage.

It is not known what the role is of the *hch/het-c* gene in the fungal cell or if the only reason of its existence is to generate a vegetative barrier between different isolates.

The *P. anserina het-c* gene is one of the three incompatibility loci involved in the non-allelic *het-c*/*het-d* and *het-c*/*het-e* system (Saupe *et al.* 1994). So far four alleles were known of the *P. anserina het-c*. Three of these known alleles were also found in our investigaton and in addition we found presumably five new alleles. However, we only examined 15 randomly chosen strains, which means only 13% of our collection, therefore many more new alleles may be present in our collection and in nature. Nevertheless, this is only a theoretical conclusion, to see whether these five (and maybe more) new variants are really functional alleles or not, a functional analysis should be performed. Estimation of the maximum number of alleles in the Wageningen *P. anserina* population with a hyperbolic model based on these data yields a predicted number of 24 alleles for the Wageningen population of *P. anserina* strains (with a confidence interval between 13 and 35).
The differences between these new and the already known alleles are mostly non-synonymous mutations (differences in nucleotides that alter the amino acid encoded by the triplet). This phenomenon was seen before in the *het-c* locus of *P. anserina* (Saupe *et al.* 1995). Certain *het-c* alleles only interact with given *het-d* and *het-e* alleles (Saupe *et al.* 1994) and these genes seem to be under positive Darwinian selection resulting in more alleles (Saupe 2000). The *hch*/*het-c* allele of *N. crassa,* however, with its three known alleles is believed to be maintained by balancing selection (Wu *et al.* 1998).
In our experiments we found several presumably new functional *het-c* alleles, for two reasons. At first, mutations in the coding part of the gene were mostly non-synonymous. The ratio between silent and non-synonymous mutations was lower in the coding region, than expected in loci under selective pressure to maintain protein function (Nei 1987). Besides, some of the new polymorphisms were in polymorphic sites known to be involved in recognition. Similar situations were described for the *het-s* locus of *P. anserina* (Deleu *et al.* 1993) and the *ref(2)P* gene in *Drosophila*. In the latter case, comparison of four alleles revealed one silent and seven non-synonymous polymorphisms (Dru *et al.* 1993). In addition to the ratio of synonymous/non-synonymous substitutions, on the 3D model of the HET-C protein the polymorphic positions are located on the right place to be involved interactions with the HET-D and HET-E proteins (Paoletti pers. comm., unpublished results).
These findings suggest a very dynamic and rapidly evolving non-allelic system, in which concerted evolution and positive selection act towards the generation of new alleles in all three genes involved: *het-c*, *het-d* and *het-e*. Similar fast evolving genes were found in numerous organisms, like immunoglobulin genes in mammals (Maizels 2005) or certain pathogen resistance genes in plants (Mondragon-Palomino and Gaut 2005). Nevertheless, at the same time the *het-c* locus in *P. anserina* must be under maintaining selection to retain the primary function

of the gene, because inactivation of the locus results in serious defects on ascospore production (Saupe *et al.* 1994).

With the obtained results we can speculate on the functionality of the monomorphic *hch* gene and the polymorphic *het-c* gene in the natural population. The *hch* gene does not seem to play an important role in incompatibility in the natural *P. anserina* population, but the fact that the gene is well-conserved in many species and can be functional when different alleles from *N. crassa* are introduced into *P. anserina* suggest that the gene has an essential function within the cell. The polymorphism in the *het-c* gene and the presumed positive selection for new alleles suggests that this gene has an important role in heterokaryon incompatibility in *P. anserina* and potentially also has one or more other functions.

The significance of heterokaryon incompatibility is still an unanswered question. Filamentous fungi often undergo cell fusion between different isolates and form heterokaryons that may give an opportunity for recombination but also opens a gate for infectious genetic elements, like mycoviruses and plasmids. During the incompatibility reaction the shared hyphae, connecting two, genetically unlike individuals, are destroyed in a rapid lytical step, therefore such hazardous genetic elements are restricted in horizontal transmission. In this meaning vegetative incompatibility is a *bona fide* self/non-self recognition system (Saupe 2000).

The other explanation is that heterokaryon incompatibility is a sort of evolutionary accident. There are many loci in which selectively neutral polymorphisms exist. When some of these are united in the same cytoplasm, they may be able to trigger cell death. This explanation is called the allorecognition hypothesis (Saupe 2000). The fact that we found several different presumably positively selected *het-c* alleles within a single natural population of *P. anserina* is in support of the former explanation that heterokaryon incompatibility is an adaptive fungal nonself recognition system. Whether, the presumed new *het-c* alleles are indeed functional in interaction to the various known *het-E* and *het-D* alleles is still unclear. Furthermore, analysis of the allelic variation of these interacting genes in this population is of great interest and requires further study.

Acknowledgments

We are grateful to Marianne Boone for her assistance with the molecular work at the Laboratory of Genetics, Wageningen University. We thank Sven J. Saupe and Mathieu Paoletti from the Université Bordeaux for their cooperation and useful suggestions. We thank Arjan de Visser for of the hyperbolic modelling of our data.

References

Deleu C, Clavé C, Bégueret J, 1993. A single amino-acid difference is sufficient to elicite vegetative incompatibility in the fungus *Podospora anserina*. *Genetics* 135, 45-52.

Dru P, Bras F, Dezélée S, Gay P, Petitjean AM, Pierre-Denebourg A, Teninges D, Contamine D, 1993. Unusual variability of the *Drosophila melanogaster ref(2)P* protein which controls the multiplication of sigma rhabdovirus. *Genetics* 133, 943-954.

Espagne E, Balhadère P, Bégueret J, Turcq B, 1997. Reactivity in vegetative incompatibility of the HET-E protein of the fungus *Podospora anserina* is dependent on GTP-binding activity and a WD40 repeated domain. *Molecular and General Genetics* 256, 620-627.

Espagne E, Balhadere P, Penin M-L, Barreau C, Turcq B, 2002. HET-E and HET-D belong to a new subfamily of WD40 proteins involved in vegetative incompatibility specificity in the fungus *Podospora anserina. Genetics* 161, 71–81.

Esser K, 1974. *Podospora anserina*. In: King RC (ed) Handbook of genetics. Plenum Press, New York London, pp 531-551.

Glass NL, Kaneko I, 2003. Fatal attraction: nonself recognition and heterokaryon incompatibility in filamentous fungi. *Eukaryotic Cell* 2, 1-8.

Maas FPMM, 2005. On the ecology and evolution of fungal senescence. PhD Thesis, Wageningen University, The Netherlands.

Maizels N, 2005. Immunoglobulin gene diversification. *Annual Review of Genetics* 39, 23-46.

Mondragon-Palomino M, Gaut BS, 2005. Gene conversion and the evolution of three leucine-rich repeat gene families in *Arabidopsis thaliana*. *Molecular Biology and Evolution* 22, 2444-2456.

Nei M, 1987. Molecular evolutionary genetics. Columbia University Press. New York.

Saupe S, Descamps C, Turcq B, Bégueret J, 1994. Inactivation of the *Podospora anserina* vegetative incompatibility locus *het-c*, whose product resembles a glycolipid transfer protein, drastically impairs ascospore production. *Proceedings of the National Academy of Sciences of the USA* 91, 5927-5931.

Saupe S, Turcq B, Bégueret J, 1995. Sequence diversity and unusual variability at the *het-c* locus involved in vegetative incompatibility in the fungus *Podospora anserina*. *Current Genetics* 27, 466-471.

Saupe SJ, Kuldau GS, Smith ML, Glass NL, 1996. The product of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa* has characteristics of a glycine-rich cell wall orotein. *Genetics* 143, 1589–1600.

Saupe SJ, Glass NL, 1997. Allelic specificity at the *het-c* heterokaryon incompatibility locus of *Neurospora crassa* is determined by a highly variable domain. *Genetics* 146, 1299-1309.

Saupe SJ, 2000. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews* 64, 489-502.

Saupe SJ, Clavé C, Sabourin M, Bégueret J, 2000. Characterization of *hch*, the *Podospora anserina* homolog of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa*. *Current Genetics* 38, 39-47.

Sambrook J, Fritsch EF, Maniatis T, 1989. Molecular cloning – a laboratory manual. 2nd edition. Cold Spring Harbor Laboratory Press

van der Gaag M, Debets AJM, Osiewacz HD, Hoekstra RF, 1998. The dynamics of pAL2-1 homologous linear plasmids in *Podospora anserina*. *Molecular and General Genetics* 258, 521-529.

van der Gaag M, 2005. Genomic conflicts in *Podospora anserina*. PhD Thesis, Wageningen University, The Netherlands.

Wu J, Saupe SJ, Glass NL, 1998. Evidence for balancing selection operating at the *het-c* heterokaryon incompatibility locus in a group of filamentous fungi. *Proceedings of the National Academy of Sciences of the USA* 95, 12398-12403.

CHAPTER 5

Sexual genes in the asexual filamentous fungus *Aspergillus niger* and related black Aspergilli

Authors:

Károly Pál[1, 2], Sándor Kocsubé[2], Anne D. van Diepeningen[1], János Varga[2],

Alfons J.M. Debets[1] and Rolf F. Hoekstra[1]

[1]) Laboratory of Genetics, Plant Sciences, Wageningen University, Wageningen, The Netherlands. [2]) University of Szeged, Faculty of Science, Dept. of Microbiology, Szeged, H-6701 Hungary

Abstract

Sexual reproduction is a general feature in higher eukaryotes, although the genetic machinery to produce sexually derived offspring can be very different. Contrary to the megabase-sized sexual chromosomes of animals, fungal genomes harbour much shorter sequences located at the *mat* locus which governs the mating processes. For heterothallic fungi, the two mating-types are characterized by the presence of either of the two idiomorphs of the *mat*-locus*, mat-1* or the *mat-2*, while homothallic species may contain both *mat* idiomorphs. The black mould *Aspergillus niger* is known as an asexual species. but some related presumed asexual species were recently shown to have a cryptic sexual cycle. The scope of this work is to screen for and compare meiosis related genes in *A. niger* and to compare them to those in related Aspergilli, other filamentous fungi and yeast. Furthermore, we screened a population of 162 wild-type black Aspergilli population for the presence of the *mat* genes. Though we find many meiosis related genes in the *A. niger* genome, we do not find any evidence for (cryptic) sexuality in nature.

Keywords: Aspergillus niger, mating type genes, asexuality, *Aspergillus fumigatus*, *Aspergillus nidulans*, *Podospora anserina*, *Neurospora crassa*, *Saccharomyces cerevisiae*

1. Introduction

In general sexual reproduction is an efficient tool to shuffle two parental genomes to promote the creation and spread of advantageous traits and to purge the genome of deleterious mutations. However, there are still continuing debates about the origin and advantages of sex (Hurst and Peck 1996, Barton and Charlesworth 1998, Cavalier-Smith 2002, De Visser and Elena 2007).
Mating processes in fungi are governed by mating type genes. Most heterothallic filamentous ascomycetes have a mating system with two alleles (called idiomorphs, as they do not share any significant sequence similarity) located in a single locus (Turgeon and Yoder 2000). One idiomorph (MAT-1) contains an ORF encoding a protein with a motif called an alpha box, which is also present in the MATα1 protein of the yeast *Saccharomyces cerevisiae* (Turgeon and Yoder 2000). The other idiomorph (MAT-2) contains a single open reading frame (ORF) encoding a regulatory protein with a DNA-binding domain of the high mobility group (HMG) type. Mating type genes regulate plasmogamy and the production of ascogenous hyphae in ascomycetes. Mating type genes also have other functions in some species, for example they can affect vegetative incompatibility in *Neurospora crassa*. Such mating type genes have been identified in a number of filamentous ascomycetes belonging to fungal groups that are widely separated in evolutionary terms (e.g. in species belonging to the Pyrenomycetes, Discomycetes and Loculoascomycetes). Recently, mating type genes have also been identified in *A. fumigatus*, *A. nidulans* and *A. oryzae* (Pöggeler 2002, Varga 2003, Dyer *et al.* 2003, Paoletti *et al.* 2005).
Knowledge about the reproductive strategy of fungi is not only of fundamental interest but is also crucial for applied research, e.g. disease management (Milgroom and Fry, 1997). Observation of a high genetic diversity in a supposedly asexual fungus might be explained by meiotic exchanges in the near past or the existence of a cryptic sexual state. For example, in different isolates of the opportunistic human pathogen *Aspergillus fumigatus* (Latgé 1999), which was known as an asexual species, both mating-type genes and meiosis-related genes were identified (Pöggeler 2002, Varga 2003, Dyer and Paoletti 2005, Paoletti *et al.* 2005).
The ascomycete *Aspergillus* genus includes many different fungi which are mostly saprobic, but human and plant pathogens are also common. So far 114 out of 186 *Aspergillus* species were found to produce only asexual mitotic conidiospores. The other 72 strains produce both asexual conidia and sexual meiotic ascospores, while two species are known to propagate only via sexual spores. Phylogenetic analysis of 15 species revealed that sexual and asexual Aspergilli are closely related (Geiser *et al.* 1996).
The black mould *A. niger* is an industrially important fungus and occasionally an opportunistic human pathogen. *Aspergillus* section *Nigri* consists of several presumably asexual species next to *A. niger*, including eg. *A. tubingensis*, *A. brasiliensis, A. japonicus, A. carbonarius* and several others (Samson *et al.* 2004). These species have minor morphological differences, but can be distinguished by mitochondrial DNA analysis (Varga *et al.* 1993, 1994, Kevei *et al.* 1996, Hamari *et al.* 1997), restriction analysis of the ITS (intergenic spacer) region (Accensi *et al.*

1999), and by sequence analysis (Samson *et al.* 2004). In general, asexual fungi can recombine via the parasexual cycle, but this seems limited in the natural isolates of the black Aspergilli as they have a high degree of heterokaryon incompatibility (van Diepeningen *et al.* 1997). Also looking at recombination between molecular phylogenies of different nuclear genes no evidence could be found for recombination between them (van Diepeningen 1999).
In this study, we examined eight, recently sequenced ascomycete fungal genomes for genes involved in meiotic processes. We focused mainly on the two sequenced *A. niger* genomes, and later broadened our examinations to four closely related sexual and asexual Aspergilli and to two sexual species from the *Sordariales* order. Our aim was to compare the genetic architecture of the mating related genes in sexual and asexual filamentous fungi with those of the yeast *Saccharomyces cerevisiae*, a model organism for sexual processes. Furthermore, we developed and tested (degenerate) primers for the *mat-1* and *mat-2* genes and tested more than 160 natural black *Aspergillus* isolates for the presence or absence of mating genes in their genomes.

2. Materials and methods

2.1 Genes involved in mating

In *S. cerevisiae* three different groups of genes were identified that are involved in mating. These groups are the meiotic proteins that are involved in meiosis itself, the pheromone processing proteins involved in signaling to the other mating partner and pheromone response proteins involved in the response to mating pheromones. A list of these *S. cerevisiae* genes based on the article of Woo *et al.* (2006) is given in Table 1.

2.2 Identification of MAT homolog sequences in the A. niger databases

We searched two *A. niger* databases: the genome of the culture collection strain CBS513.88 from DSM (prepared by shotgun sequencing, 7.5 times coverage; http://www.dsm.com/en_US/html/dfs/genomics_aniger.htm) and the genome of the culture collection strain ATCC1015 from DOE Joint Genome Institute (assembly release version 1.0 of whole genome shotgun reads was constructed with the JGI assembler, Jazz, using paired end sequencing reads at a coverage of 8.9X; http://www.jgi-psf.org/cgi-bin/runAlignment?db=Aspni1&advanced=1). The strains are heterokaryon incompatible with one another and as all *A. niger strains* presumably asexual.
For blastp searches we used the proteins of *S. cerevisiae* as listed above. For validation of the identified *A. niger* sequences, a bi-directional best hit analysis was performed, using the polypeptide sequence of the identified *A. niger* ORFs as a query for a blastp search at the *S. cerevisiae* and GenBank database (http://ncbi.nih.gov/BLAST; Altschul *et al.* 1990). As criterion for homologs we used an accepted E-value of $<e^{-10}$.

2.3 Identification of mating gene homologs in other filamentous fungi

In turn the *S. cerevisiae* genes were used to search the genomes of *A. fumigatus* Af 293, *A. nidulans* FGSC A4, *A. oryzae* RIB40, and *A. terreus* NIH2624, *P. anserina* S and *N. crassa* OR74A for homologs. Again bi-directional best hit analyses were performed with as criterion for homologs an accepted E-value of $<e^{-10}$.

The *A. fumigatus* preliminary sequence data was obtained from The Institute for Genomic Research website at http://www.tigr.org. *A. nidulans*, *N. crassa* (release 7) and *A. terreus* sequence data were from the *Aspergillus nidulans*, *Neurospora crassa* and *Aspergillus terreus* Sequencing Project, Broad Institute of MIT and Harvard (http://www.broad.mit.edu). *A. oryzae* sequences were available on the server of National Institute of Technology and Evaluation (NITE). The *P. anserina* genome was published by the Institut de Génétique et Microbiologie (Université de Paris-Sud XI / CNRS).

2.4 Comparison of homologs of different origin

Homologous polypeptides were aligned with ClustalX (Thompson *et al.* 1997) and edited with Boxshade 3.21 (written by Kay Hofmann and Michael D. Baron and available at http://www.ch.embnet.org/software/BOX_form.html). The degree of identity of proteins was described with blasting the proteins against each other by BLOSUM62 matrix (Tatusova and Madden 1999, at http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi,).

2.5 Strains and culture conditions

In total 161 black *Aspergillus* natural isolates and one *A. niger* culture collection strain were tested for the presence of the *mat-1* and/or *mat-2* mating type genes. 83 of these strains could be classified based on their mitochondrial or ITS haplotypes as *A. niger*, 65 strains as *A. tubingensis*, 9 as *A. japonicus,* one as *A. carbonarius* and one as *A. brasiliensis* (Hamari *et al.* 1997, Kevei *et al.* 1996, Varga *et al.* 1993, 1994). 13 strains were not characterised.

All strains were grown on liquid or solid complete medium (CM) (Pontecorvo *et al.* 1953) supplemented with 1 mg/l $ZnSO_4$, $FeSO_4$, $MnCl_2$ and $CuSO_4$. All incubations took place at 30ºC.

2.6 DNA isolation

A. mycelial mat was grown overnight on liquid CM and 0.1 g was taken and frozen with liquid nitrogen in a 1.5 ml Eppendorf tube. After freezing, the mycelium was grinded with sterile sand. Total nucleic acids were isolated with a phenol-chloroform extraction: a 2 ml Eppendorf tube was filled for ¼ with mycelial powder and 600 µl extraction buffer (0.5 M NaCl, 10 mM Tris-HCl, 10 mM Na_2EDTA, 1% SDS, pH 7.5) was added. The mixture was incubated for 1 hour at 65 ºC and vortexed every 15 minutes. Nucleic acids were then extracted once with 500 µl water saturated phenol (vortex and incubation for 10 minutes on room temperature) and twice with 500 µl and 300 µl SEVAG (24:1 chloroform:isoamyl alcohol), respectively. Nucleic acids were precipitated with 0.6x volume iso-propanol, then chilled in –70 ºC fridge for 30 min. The final wash was done with 70% ethanol out of –20 ºC, and the nucleic acids were resuspended in 50 µl demiwater or TE buffer (10 mM Tris/ 10 mM EDTA, pH 7.5).

2.7 PCR amplification

PCR primers for amplification of *mat-1* gene were designed on the *mat-1* sequence from the *A. niger* CBS513.88 strain. *Mat-2* primers were developed on *A. nidulans* and *A. fumigatus mat-2* sequences, therefore these contain degenerated bases (Table 2).

PCR amplifications on natural black *Aspergillus* strains to amplify *mat-1* were performed in 25 µl volume, with concentrations of the following per reaction: 0,35 units SuperTaq DNA Polymerase (HT Biotechnology, 5 units/µl), 200µM dNTPs (Promega), 0.2 µM primers (MWG Biotech), 2.5 µl PCR buffer (10 mM Tris-HCl pH=9,0, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1 % (w/v) Triton X-100, 0.01 % (w/v) Gelatin). PCR program was performed with an initial temperature of 95 °C for 1 min, 35 cycles at 95 °C for 1 min, 56 °C for 1 min, and 72 °C for 2 min were carried out with a final step at 72 °C for 5 min. PCR amplifications of *mat-2* were performed in 50 µl volume. Final concentration of the ingredients is the following per reaction: 1 µl JumpStart Taq DNA Polymerase (Sigma), 5 µl PCR buffer (Sigma, 10 mM Tris-HCl pH=8.3 at 25 °C, 50 mM KCl, 1.5 mM Mg Cl_2, 0.01 % (w/v) gelatin), 200µM dNTPs (Sigma), 100 pmol primers (Biological Research Centre, Szeged). PCR program was performed with an initial temperature of 95 °C for 4 minutes, 35 cycles at 95 °C for 30 seconds, 52 °C for 40 seconds and 72 °C for 1 min were carried out with a final step at 72 °C for 2 minutes.

Amplifications were done using a Biorad iCycler thermal cycler (Bio-Rad Laboratories, Veenendaal, The Netherlands).

Products of PCR reactions were then separated in a 1% agarose gel in 0.5x TBE (tris-borate-EDTA) buffer. Products were stained with 0.1 µg/ml Ethidium-Bromide (Sigma-Aldrich) and visualized by UV trans-illumination.

2.8 Sequencing

Sequencing of *hch* and *het-c* fragments was done with DETT 2 chemicals (Amerham): 4 µl in 10 µl reaction mixture with 200 ng DNA and 5 ng primers. The sequencing reaction was performed in a 16 capillaries ABI PRISM Genetic Analyzer (Applied Biosystems).

2.9 Statistics

The presence (and absence if no genes were detected with our PCR anlysis) of both mating type genes *mat-1* and *mat-2* in the different tested black *Aspergillus* species was compared with Chi-square tests.

Table 2. PCR primers for the *mat-1* and *mat-2* genes based on the sequences in *A. niger CBS513.88*. 'I' means special base, able to pair with all the four regular bases.

Primer pairs	Sequence (5' → 3')	Melting temperature (°C)	Length of PCR product (bp)
*mat-1*var forward *mat-1*var reverse	ATG CCC CCC CAG CAA CTC GAA AGC CAT GGT GTA CTG CTG CTG	56	542
mat-2 forward *mat-2* reverse	AA(AG) (AG)TI CCI (AC)GI CCI CCI AA(CT) GC TTN C(GT)I GGI GT(AG) TAI TG(AG) TA(AG) TC(AGCT) GG	52	~300*

* the primers were designed on different fungi, we could not tell the exact size of the expected product. We isolated the strongest band from the gel, which was sequenced and showed similarity to the *A. fumigatus* MAT-2.

Table 1. Mating related proteins in *S. cerevisiae* and their functions. The genes fall into three functional classes: the meiotic proteins, the pheromone processing proteins and the pheromone response proteins (Woo *et al.* 2006)

	Name and ID of protein	Function
Meiotic proteins	DMC1 (YER179W)	meiosis-specific protein, repair of double-strand breaks
	HOP1 (YIL072W)	meiosis-specific DNA binding protein
	HOP2 (YGL033W)	meiosis-specific protein, prevents synapsis between nonhomologous chromosomes
	MLH1 (YMR167W)	required for mismatch repair in mitosis and meiosis
	MLH2 (YLR035C)	required for DNA mismatch repair in mitosis and meiosis
	MLH3 (YPL164C)	involved in DNA mismatch repair
	MND1 (YGL183C)	required for recombination and meiotic nuclear division
	MRE11 (YMR224C)	subunit of a complex that functions in repair of DNA double-strand breaks and in telomere stability
	MSH2 (YOL090W)	forms heterodimers with Msh3p and Msh6p that bind to DNA mismatches to initiate the mismatch repair process
	MSH4 (YFL003C)	involved in meiotic recombination
	MSH5 (YDL154W)	forms a dimer with Msh4p that facilitates crossovers between homologs during meiosis
	MSH6 (YDR097C)	required for mismatch repair in mitosis and meiosis
	PMS1 (YNL082W)	ATP-binding protein required for mismatch repair in mitosis and meiosis
	RAD50 (YNL250W)	subunit of MRX complex, involved in processing double-strand DNA breaks in vegetative cells
	RAD51 (YER095W)	DNA strand exchange protein
	RAD52 (YML032C)	stimulates DNA strand exchange
	SPO11 (YHL022C)	meiosis-specific protein, initiates meiotic recombination

Table 1. (Continued) Mating related proteins in *S. cerevisiae* and their functions. The genes fall into three functional classes: the meiotic proteins, the pheromone processing proteins and the pheromone response proteins (Woo *et al.* 2006)

	Name and ID of protein	Function
Pher. proc. proteins	KEX1 (YGL203C)	protease involved in the processing of killer toxin and alpha factor precursor
	KEX2 (YNL238W)	subtilisin-like protease (proprotein convertase)
	RAM1 (YDL090C)	beta subunit of the CAAX farnesyltransferase (FTase)
	RAM2 (YKL019W)	alpha subunit of both the farnesyltransferase and type I geranylgeranyltransferase
	RCE1 (YMR274C)	type II CAAX prenyl protease involved in the proteolysis and maturation of Ras and the a-factor mating pheromone
	STE6 (YKL209C)	ATP-binding cassette (ABC) transporter required for the export of a-factor
	STE13 (YOR219C)	dipeptidyl aminopeptidase, required for maturation of alpha factor
	STE14 (YDR410C)	farnesyl cysteine-carboxyl methyltransferase
	STE23 (YLR389C)	metalloprotease involved in N-terminal processing of pro-a-factor to the mature form
	STE24 (YJR117W)	zinc metalloprotease that functions in two steps of a-factor maturation
Pheromone response proteins	BEM1 (YBR200W)	protein containing SH3-domains, involved in establishing cell polarity and morphogenesis
	CDC24 (YAL041W)	guanine nucleotide exchange factor, required for polarity establishment and maintenance
	CDC42 (YLR229C)	small rho-like GTPase, essential for establishment and maintenance of cell polarity
	DIG1 (YPL049C)	regulatory protein of unknown function
	DIG2 (YDR480W)	regulatory protein of unknown function
	FAR1 (YJL157C)	cyclin-dependent kinase inhibitor that mediates cell cycle arrest in response to pheromone
	FUS3 (YBL016W)	mitogen-activated protein kinase involved in mating pheromone response
	GPA2 (YER020W)	nucleotide binding alpha subunit of the heterotrimeric G protein that interacts with the receptor Gpr1p
	KSS1 (YGR040W)	mitogen-activated protein kinase (MAPK) involved in signal transduction pathways that control filamentous growth and pheromone response
	STE2 (YFL026W)	receptor for alpha-factor pheromone
	STE3 (YKL178C)	receptor for a factor receptor
	STE4 (YOR212W)	G protein beta subunit, forms a dimer with Ste18p to activate the mating signaling pathway
	STE5 (YDR103W)	scaffold protein that, in response to pheromone, shuttles from the nucleus to the plasma membrane
	STE7 (YDL159W)	signal transducing MAP kinase kinase involved in pheromone response
	STE11 (YLR362W)	signal transducing MEK kinase involved in pheromone response
	STE12 (YHR084W)	transcription factor, activates genes involved in mating or pseudohyphal/invasive growth pathways
	STE18 (YJR086W)	G protein gamma subunit, forms a dimer with Ste4p to activate the mating signaling pathway
	STE20 (YHL007C)	signal transducing kinase of the PAK (p21-activated kinase) family, involved in pheromone response
	STE50 (YCL032W)	protein involved in mating response

3. Results

3.1 Comparison of yeast and filamentous fungi

Homologs could be found in the genomes of the filamentous fungi for the majority of *S. cerevisiae* mating related genes (Table 3). However, 9 out of 46 (20%) of the *S. cerevisiae* mating related genes could not be detected in any of the filamentous fungi. The number of genes that were missing from the genomes of different *Aspergillus* species was between 11 and 13, while 18 and 16 genes were missing from the *N. crassa* and *P. anserina* genomes, respectively.

Some of the proteins were highly conserved in all genomes examined even of the presumably asexual fungi (with E-values $<e^{-100}$), and may have an additional role next to their function in mating processes (i.e. DNA repair). In other cases the E-values of proteins of filamentous fungi were just above the threshold when we did the blast search with the *S. cerevisiae* proteins, but comparing the homologs in the different filamentous fungi these proteins showed higher similarity. In most of the sequenced genomes of the filamentous fungi only one – either *mat-1* or *mat-2* – of the mating-type genes was found per genome. Only the genome of the homothallic *A. nidulans* contained both mating-type genes.

3.2 Comparing the different filamentous fungi

N. crassa and *P. anserina* lacked more of the *S. cerevisiae* mating related genes (18 and 16 respectively) than did the different *Aspergillus* species (11-13). Especially DMC1, involved in double-strand break repair in *S. cerevisiae*, and STE7, a kinase involved in the pheromone response, did notably have no homologs in the two *Sordariales* genomes. There was no gene present in the *Sordariales* that was not present in the Aspergilli (Table 3).

The number of protein homologs was generally not different between the different orders or between presumed sexual and asexual lines, except for two proteins: the protease KEX1 and the ABC transporter STE6 (Table 4). Both of these genes with a presumably more general function have multiple homologs. KEX1 has 11-14 homologs in the asexual Aspergilli (including *A. fumigatus* that may have a sexual cycle), only 5 in sexual *A. nidulans*, 4 in *N. crassa* and 5 in *P. anserina*. We identified 24 to 45 homologs of STE6 in the Aspergilli, whereas slightly less were found in *N. crassa* and *P. anserina* (19 and 21, respectively).

The SPO11 protein, involved in the initiation of meiotic recombination, had only a homolog in *A. nidulans,* with a rather low homology (Table 4).

Table 3. *S. cerevisiae*, *N. crassa* and *A. nidulans* mating related proteins and whether they have homologs (threshold E-value is of $<e^{-10}$) in different filamentous fungal species.

			Strain								
			A. fumigatus (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC1055)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (Strain S)	*S. cerevisiae* (S288C)
S. cerevisiae **mating related proteins**	Meiotic proteins	DMC1 (YER179W)	•	•	•	•	•	•			•
		HOP1 (YIL072W)	•	•	•	•		•			•
		HOP2 (YGL033W)									•
		MLH1 (YMR167W)	•	•	•	•	•	•	•	•	•
		MLH2 (YLR035C)									•
		MLH3 (YPL164C)	•	•	•	•	•	•		•	•
		MND1 (YGL183C)		•			•				•
		MRE11 (YMR224C)	•	•	•	•	•	•	•	•	•
		MSH2 (YOL090W)	•	•	•	•	•	•	•	•	•
		MSH4 (YFL003C)	•		•		•	•	•		•
		MSH5 (YDL154W)	•		•	•		•	•	•	•
		MSH6 (YDR097C)	•	•		•	•	•	•	•	•
		PMS1 (YNL082W)	•	•	•	•	•	•	•	•	•
		RAD50 (YNL250W)	•	•	•	•	•	•	•	•	•
		RAD51 (YER095W)	•	•	•	•	•	•	•	•	•
		RAD52 (YML032C)	•	•	•	•	•	•	•	•	•
		SPO11 (YHL022C)		•							•
	Pher. p. proteins	KEX1 (YGL203C)	•	•	•	•	•	•	•	•	•
		KEX2 (YNL238W)	•	•	•	•	•	•	•	•	•
		RAM1 (YDL090C)	•	•	•	•	•	•	•	•	•
		RAM2 (YKL019W)	•	•	•	•	•	•	•	•	•
		RCE1 (YMR274C)	•	•	•	•	•	•		•	•
		STE6 (YKL209C)	•	•	•	•	•	•	•	•	•
		STE13 (YOR219C)									•
		STE14 (YDR410C)	•	•	•	•	•	•	•	•	•
		STE23 (YLR389C)	•	•	•	•	•	•	•	•	•
		STE24 (YJR117W)	•	•	•	•	•	•	•	•	•
	Pheromone response pathway proteins	BEM1 (YBR200W)	•	•	•	•	•	•	•	•	•
		CDC24 (YAL041W)	•	•	•	•	•	•	•	•	•
		CDC42 (YLR229C)	•	•	•	•	•	•	•	•	•
		DIG1 (YPL049C)									•
		DIG2 (YDR480W)									•
		FAR1 (YJL157C)									•
		FUS3 (YBL016W)	•	•	•	•	•	•	•	•	•
		GPA2 (YER020W)	•	•	•	•	•	•	•	•	•
		KSS1 (YGR040W)									•
		STE2 (YFL026W)	•	•	•		•	•			•
		STE3 (YKL178C)	•	•	•	•	•	•		•	•
		STE4 (YOR212W)	•	•	•	•	•	•	•	•	•
		STE5 (YDR103W)									•
		STE7 (YDL159W)	•	•	•	•	•	•			•
		STE11 (YLR362W)	•	•	•	•	•	•	•	•	•
		STE12 (YHR084W)	•	•	•	•	•	•	•	•	•
		STE18 (YJR086W)									•
		STE20 (YHL007C)	•	•	•	•	•	•	•	•	•
		STE50 (YCL032W)	•	•	•	•	•	•	•	•	•
N. c		MAT-A1		•	•	•	•	•	•		•
		MAT-A2							•		
		MAT-A3							•		
A[1]		MAT-1	[2]	•		•	•	•	•		•
		MAT-2	•	•						•	

[1] *A. nidulans* proteins; MAT-1 and MAT-2 proteins have identical function as the *N. crassa* MAT-A1 and MAT-A3, respectively
[2] the MAT-1 was found in different isolates of *A. fumigatus* (see references, Paoletti et al. 2005)

Table 4. Homologs of *S. cerevisiae* proteins involved in meiosis in filamentous ascomycetes. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

	Homologs in fungi and ID of homologous protein with E-value							
Protein	*A. fumigatus* (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC1055)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (Strain S)
Meiotic proteins								
DMC1 (YER179W)	3 Afu7g02200 (0.0)	3 AN9092.3 (0.0)	3 An12g00460 (0.0)	?[3] 43615 (0.0)	3 AO090038000596 (7e-99)	3 ATEG_01853.1 (0.0)	0	0
HOP1 (YIL072W)	1 Afu6g13050 (3e-12)	1 AN5516.3 (3e-14)	1 An08g10440 (8e-19)	1 175768 (2e-11)	0	1 ATEG_03335.1 (2e-14)	0	0
HOP2 (YGL033W)	0	0	0	0	0	0	0	0
MLH1 (YMR167W)	3 Afu5g11700 (0.0)	3 AN0126.3 (0.0)	3 An18g03030 (0.0)	3 42780 (0.0)	3 AO090120000288 (0.0)	3 ATEG_02100.1 (0.0)	3 NCU08309.2 (0.0)	3 Pa_6_2940 (0.0)
MLH2 (YLR035C)	0	0	0	0	0	0	0	0
MLH3 (YPL164C)	4 Afu4g06490 (1e-34)	3 AN4365.3 (6e-28)	3 An04g00870 (8e-29)	3 54646 (2e-34)	3 AO090023000933 (1e-30)	3 ATEG_05571.1 (3e-32)	0	3 Pa_3_9800 (4e-29)
MND1 (YGL183C)	0	1 AN1843.3 (6e-12)	0	0	1 AO090003001276 (8e-15)	0	0	0
MRE11 (YMR224C)	1 Afu6g11410 (0.0)	1 AN0556.3 (0.0)	1 An08g07330 (0.0)	1 175727 (0.0)	1 AO090023000471 (0.0)	1 ATEG_05022.1 (0.0)	1 NCU08730.2 (0.0)	1 Pa_6_11830 (0.0)

[3] Exact number of homologs is not known, since only blasting in the non-filtered database resulted hits, which were redundant

Table 4. (Continued) Homologs of *S. cerevisiae* proteins involved in meiosis in filamentous ascomycetes. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

	Homologs in fungi and ID of homologous protein with E-value							
Protein	*A. fumigatus* (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC1055)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (Strain S)
Meiotic proteins								
MSH2 (YOL090W)	6 Afu3g09850 (0.0)	4 AN10621.3 (0.0)	6 An16g03520 (0.0)	5 183201 (0.0)	5 AO090005001261 (0.0)	6 ATEG_04508.1 (0.0)	6 NCU02230.2 (0.0)	7 Pa_5_9490 (0.0)
MSH4 (YFL003C)	5 Afu1g02000 (5e-66)	0	6 An01g06260 (0.0)	0	5 AO090005000895 (5e-98)	6 ATEG_04890.1 (0.0)	6 NCU10895.2 (3e-43)	0
MSH5 (YDL154W)	6 Afu1g11170 (2e-45)	0	6 An08g03470 (7e-73)	5 122070 (1e-43)	0	6 ATEG_00319.1 (0.0)	6 NCU09384.2 (0.0)	6 Pa_1_13960 (2e-38)
MSH6 (YDR097C)	1 Afu4g08300 (0.0)	4 AN1708.3 (0.0)	0	4 50490 (0.0)	5 AO090023000729 (0.0)	6 ATEG_05365.1 (0.0)	6 NCU08135.2 (0.0)	7 Pa_5_3880 (0.0)
PMS1 (YNL082W)	1 Afu2g13410 (0.0)	3 AN6316.3 (0.0)	3 An02g02280 (0.0)	4 36742 (0.0)	3 AO090026000318 (9e-79)	3 ATEG_01317.1 (0.0)	3 NCU08020.2 (0.0)	4 Pa_6_1450 (3e-83)
RAD50 (YNL250W)	7 Afu4g12680 (0.0)	8 AN3619.3 (0.0)	15 An01g08180 (0.0)	?[4] 138700 (0.0)	6 AO090009000296 (0.0)	6 ATEG_03151.1 (0.0)	10 NCU00901.2 (0.0)	6 Pa_1_1880 (0.0)
RAD51 (YER095W)	2 Afu1g10410 (0.0)	2 AN1237.3 (0.0)	3 An08g02350 (0.0)	?[4] 47657 (0.0)	2 AO090038000386 (0.0)	3 ATEG_00230.1 (0.0)	1 NCU02741.2 (0.0)	1 Pa_1_13570 (0.0)
RAD52 (YML032C)	1 Afu4g06970 (4e-53)	1 AN4407.3 (0.0)	1 An04g01290 (4e-47)	1 44271 (3e-32)	1 AO090023000890 (1e-47)	1 ATEG_05536.1 (0.0)	1 NCU04275.2 (0.0)	1 Pa_2_12450 (2e-46)

[4] Exact number of homologs is not known, since only blasting in the non-filtered database resulted hits, which were redundant

Table 4. (Continued) Homologs of *S. cerevisiae* proteins involved in meiosis and pheromone processing in filamentous ascomycetes. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

	Homologs in fungi and ID of homologous protein with E-value							
Protein	*A. fumigatus* (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC1055)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (Strain S)
Meiotic proteins								
SPO11 (YHL022C)	0	1 AN8259.3 (1e-12)	0	0	0	0	0	0
Proteins involved in pheromone processing								
KEX1 (YGL203C)	11 Afu1g08940 (1e-83)	5 AN10184.3 (0.0)	14 An08g00430 (2e-98)	12 208486 (2e-94)	12 AO090005001632 (6e-92)	11 ATEG_08515.1 (0.0)	4 NCU04316.2 (0.0)	5 Pa_2_10030 (7e-84)
KEX2 (YNL238W)	1 Afu4g12970 (0.0)	1 AN3583.3 (0.0)	1 An01g08530 (0.0)	1 55344 (0.0)	1 AO090009000291 (0.0)	1 ATEG_03179.1 (0.0)	1 NCU03219.2 (0.0)	1 Pa_1_1930 (0.0)
RAM1 (YDL090C)	2 Afu4g10330 (4e-47)	2 AN2002.3 (0.0)	3 An04g06620 (2e-59)	2 133160 (3e-47)	2 AO090003001188 (5e-37)	2 ATEG_00696.1 (3e-40)	2 NCU05999.2 (1e-39)	2 Pa_4_7760 (3e-47)
RAM2 (YKL019W)	1 Afu4g07800 (9e-31)	2 AN3867.3 (1e-45)	2 An04g02210 (4e-43)	2 190580 (4e-39)	2 AO090023000791 (1e-41)	1 ATEG_05421.1 (4e-43)	1 NCU03632.2 (3e-39)	
RCE1 (YMR274C)	1 Afu6g04890 (1e-18)	1 AN6528.3 (1e-19)	1 An14g03420 (2e-17)	1 211108 (1e-12)	1 AO090701000050 (5e-20)	1 ATEG_06959.1 (7e-16)	0	

Table 4. (Continued) Homologs of *S. cerevisiae* proteins involved in pheromone processing in filamentous ascomycetes. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

	Homologs in fungi and ID of homologous protein with E-value							
Protein	*A. fumigatus* (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC1055)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (Strain S)
Proteins involved in pheromone processing								
STE6 (YKL209C)	29 Afu4g08800 (0.0)	24 AN2300.3 (0.0)	25 An04g03690 (0.0)	35 190859 (0.0)	45 AO090023000664 (0.0)	25 ATEG_09424.1 (0.0)	19 NCU09975.2 (0.0)	21 Pa_7_7770 (0.0)
STE13 (YOR219C)	0	0	0	0	0	0	0	0
STE14 (YDR410C)	1 Afu2g08420 (1e-43)	1 AN6162.3 (0.0)	1 An12g03660 (5e-44)	1 186640 (3e-48)	1 AO090011000860 (2e-37)	1 ATEG_09679.1 (0.0)	1 NCU00034.2 (0.0)	1 Pa_7_9690 (2e-32)
STE23 (YLR389C)	1 Afu5g02010 (0.0)	1 AN8044.3 (0.0)	2 An16g01860 (0.0)	2 41311 (0.0)	1 AO090003001317 (0.0)	1 ATEG_09820.1 (0.0)	1 NCU00481.2 (0.0)	1 Pa_7_6640 (0.0)
STE24 (YJR117W)	1 Afu4g07590 (0.0)	1 AN11231.3 (0.0)	1 An04g01950 (0.0)	1 50547 (0.0)	1 AO090023000816 (0.0)	1 ATEG_05444.1 (0.0)	1 NCU03637.2 (0.0)	1 Pa_2_9310 (0.0)

Table 4. (Continued) Homologs of *S. cerevisiae* proteins involved in pheromone response in filamentous ascomycetes. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

	Homologs in fungi and ID of homologous protein with E-value							
Protein	*A. fumigatus* (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC1055)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (Strain S)
Proteins involved in pheromone response								
BEM1 (YBR200W)	1 Afu4g04120 (1e-48)	1 AN7030.3 (0.0)	1 An14g00710 (9e-67)	1 53737 (3e-53)	1 AO090206000084 (7e-55)	1 ATEG_10121.1 (0.0)	1 NCU06593.2 (0.0)	1 Pa_6_570 (1e-56)
CDC24 (YAL041W)	1 Afu4g11450 (1e-46)	1 AN5592.3 (3e-38)	2 An04g05150 (4e-41)	1 129585 (1e-44)	1 AO090003001078 (1e-39)	2 ATEG_03936.1 (7e-39)	1 NCU06067.2 (9e-41)	1 Pa_5_10060 (1e-31)
CDC42 (YLR229C)	17 Afu2g05740 (1e-85)	15 AN7487.3 (0.0)	17 An02g14200 (7e-90)	9 52477 (0.0)	14 AO090001000693 (4e-90)	11 ATEG_06763.1 (0.0)	12 NCU02160.2 (0.0)	15 Pa_7_9860 (6e-89)
DIG1 (YPL049C)	0	0	0	0	0	0	0	0
DIG2 (YDR480W)	0	0	0	0	0	0	0	0
FAR1 (YJL157C)	0	0	0	0	0	0	0	0
FUS3 (YBL016W)	63 Afu6g12820 (0.0)	66 AN3719.3 (0.0)	25[5] An08g10670 (0.0)	51 207710 (0.0)	64 AO090003000402 (0.0)	63 ATEG_03316.1 (0.0)	62 NCU02393.2 (0.0)	60 Pa_5_5680 (0.0)
GPA2 (YER020W)	3 Afu1g12930 (0.0)	3 AN1016.3 (0.0)	4 An08g05820 (0.0)	4 55775 (0.0)	4 AO090012000600 (0.0)	3 ATEG_00488.1 (0.0)	3 NCU05206.2 (0.0)	4 Pa_1_23950 (0.0)

[5] Maximum number of hits was limited in 25 hits per blast search.

Table 4. (Continued) Homologs of *S. cerevisiae* proteins involved in pheromone response in filamentous ascomycetes. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

	Homologs in fungi and ID of homologous protein with E-value							
Protein	*A. fumigatus* (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC1055)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (Strain S)
Proteins involved in pheromone response								
KSS1 (YGR040W)	0	0	0	0	0	0	0	0
STE2 (YFL026W)	1 Afu3g14330 (3e-17)	1 AN2520.3 (1e-18)	1 An09g04180 (5e-17)	0	1 AO090701000605 (4e-16)	1 ATEG_03500.1 (5e-15)	0	0
STE3 (YKL178C)	1 Afu5g07880 (7e-38)	1 AN7743.3 (3e-15)	1 An03g03890 (2e-37)	1 44420 (3e-19)	1 AO090701000699 (4e-38)	1 ATEG_08338.1 (8e-36)	0	1 Pa_7_9070 (2e-11)
STE4 (YOR212W)	11 Afu5g12210 (2e-55)	14 AN0081.3 (0.0)	13 An18g02090 (7e-78)	8 54102 (1e-76)	15 AO090120000339 (4e-77)	11 ATEG_02052.1 (0.0)	4 NCU00440.2 (0.0)	14 Pa_7_6570 (8e-66)
STE5 (YDR103W)	0	0	0	0	0	0	0	0
STE7 (YDL159W)	39 Afu3g05900 (2e-57)	50 AN3422.3 (0.0)	25[6] An11g10690 (3e-63)	24 209137 (4e-66)	49 AO090020000060 (4e-64)	49 ATEG_08950.1 (0.0)	0	0
STE11 (YLR362W)	58 Afu5g06420 (4e-91)	61 AN2269.3 (0.0)	25[7] An17g01280 (4e-89)	55 214017 (4e-99)	61 AO090009000610 (1e-79)	59 ATEG_09389.1 (0.0)	57 NCU06182.2 (0.0)	54 Pa_7_8030 (2e-80)
STE12 (YHR084W)	1 Afu5g06190 (2e-60)	1 AN2290.3 (0.0)	1 An17g01580 (2e-60)	1 57283 (3e-66)	1 AO090009000638 (3e-62)	1 ATEG_09411.1 (0.0)	1 NCU00340.2 (0.0)	1 Pa_7_1730 (1e-59)

[6] Maximum number of hits was limited in 25 hits per blast search.

Table 4. (Continued) Homologs of *S. cerevisiae* proteins involved in pheromone response in filamentous ascomycetes. Only proteins with E-value of $<e^{-10}$ are shown. E-values below e^{-100} are shown as 0.0.

	Homologs in fungi and ID of homologous protein with E-value							
Protein	*A. fumigatus* (Af 293)	*A. nidulans* (FGSC A4)	*A. niger* (CBS513.88)	*A. niger* (ATCC1055)	*A. oryzae* (RIB40)	*A. terreus* (NIH 2624)	*N. crassa* (OR74A)	*P. anserina* (Strain S)
Proteins involved in pheromone response								
STE18 (YJR086W)	0	0	0	0	0	0	0	0
STE20 (YHL007C)	60 Afu2g04680 (0.0)	58 AN2067.3 (0.0)	25[7] An11g04320 (0.0)	43 208780 (0.0)	59 AO090003000267 (0.0)	55 ATEG_06035.1 (0.0)	57 NCU03894.2 (0.0)	57 Pa_4_7000 (0.0)
STE50 (YCL032W)	1 Afu2g17130 (4e-19)	1 AN7252.3 (2e-16)	1 An04g09220 (2e-18)	1 51404 (3e-17)	1 AO090102000104 (7e-17)	1 ATEG_10030.1 (5e-16)	1 NCU00455.2 (2e-14)	1 Pa_7_6830 (2e-15)

[7] Maximum number of hits was limited in 25 hits per blast search.

3.3 Comparison of two A. niger genomes

In Table 4 three of the *S. cerevisiae* proteins are shown only in one of the two *A. niger* strains (meiotic recombination protein MSH4 (no hit) and alpha factor pheromone receptor STE2 (E-value below threshold) are not present in strain ATCC1015 and mismatch repair protein MSH6 (counterblasting with the *A. niger* protein resulted SRP40 protein as best hit in the *S. cerevisiae* database) is not present in CBS513.88). In addition, there are nine more genes which were found neither in any of the two *A. niger*s nor in other filamentous fungi. The two *A. niger* strains possessed a very similar set of mating-related genes. The investigated protein homologs in the *A. niger* strains showed no or little polymorphism between each other, only a few amino acid substitutions or single missing amino acids generate slight polymorphism (Table 5). When proteins from the two *A. niger* databases were compared we found many, smaller or bigger annotation errors in both genomes. As a solution to this problem we aligned the DNA sequences from both *A. niger* genomes and changed the ORFs of the shorter protein according to the longer protein.

Table 5. Differences between mating related proteins of the two *A. niger* (CBS513.88 and ATCC1015) strains. If two proteins differ in size, the longer one is the basis for counting percentage of identities, similarities and gaps. Gaps are counted only in the homologous region.

Protein	Type and size of difference	Identical	Gaps
Meiotic proteins			
DMC1 (YER179W)	no difference	344/344 (100%)	(0%)
HOP1 (YIL072W)	no difference	783/783 (100%)	(0%)
MLH1 (YMR167W)	substitutions: 251, 376, 526, 537, 548, 665, 726, 731-732, 744, 760-761, 765	759-772 (98%)	(0%)
MLH3 (YPL164C)	no difference	943/943 (100%)	(0%)
MRE11 (YMR224C)	indel region: 644 substitutions: 506, 584, 732, 742	737/742 (99%)	1/742 (0%)
MSH2 (YOL090W)	substitution: 511	944/945 (99%)	(0%)
MSH4 (YFL003C)	only present in CBS513.88	–	–
MSH5 (YDL154W)	substitutions: 8, 141, 487, 664	892/896 (99%)	(0%)
MSH6 (YDR097C)	only present in ATCC1015	–	–
PMS1 (YNL082W)	substitutions: 6, 395, 490, 518, 576, 591, 610, 656, 771, 1029, 1050	1044/1055 (99%)	(0%)
RAD50 (YNL250W)	substitutions: 425, 479, 482, 512, 553, 565, 597, 729, 775, 832, 929, 937, 942, 991,	1280/1294 (99%)	(0%)
RAD51 (YER095W)	no difference	–	–
RAD52 (YML032C)	substitutions: 234, 291, 395	595/598 (99%)	(0%)

Table 5. (Continued) Differences between mating related proteins of the two *A. niger* (CBS513.88 and ATCC1015) strains. If two proteins differ in size, the longer one is the basis for counting percentage of identities, similarities and gaps. Gaps are counted only in the homologous region.

Protein	Type and size of difference	Identical	Gaps
Pheromone processing proteins			
KEX1 (YGL203C)	substitutions: 673, 721, 811	820/823 (99%)	(0%)
KEX2 (YNL238W)	no difference	–	–
RAM1 (YDL090C)		523/523 (100%)	(0%)
RAM2 (YKL019W)	no difference	–	–
RCE1 (YMR274C)	substitutions: 105, 113, 215	329/332 99%)0	(0%)
STE6 (YKL209C)	substitutions: 7, 125, 148, 208, 243, 247-248, 262, 452, 454, 468, 594, 634, 641, 658, 672, 740, 819, 912, 981-982, 985, 990, 1008, 1052, 1123	1306/1332 (98%)	(0%)
STE14 (YDR410C)	no difference	–	–
STE23 (YLR389C)	1st substitution: 774 2nd pair substitutions: 269, 1019 The two allele pairs differ from each other.	1166/1167 (99%) 1035/1037 (99%)	(0%) (0%)
STE24 (YJR117W)	no difference	–	–
Pheromone response proteins			
BEM1 (YBR200W)	substitutions: 96	600/601 (99%)	(0%)
CDC24 (YAL041W)	no difference	–	–
CDC42 (YLR229C)	1st pair: no difference 2nd pair: substitution: 48	250/250 (100%) and 198/199 (99%)	(0%) and (0%)
FUS3 (YBL016W)	no difference	–	–
GPA2 (YER020W)	no difference	–	–
STE2 (YFL026W)	only present in CBS513.88	–	–
STE3 (YKL178C)	substitutions: 134, 307	521/523 (99%)	(0%)
STE4 (YOR212W)	no difference	–	–
STE7 (YDL159W)	no difference	–	–
STE11 (YLR362W)	indel region: 87-88	903/905 (99%)	2/905 (0%)
STE12 (YHR084W)	no difference	–	–
STE20 (YHL007C)	substitution: 92	837/838 (99%)	0/838 (0%)
STE50 (YCL032W)	no difference	–	–

3.4 Screening for MAT homologs among black Aspergilli

Primers were developed for the detection of the *mat-1* locus in the black Aspergilli based on the *mat-1* homologs in the two sequenced *A. niger* genomes (Table 2). Degenerate primers were made for the *mat-2* locus based on conserved parts of the *mat-2* locus as found in *A. nidulans* and *A. fumigatus*. Blasting the genome of the *A. niger* ATCC1015 strain with the *A. nidulans* MAT-2 protein (AN4734.3) revealed the presence of two homologs with E-values of e^{-22} and e^{-21} respectively. Using these two proteins to counterblast in the GenBank, the best hit was the *A. fumigatus* MAT-2 protein (EAL89707.1) with E-values of 0.0 and e^{-177}. Although these values are low enough to show the presence of *mat-2* homologs in one of the sequenced *A. niger* strains (ATCC1015) and might suggest that *A. niger* is of heterothallic origin, the *mat-2* homolog loci did not bear all the essential key features important in real *mat-2* genes and thus may not be real function mating genes (Paul Dyer, personal communication). The testing of these degenerate primers on a subset of the wild-type black *Aspergillus* isolates did not yield clear results (multiple PCR products were found) on the presence of the *mat-2* locus in the majority of the strains, with the exception of two isolates, which resulted 1-1 fragment, showing similarity to *mat-2*.

For testing the presence of *mat-1*, we randomly chose 162 black *Aspergillus* isolates in a way that strains were originated from more than 20 countries worldwide and belong to five known black *Aspergillus* species (*A. brasiliensis*, *A. carbonarius*, *A. japonicus*, *A. niger* and *A. tubingensis*). We found that 137 (84.5%) isolates contained *mat-1* and 25 (15.5%) did not show the presence of this gene.

In the different black *Aspergillus* species we did find significantly different numbers of strains in which we could or could not detect the *mat-1* (Table 6). *A. niger* did show the highest number of strains with a detectable *mat-1* (92.7 %) relative to the number of strains were *mat*-1 could not be detected (7.3 %). Both *A. tubingensis* and *A. japonicus* differ significantly from the ratio found in *A. niger* and from each other (Table 7). *A. japonicus* even showed a nearly half to half ratio of strains with a detectable versus a non-detectable *mat*-1.

We found one isolate, the *A. brasiliensis* in which *mat-1* was not detectable, but produced a small fragment with *mat-2* PCR. This product was sequenced and the translated sequence was validated with blasting in the GenBank), and showed similarity to a small fragment (41 out of the total length of 382 amino acids) of the *A. fumigatus* MAT-2 (68% identity and 90 similarity in a 41 amino acid long range). After comparing the *A. brasiliensis* MAT-2 polypeptide fragment to the *A. fumigatus* MAT-2 protein we found that the *A. brasiliensis* MAT-2 protein shows conserved regions to the latter one, but these regions show similarity to four different regions of the *A. fumigatus* MAT-2 protein. Between these homologous regions there are long gaps.

We did not find remarkable differences in the number of *mat-1* positive and negative isolates between different countries (data not shown).

Table 6. Presence and absence (not detectable with the PCR test)of *mat-1* and *mat-2* mating type alleles in natural black *Aspergillus* isolates. *Mat-1* and *mat-2* alleles were found in different isolates of *A. tubingensis*.

Species	**Nr of isolates**	***mat-1 var***		***mat-2***	
		detectable	not detected	detectable	not detected
A. brasiliensis	1		1-	1+	-
A. carbonarius	1	1+		-	1-
A. japonicus	10	6+	4-	-	10-
A. niger	83	77+	6-	-	83-
A. tubingensis	65	53+	12-	1+	64-
unknown	3	2+	1-	-	3-
Total	162	137	25	2+	160-

Table 7. P-values of Chi-square test of wild-type isolates in which *mat-1* was detected. Significanly different values are marked with *.

Chi-square	**A. brasil.**	**A. carb.**	**A. japonicus**	**A. niger**	**A. tubing.**	**unknowns**
A. brasiliensis	1	0.00*	0,37	0,00*	0,02*	0,16
A. carbonarius		1	0,26	0,78	0,66	0,48
A. japonicus			1	0.00*	0,00*	0,16
A. niger				1	0,03*	0.00*
A. tubingensis					1	0,01*
unknowns						1

4. Discussion

Sexual and asexual propagation are two alternative ways of fungal reproduction. Whereas certain strains are able to undergo both, others reproduce exclusively sexually or asexually. The black mould *A. niger* and several closely related species constitute the black *A. niger* aggregate and are generally considered to be strictly asexual fungi. These asexual *Aspergillus* species are of high importance as human pathogens and as industrial microorganisms. It is crucial to know more about their reproductive cycles for better disease control or for applications in industry. Recently sequenced genomes are adequate tools to compare fungi with different life cycles (i.e. sexual/asexual) and make conclusions about the way of reproduction of related species.

In this study we investigated the sequenced genomes of eight ascomycete filamentous fungi and especially the two genomes of two heterokaryon incompatible presumed asexual *A. niger* strains for the presence of mating-related genes. In addition we also performed PCR on wild-type isolates belonging to *Aspergillus* section *Nigri* to search for the presence of the mating-type genes *mat-1* and *mat-2*. We found that the majority of the genes present in the sexual *A. nidulans* are also present in asexual *Aspergillus* species, such as the two *A. niger* isolates. There are a few genes in Table 3 and Table 4, which are present only in one of the *A. niger* strains. It does not always mean that we did not find homolog of the given gene, but the E-value of the hit was

below threshold. The presence of homologs of majority of *S. cerevisiae* genes in the filamentous fungi strongly supports the hypothesis that sexual reproduction arose early in a common ancestor to all living eukaryotes (Ramesh *et al.* 2005). We did not find remarkable differences between the genes and proteins of the different *A. niger* isolates, which is interesting in a supposedly asexual fungus and lead to the question: why are meiosis-related genes preserved in an asexual species? Many genes having a role in meiotic processes in *S. cerevisiae* were found to be preserved in *A. niger*. This may suggest a possible sexual cycle, but these proteins may also be involved in other processes. Whether these genes are expressed in *A. niger* is not known. These *in silico* results are not conclusive, laboratory experiments were required to get a better insight into possible sexuality of *A. niger*.

The primers that we used for the detection of the *mat-1* locus in the black Aspergilli were developed based on homologs in the two sequenced *A. niger* genomes. In the different black *Aspergillus* species we did find significantly different numbers of strains in which we could or could not detect *mat-1*. *A. niger* showed the highest number of strains with a detectable *mat-1* homolog relative to the number of strains where *mat-1* could not be detected. In *A. japonicus* however in about half of the strains *mat-1* was detected. This may mean that the developed primers function better on the *A. niger mat-1* sequences, then on some of the other black *Aspergillus* species like *A. japonicus*.

The primers used for the *mat-2* locus were designed on conserved parts of the *mat-2* locus found in the genomes of *A. nidulans* and *A. fumigatus*. Testing these primers *in silico* on the two sequenced *A. niger* genomes did show some homologous sequences only in *A. niger* ATCC1015, but no complete and functional *mat-2* alleles (Paul Dyer, pers. comm.) and none in the other sequenced strain. Testing the degenerate primers on wild-type black *Aspergillus* isolates also did not yield clear results on the presence of the *mat-2* locus in the majority of the isolates either in a heterothallic way (with different sexes harbouring different mating types) or in a homothallic way (with strains harbouring both mating types). However, *mat-2* could only be detected in two isolates. We found but one strain (the *A. brasiliensis* isolate) that yielded a *mat-2* PCR product, which translated to amino acids exhibited similarity to the *A. nidulans* MAT-2. However, the resulting fragment was very short and fragmented compared to that of the *A. fumigatus* MAT-2, therefore it is doubtful that the *mat-2* locus still functions in this strain. The lack of clear results (PCR products) of the *mat-2* locus in the black *Aspergillus* isolates could be due to lack of recognition of the degenerated primers or due to the fact that *mat-2* genes are really absent.

There are two possible explanations for the high genetic diversity and vegetative incompatibility observed amongst natural isolates of the black *Aspergillus* aggregate (van Diepeningen *et al.* 1997). A cryptic sexual cycle or - in evolutionary terms - recently lost sexuality. Looking at recombination between molecular phylogenies of different nuclear genes no evidence was found for recombination between them (van Diepeningen 1999). Also looking at the mating genes, we saw an unequal distribution in the number of *mat-1* and (maybe) *mat-2*. For instance in *A. niger* we saw that 93% of the isolates harbour the *mat-1*, whereas in *A. tubingensis* this is 81.5 %. Such

a strong deviation from the expected ratio indicates that the examined isolates are reproducing asexually and that sexuality is lost in these species, because in sexual species different mating-types are present almost in an equal ratio. In the *A. fumigatus*, which is presumably a cryptic sexual species, they found a quite balanced distribution of *mat-1* and *mat-2* idiomorphs (43 % and 57 %, respectively) (Paoletti *et al.* 2005).
Both homothallic (*A. nidulans*) and heterothallic (*A. heterothallica* – Kwon and Raper 1967) species are present in the *Aspergillus* genus. Many asexual isolates investigated in this paper show heterothallic features (lack of either *mat-1* or *mat-2*), but more detailed analysis is needed.
We searched the genomes of *A. niger* strains and several other (filamentous) fungi and designed and tested *mat-1* and *mat-2* specific primers on more than 160 black *Aspergillus* isolates. Quite a large number of meiosis related genes proved well conserved, hinting either on meiosis occurring or on well conserved other functions for these genes. We found that the ratio of isolates showing the presence of *mat-1* is species dependent, whereas *mat-2* PCR results were not as clear (due to the degenerated primers or heterothallism, the lack of *mat-2* locus), and we found only one *mat-2* positive isolate. Our findings together with previous results indicate that the black Aspergilli are probably indeed asexuals. We suggest further analysis of the *A. niger* genomes supported with functional analysis of mating-related genes, and investigation of more black *Aspergillus* isolates with PCR or Southern-blot to once and for all solve the question of asexuality.

Acknowledgments

We thank the following institutes and companies and individuals involved in the genome projects of the species used in this study: genome of the culture collection strain CBS513.88 from DSM (prepared by shotgun sequencing, 7.5 times coverage; http://www.dsm.com/en_US/html/dfs/genomics_aniger.htm) and genome of the culture collection strain ATCC1015 from DOE Joint Genome Institute (assembly release version 1.0 of whole genome shotgun reads was constructed with the JGI assembler, Jazz, using paired end sequencing reads at a coverage of 8.9X; http://www.jgi-psf.org/cgi-bin/runAlignment?db=Aspni1&advanced=1). The *A. fumigatus* preliminary sequence data was obtained from The Institute for Genomic Research website at http://www.tigr.org. *A. nidulans*, *N. crassa* (release 7) and *A. terreus* sequence data were from the *Aspergillus nidulans*, *Neurospora crassa* and *Aspergillus terreus* Sequencing Project, Broad Institute of MIT and Harvard (http://www.broad.mit.edu). *A. oryzae* sequences were available on the server of National Institute of Technology and Evaluation (NITE). The *P. anserina* genome was published by the Institut de Génétique et Microbiologie (Université de Paris-Sud XI / CNRS)
We thank the lab assistance in the Laboratory of Genetics, Wageningen University and Department of Microbiology, University of Szeged. We thank Paul Dyer and Mathieu Paoletti for comments and recommendations.

References

Acccensi F, Cano J, Figuera L, Abarca ML, Cabañes FJ, 1999. New PCR method to differentiate species in the *Aspergillus niger* aggregate. *FEMS Microbiology Letters* 180, 191-196.

Barton NH, Charlesworth B, 1998. Why sex and recombination? *Science* 281, 1986-1990.

Cavalier-Smith T, 2002. Origins of the machinery of recombination and sex. *Heredity* 88, 125-141.

de Visser JGM, Elena SF, 2007. The evolution of sex: empirical insights into the roles of epistasis and drift. *Nature Reviews Genetics* 8, 139-149.

Dyer PS, Paoletti M, Archer DB, 2003. Genomics reveals sexual secrets of *Aspergillus*. *Microbiology* 149, 2301-2303.

Dyer PS, Paoletti M, 2005. Reproduction in *Aspergillus fumigatus*: sexuality in a supposedly asexual species? *Medical Mycology Supplement* 1, 43, S7_/S14)

Geiser DM, Timberlake WE, Arnold ML, 1996. Loss of meiosis in *Aspergillus*. *Molecular Biology and Evolution* 13, 809-817.

Glass NL, Jacobson DJ, Shiu PKT, 2000. The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycete fungi. *Annual Review of Genetics* 34, 165-186.

Hamari Z, Kevei F, Kovács É, Varga J, Kozakiewicz Z, Croft JH, 1997. Molecular and phenotypic characterization of *Aspergillus japonicus* strains with special regard to their mitochondrial DNA polymorphisms. *Antonie van Leeuwenhoek* 72, 337-347.

Hurst LD, Peck JR, 1996. Recent advances in understanding of the evolution and maintenance of sex. *Trends in Ecology and Evolution* 11, 46-52.

Kevei F, Hamari Z, Varga J, Kozakiewicz Z, Croft JH, 1996. Molecular polymorphism and phenotypic variation in *Aspergillus carbonarius*. *Antonie van Leeuwenhoek* 70, 59-66.

Kwon KJ, Raper KB, 1967. Sexuality and cultural characteristics of *Aspergillus heterothallicus*. *American Journal of Botany* 54, 36-48.

Latgé JP, 1999. *Aspergillus fumigatus* and aspergillosis. *Clinical Microbiology Reviews* 12, 310-350.

Milgroom MG, Fry WE, 1997. Contributions of population genetics to plant disease epidemiology and management. *Advances in Botanical Research* 24, 1-30.

Paoletti M, Rydholm C, Schwier EU, Anderson MJ, Szakacs G, Lutzoni F, Debeaupuis JP, Latgé JP, Denning DW, Dyer PS, 2005. Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Current Biology* 15, 1242–1248.

Pontecorvo G, Roper JA, Forbes E, 1953. Genetic recombination without sexual reproduction in *Aspergillus niger*. *Journal of General Microbiology* 8, 198–210.

Pontecorvo G, 1956. The parasexual cycle in fungi. *Annual Review of Microbiology* 10, 393-400.

Pöggeler S, 2002. Genomic evidence for mating abilities in the asexual pathogen *Aspergillus fumigatus*. *Current Genetics* 42, 153-160.

Ramesh MA, Malik SB, Logsdon Jr JM, 2005. A phylogenomic inventory of meiotic genes: evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. *Current Biology* 15, 185–191.

Samson RA, Houbraken JAMP, Kuijpers AFA, Frank JM, Frisvad JC, 2004. New ochratoxin or sclerotium producing species in *Aspergillus* section *Nigri*. *Studies in Mycology* 50, 45-61.

Saupe SJ, 2000. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews* 64, 489-502.

Turgeon BG, Yoder OC, 2000. Proposed nomenclature for mating type genes of filamentous ascomycetes. *Fungal Genetics and Biology* 31, 1-5.

Van Diepeningen AD, Debets AJM, Hoekstra RF, 1997. Heterokaryon incompatibility blocks virus transfer among natural isolates of black *Aspergilli*. *Current Genetics* 32, 209-217.

Van Diepeningen AD, 1999. Horizontal transfer of genetic elements in the black Aspergilli. PhD Thesis, Wageningen University, The Netherlands.

Varga J, Kevei F, Fekete C, Coenen A, Kozakiewicz Z, Croft JH, 1993. Restriction fragment length polymorphisms in the mitochondrial DNAs of the *Aspergillus niger* aggregate. *Mycological Research* 97, 1207-1212.

Varga J, Kevei F, Vriesema A, Debets F, Kozakiewicz Z, Croft JH, 1994. Mitochondrial DNA restriction fragment length polymorphisms in field isolates of the *Aspergillus niger* aggregate. *Canadian Journal of Microbiology* 40, 612-621.

Varga J, 2003. Mating type gene homologs in *Aspergillus fumigatus*. *Microbiology* 149, 816-819.

Woo PCY., Chong KTK., Tse H., Cai JJ., Lau CCY., Zhou AC., Lau SKP, Yuen KY 2006. Genomic and experimental evidence for a potential sexual cycle in the pathogenic thermal dimorphic fungus *Penicillium marneffei*. *FEBS Letters* 580, 3409-3416.

CHAPTER 6

General Discussion

The study presented in this thesis was aimed to get a better insight into the background of both vegetative incompatibility and mating in filamentous fungi, with emphasis on the species *A. niger* and *P. anserina*. As for all organisms reproduction, both asexual and sexual, is of supreme importance for fungi.

The black mould *A. niger* is a real cosmopolitan species and is very versatile in its metabolism, and thus widely used by the industry. It has a 'Generally Recognised As Safe' (GRAS) status, given by the US Food and Drug Administration, but occasionally it is known to act as opportunistic pathogen. It occurs worldwide mainly as a saprobic fungus with a preference for subtropical and tropical regions. In soil the number of propagules of *A. niger* and related species can reach very high densities. *A. niger* and the other black Aspergilli have the unique capability to grow on very high concentrations of tannic acid. Plants can contain 5-20 % tannin by weight, and in certain experiments the *A. niger* was able to grow on 80 % tannin by weight. It is difficult to mineralize tannin-protein complexes, therefore this tannine-degrading feature of *A. niger* can be of high importance for the N-release in soil. Thus, *A. niger* probably has a unique ecological niche (van Diepeningen *et al.* 2004).

A. niger is known as an asexual fungus and at the same time natural isolates of this species show a high degree of vegetative incompatibility with each other. Vegetative incompatibility needs to be tested in this species genetically as no visible manifestations like barrges are formed. There are different explanations for the role of heterokaryon incompatibility (Saupe 2000). In *A. niger* there are two main reasons why vegetative incompatibility may be adaptive. At first, it seems to be very efficient against infection of mycoviruses. Approximately 10 % of the natural isolates of *A. niger* is infected with dsRNA viruses, which may decrease the fitness of their hosts. It is almost impossible to get rid of mycovirus infection through asexual reproduction, and it is very rare that conidiospores produced by infected mycelia do not contain the virus (van Diepeningen *et al.* 1997). Another reason to maintain incompatibility is related to the life-style. In tropical regions every gram of soil may contain many spores, even up to 250 in a gram. This means a strong competition between the colonies arising from the spores, in which preserving identity is very important to avoid resource exploitation. Therefore, heterokaryon incompatibility may have a crucial role for *A. niger*, to preserve its fitness and identity.

The coprophilic ascomycete *P. anserina* is one of those exceptional fungal species, which senesce. All natural isolates of this species die within a few weeks. The explanation for this phenomenon probably lies in its life on an ephemeral substrate. *P. anserina* obligately grows on herbivore dung. Many fungal species compete at the same time for the limited resources and the dung dries out quickly. The competition and limited nutrition source force the *P. anserina* to speed up its life cycle (Maas 2005). *P. anserina* is an obligately sexual organism in which pseudo-homothallism ensures that single spores containing both mating types in two different nuclei are able to reproduce as soon as possible. In the environment formed by herbivore dung vegetative incompatibility (manifesting as 'barrage') can be an efficient tool against resource exploitation, like in the previously mentioned *A. niger*. Anyhow, heterokaryon incompatibility is

here not an absolute barrier against the horizontal transfer of genetic materials as e.g. the mitochondrial pAL2-1 plasmid is transferred easily between colonies. The effective horizontal transfer may make up the loss due to the inefficient vertical transfer (via sexual reproduction), especially in outcrossing (van der Gaag *et al.* 1998).

In **Chapter 1** a general overview is given about vegetative incompatibility and fungal mating. The life cycles and some other features of both *A. niger* and *P. anserina* are described, and aims of this thesis are represented.

Chapter 2 describes a comparative *in silico* genome analysis, performed in order to reveal (dis)similarities in heterokaryon incompatibility (*het*) and apoptosis (PCD) related genes between eight fungal species. Seven filamentous and one yeast Ascomycete genomes were investigated. We searched for homologs of *N. crassa het*-genes, *P. anserina het*, modifier and suppressor genes and apoptosis related genes of *S. cerevisiae* in the genome databases. Most of the genes were present both in filamentous fungi and yeast. Nine genes were present only in filamentous fungi and we did not find any proteins specific for yeast. Differences between filamentous fungi are often present only in the different number of alleles. *P. anserina* harboured three *idi* (*i*nduced *d*uring *i*ncompatibility) genes, which were not found in any other species. Next to this, the results show that many genes are well conserved, but only a functional analysis could give a more precise picture, whether these genes still function in each of the strains or not and what these functions are. We compared the *het*-gene homologs from two *A. niger* strains that we show to be incompatible, and found minor variability between their genes and the proteins encoded by them. These small differences may be sufficient to elicit incompatibility, but only further functional analysis could truelly prove this. The main result of our research in this chapter is the description of the presence of *het*-genes in the *A. niger* isolates. Based on the sequence data collected in this chapter, practical examinations (isolation/deletion of genes, etc.) can be performed in the black *A. niger*.

Analysis of a single *A. niger* gene, the *N. crassa het-c* homolog, was the focus of **Chapter 3**. This gene has three alleles in *N. crassa* and in related *Sordaria* species, showing trans-species polymorphism and balancing selection (Wu *et al.* 1998). The three alleles of *het-c* in *N. crassa* form an allelic incompatibility system (Saupe and Glass 1997). Natural isolates of *A. niger* are highly incompatible with each other, but the genetic background of incompatibility is not known. Due to the detailed analysis, the *N. crassa het-c* gene is well characterized . The *het-c* gene was first identified *in silico* in the *A. niger* genome, and we could design primers to carry out PCR based analyses to detect size polymorphisms in 99 natural black *Aspergillus* isolates. Next to that, the *het-c* gene was partially sequenced from nine randomly choosen strains. All the polymorphic sites known from *N. crassa* were the uniform in *A. niger.* But, we found one polymorphic site in the *A. niger het-c*, with different numbers of glutamine repeats. Glutamine repeats may originate

from polymerase slippage, resulting in a variable number of glutamine residues (Zoghbi and Orr 2000). Such variable number of glutamine repeats e.g. found in neurodegenerative human diseases, like Huntington disease, where. the role of the polyglutamine units might be to cause protein aggregation (Perutz *et al.* 1994, Rega *et al.* 2001, Masino and Pastore 2002). A function of the repeats in *A. niger's het-c* is unknown. A further functional analysis of the *het*-c gene was performed in *A. niger*, where a laboratory strain was transformed with each of the 3 known *het-c* alleles of *N. crassa* separately. Only the *het-c*^{Panama} allele triggered incompatibility and no viable colonies recovered after transformation with it. This finding corresponds with a similar study done in *P. anserina* (Saupe *et al.* 2000).
A possible sequel of our examinations can be the interruption of *het-c* in *A. niger*, to see what cellular functions are coupled to this locus. The introduction of *het-c* genes from different isolates with different number of glutamine-repeats into a tester strain may show whether this motif does have any influence on vegetative incompatibility.

Chapter 4 reports on the analysis of *hch* and *het-c* in *P. anserina*. The *hch* is the homolog of the allelic *N. crassa het-c*, but a small screening of *P. anserina* strains did not result in observed polymorphism (Saupe *et al.* 2000). We examined more than 100 *P. anserina* isolates (collected around Wageningen, The Netherlands; van der Gaag 2005), and our results confirmed those found by Saupe *et al.* (2000). Neither the CAPS analysis nor the sequencing of 11 randomly choosen isolates revealed any polymorphism. So the absence of polymorphy in the *hch / N. crassa het-c* homologs, in wildtypes of both *A. niger* and *P. anserina* suggests that the gene may only be functioning as heterokaryon incompatibility system in *N. crassa* and that in the other species it may have only another, perhaps conserved, function. As a *het* locus can be any locus, at which heteroallelism is not tolerated in a heterokaryon (Saupe 2000), different fungal species may have selected totally different sets of heterokaryon incompatibility genes.
The *P. anserina het-c* gene, which is only in its name identical to that of in *N. crassa*, is involved in non-allelic incompatibility in the *het-c*/*het-d* and *het-c*/*het-e* interactions. Components of these two non-allelic systems seem to be under positive selection (Saupe *et al.* 1995, Saupe 2000), and suggest that rapid, concerted evolution is acting on these loci. As a proof of this rapid evolutionary process, high variability in the *het-d* and *het-e* loci were found (Paoletti *et al.* unpublished), but only a limited number (four) of *het-c* alleles. In our experiments we sequenced 15 wild-type isolates of our collection and we found eight different alleles among which five so far unknown. The polymorphisms were located at certain amino acid positions, which are crucial in generation of new alleles and the gene seems to be under positive selection. In the future an interesting experiment could be the sequencing of more of the available 100 *P. anserina* isolates in our collection looking at *het-c* as well as *het-d* and *het-e*. Also a functional analysis of the new found *het-c* alleles and interacting *het-d* and *het-e* alleles is interesting.

In **Chapter 5** we focused on the mating/meiosis related genes in Ascomycetes, including species related to the *A. niger*. Genomes of sexual and asexual filamentous fungi and the yeast *S. cerevisiae* were compared, to find out whether asexual species do still bear some meiotic genes. We only found 9 loci present in sexual *S. cerevisiae* which are not present in the examined filamentous fungi, including the asexual species. If these genes are still functional, this finding could hint that these species have (cryptic) sexual cycles (as it was found in *A. fumigatus*; Varga 2003, Dyer *et al.* 2003, Paoletti *et al.* 2005) or more likely that these meiotic genes have further roles also in asexual species. Another possibility, is that though the found genes are very conserved, they are not active anymore as the asexual species lost their sexuality but recently (in evolutionary term).

We also investigated the presence of *mat-1* and *mat-2* idiomorphs of the mating type genes known from sexual *Aspergillus* species in 162 black *Aspergillus* isolates and found that the majority (84.5 %) of the strains harbour *mat-1,* quite a few contain no mating type homologue and only two examined strains bear some remnants of the *mat-2* idiomorph. We did find some differences between the ratios of the mating gene ideomorphs between different black Aspergillus species. In *A. niger* and *A. tubingensis* isolates the *mat-1* dominates (around 90%), whereas in the *A. japonicus* we could detect the *mat-1* only in 6 samples out of 10. In the previously considered asexual *A. fumigatus*, a quite balanced distribution of *mat-1* and *mat-2* idiomorphs was found (43 % and 57 %, respectively) and that species is now presumed to be a cryptic sexual species (Paoletti *et al.* 2005). Our findings of a very skewed mating type ration, the absence of both idiomorphs in quite a number of strains and rare presence of a degenerated *mat-2* suggest that the *A. niger* and *A. tubingensis* truelly lost their sexuality.

A further functional analysis can reveal if the found meiotic genes are still functional and why they are still present in an asexual species. Another possibility is to introduce a functional *mat-2* idiomorph into *A. niger* and investigate if sexuality is still possible in an asexual species.

Both of the *A. niger* and *P. anserina* show clear signs of vegetative incompatibility. In consequence of the environment they live in, heterokaryon incompatibility is determinative in their life. Both species have to protect their integrity, to avoid resource exploitation and/or spread of harmful genetic elements. While the genetic background of heterokaryon incompatibility is relatively well studied in *P. anserina*, the phenomena was largely unknown in the asexual *A. niger*. In this thesis I tried to reveal some secrets of the background of vegetative incompatibility in *A. niger*, and at the same time I aimed to get a better understanding on the sexuality related genes in this asexual species.

References

Dyer PS, Paoletti M, Archer DB, 2003. Genomics reveals sexual secrets of *Aspergillus*. *Microbiology* 149, 2301-2303.

Maas FPMM, 2005. On the ecology and evolution of fungal senescence. PhD Thesis, Wageningen University, The Netherlands.

Masino L, Pastore A, 2002. Glutamine repeats: structural hypotheses and neurodegeneration. *Biochemical Society Transactions* 30, 548-551.

Paoletti M, Rydholm C, Schwier EU, Anderson MJ, Szakacs G, Lutzoni F, Debeaupuis JP, Latgé JP, Denning DW, Dyer PS, 2005. Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. *Current Biology* 15, 1242–1248.

Perutz MF, Johnson T, Suzuki M, Finch JT, 1994. Glutamine repeats as polar zippers: their role in inherited neurodegenerative diseases. *Proceedings of the National Academy of Sciences of the USA* 91, 5355-5358.

Rega S, Stiewe T, Chang DI, Pollmeier B, Esche H, Bardenheuer W, Marquitan G, Putzer BM, 2001. Identification of the full-length huntingtin-interacting protein p231HBP/HYPB as a DNA-binding factor. *Molecular and Cellular Neuroscience* 18, 68-79.

Saupe S, Descamps C, Turcq B, Bégueret J, 1994. Inactivation of the *Podospora anserina* vegetative incompatibility locus *het-c*, whose product resembles a glycolipid transfer protein, drastically impairs ascospore production. *Proceedings of the National Academy of Sciences of the USA* 91, 5927-5931.

Saupe S, Turcq B, Bégueret J, 1995. Sequence diversity and unusual variability at the *het-c* locus involved in vegetative incompatibility in the fungus *Podospora anserina*. *Current Genetics* 27, 466-471.

Saupe SJ, Glass NL, 1997. Allelic specificity at the *het-c* heterokaryon incompatibility locus of *Neurospora crassa* is determined by a highly variable domain. *Genetics* 146, 1299-1309.

Saupe SJ, 2000. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews* 64, 489-502.

Saupe SJ, Clave C, Sabourin M, Begueret J, 2000. Characterization of *hch*, the *Podospora anserina* homolog of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa*. *Current Genetics* 38, 39-47.

van Diepeningen AD, Debets AJM, Hoekstra RF, 1997. Heterokaryon incompatibility blocks virus transfer among natural isolates of black Aspergilli. *Current Genetics* 32, 209-17.

van Diepeningen AD, Debets AJM, Varga J, van der Gaag M, Swart K, Hoekstra RF, 2004. Efficient degradation of tannic acid by black *Aspergillus* species. *Mycological Research* 108, 919-925.

van der Gaag M, Debets AJM, Osiewacz HD, Hoekstra RF, 1998. The dynamics of pAL2-1 homologous linear plasmids in *Podospora anserina*. *Molecular and General Genetics* 258, 521-529.

van der Gaag M, 2005. Genomic conflicts in *Podospora anserina*. PhD Thesis, Wageningen University, The Netherlands.

Varga J, 2003. Mating type gene homologues in *Aspergillus fumigatus*. *Microbiology* 149, 816-819.

Wu J, Saupe SJ, Glass NL, 1998. Evidence for balacing selection operating at the *het-c* heterokaryon incompatibility locus in a group of filamentous fungi. *Proceedings of the National Academy of Sciences of the USA* 95, 12398-12403.

Zoghbi HY, Orr HT, 2000. Glutamine repeats and neurodegeneration. *Annual Review of Neuroscience* 23, 217-247.

Summary

This thesis deals with a comparative analysis of genetic incompatibility in two filamentous fungi. For the asexual black mould *Aspergillus niger* we looked at heterokaryon incompatibility and meiosis related genes. For *Podospora anserina*, an obligate coprophilous and sexual species, we focused on heterokaryon incompatibility.

A. niger is a versatile organism used in industry and sometimes found as opportunistic pathogen. The fungus is found world-wide with a preference to (sub)tropical regions. In nature *A. niger* has a mainly saprobic lifestyle and probably has a special niche in degrading plant polymers like tannic acid and possibly in the control of the release of Nitrogen compounds found in complexes with these plant polymers. *A. niger* is an asexual species, but heterokaryon compatible strains are known to be able to go through a parasexual cycle. However, most natural isolates proved to be heterokaryon incompatible. Both incompatibility and asexuality have so far hindered genetic analysis of the species.

Now, two strains of *A. niger* have been sequenced and we used their genomes to analyze them and several other filamentous fungi and the yeast *Saccharomyces cerevisiae in silico* for heterokaryon incompatibility, apoptosis- and meiosis related genes. Experimentally we show the two sequenced *A. niger* strains to be heterokaryon incompatible and the observed small differences in their sequences homologous to known heterokaryon incompatibility genes, may be the reason of this observed incompatibility. The majority of heterokaryon incompatibility, apoptosis and meiosis-related genes proved well conserved in *A. niger* as well as in the other sexual and asexual filamentous fungi.

Of the mating type alleles we found *mat-1* to be present in both sequenced *A.nger* genomes. Remnants of the other *mat-2* allele were found in one of the sequenced genomes, but thus degenerated as to be probably non-functional. A screening of more than 160 wild-type black Aspergillus isolates with degenerate primers for *mat-2* showed only a presence of some faintly homologous sequences in two strains. *Mat-1* proved present in approximately 85% of the isolates, and in the remainder of strains no mating type genes could be detected. This much skewed ratio between the two mating type alleles indicates true asexuality of *A. niger*.

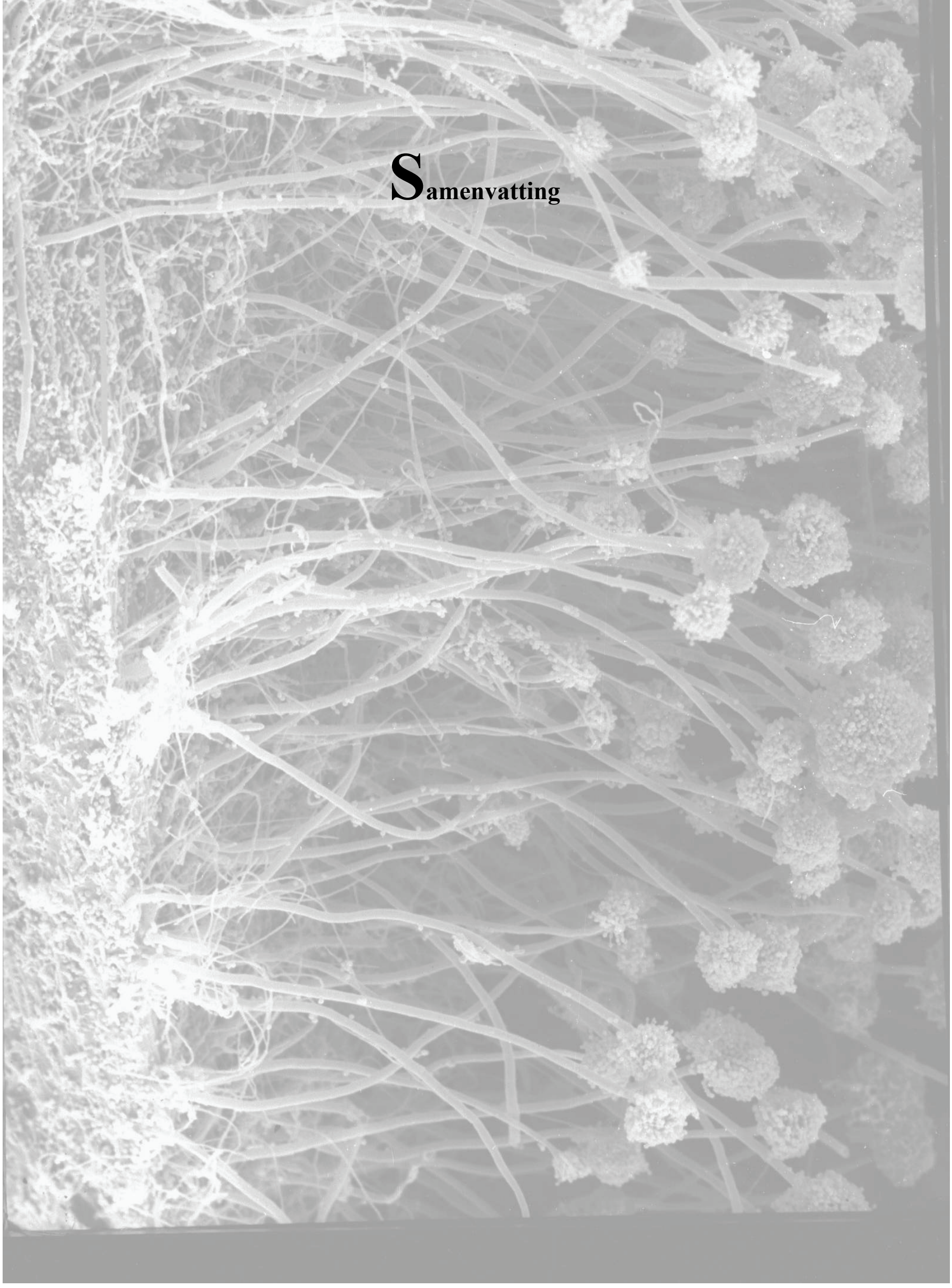
One putative heterokaryon incompatibility gene, the allelic *het-c* gene, was studied in depth. Both *het-c* and known modifier genes can be found in the *A. niger* genome. A screening of a natural population of 99 strains, showed no variation in the polymorphic site of the gene known from *Neurospora crassa*. Some polymorphisms observed in a glutamine rich area of the gene are probably the result of polymerase slippage, but it is unsure whether they have a function in incompatibility. Introduction of the *N. crassa het-c*^{Panama} allele showed the *het-c* gene can function as a lethal heterokaryon incompatibility gene, but the other results indicate that in the *A. niger* population it probably does not function as a heterokaryon incompatibility gene.

More is known about heterokaryon incompatibility in *P. anserina*, where both allelic and non-allelic heterokaryon incompatibility systems are found. Pseudo-homothallic *P. anserina* grows on the ephemeral substrate of herbivore dung and heterokaryon incompatibility may be important to protect a strain's identity and may limit the spread of infectious elements. *P. anserina* is one of the few fungal species known to senesce, perhaps due to the normally limited lifespan of its substrate. We studied two (putative) incompatibility genes in depth in *P. anserina*.

The *hch*-gene is homologous to the allelic *het-c* gene in *N. crassa* and *A. niger*. In the genome databases the gene proves to be well conserved and also known modifier genes have homologs in the genome. We screened 113 *P. anserina* wild-type isolates for polymorphism in this gene, but found none. This indicates that the *hch*-gene may not function as incompatibility gene in this species.

The *P. anserina het-c* locus is known to form dynamic non-allelic incompatibility systems with both the *het-d* and *het-e* genes. Of all three genes several alleles are known. Screening of a part of our wild-type population yielded at least five new *het-c* alleles with the indication that many more may be present. More non-synonymous than synonymous substitutions were found in the new alleles and especially sites at the outside of the HET-C protein were favored. Thus, the *het-c* gene seems to be under positive selection in *P. anserina*.

Samenvatting



Dit proefschrift handelt over een vergelijkende analyse van genetische incompatibiliteit in twee filamenteuze schimmels. Bij de asexuele zwarte schimmel *Aspergillus niger* is gekeken naar heterokaryon incompatibiliteit en meiose gerelateerde genen. Bij *Podospora anserina*, een obligate coprofiele en sexuele soort lag de focus op heterokaryon incompatibiliteit.

A.niger is een veelzijdig organisme dat in de industrie wordt gebruikt maar soms ook als opportunistische pathogeen wordt aangetroffen. De schimmel heeft een werelwijde verspreiding met een voorkeur voor de (sub-) tropische regio's. *A.niger* heeft in de natuur een voornamelijk saprobe levensstijl en vindt waarschijnlijk een speciale niche in de afbraak van plant polymeren zoals tannines en de controle van vrijkomen van stikstof uit complexen met deze plant polymeren. *A.niger* is een asexuele soort, maar heterokaryon compatibele stammen zijn in staat tot parasexuele recombinatie. De meeste natuurlijke isolaten zijn echter onderling heterokaryon incompatibel zodat klassiek genetische analyse niet mogelijk is.

De genoom sequenties van twee heterokaryon incompatibele isolaten van A.*niger* , enkele andere filamenteuze schimmels en *Saccharomyces cerevisiae* zijn geanalyseerd voor genen betrokken bij heterokaryon incompatibiliteit, apoptose en meiose. De meeste van deze genen bleken zowel in *A.niger* als in de andere sexuele en asexuele filamenteuze schimmels geconserveerd te zijn.

Van de mating-type allelen is *mat-1* aangetroffen in beide *A.niger* genomen, van *mat-2* zijn alleen gedegenereerde restanten aangetroffen in een van de genomen. Bij een screening van 160 natuurlijke isolaten van zwarte aspergilli werd *mat-1* aangetroffen in ongeveer 85% van de isolaten. Met behulp van gedegenereerde primers werden in twee isolaten zwak homologe sequenties van *mat-2* gevonden. Deze scheve verdeling van de twee mating type allelen suggereert dat *A.niger* werkelijk asexueel is.

Het *Neurospora crassa* heterokaryon incompatibiliteits gen, *het-c* en bekende *modifier* genen zijn ook aangetroffen in het genoom van *A.niger.* Analyse van 99 natuurlijke isolaten toonde echter geen variatie in de polymorfe sites zoals die bij *N.crassa* waren gevonden. Introductie van een van de drie *N.crassa het-c* allelen (*het-c*^{pa}) leidde tot een lethale reactie wat suggereert dat *het-c* in *A.niger* kan functioneren als heterokaryon incompatibiliteitsgen. De afwezigheid van allelische variatie in de natuurlijke isolaten maakt een dergelijke functionele rol echter onwaarschijnlijk.

Podospora anserina heeft zowel allelische als non-allelische heterokaryon incompatibiliteit. Twee vermoedelijke incompatibiliteitsgenen zijn in meer detail bestudeerd.

Het *hch*-gen van *P.anserina* is homoloog met het allelische *het-c* gen in *N.crassa en* A.niger. In de genomische databases is dit gen evenals gerelateerde *modifier* genen goed geconserveerd. Bij een analyse van 113 natuurlijke *P.anserina* isolaten werd geen polymorfisme voor *hch* aangetroffen. Dit geeft aan dat dit gen geen rol speelt bij heterokaryon incompatibiliteit in deze soort.

Het *P.anserina het-c* locus is betrokken bij een dynamisch non-allelisch incompatibiliteits systeem met de *het-d* en *het-e* genen. Van elk van deze drie genen zijn verscheiden allelen bekend. Bij analyse van een deel van de beschikbare natuurlijke isolaten werden ten minste vijf nieuwe *het-c* allelen aangetroffen. Deze nieuwe allelen hadden meer non-synonieme dan synonieme substituties vooral in sequenties betreffende de buitenkant van het HET-C eiwit. Hieruit kan worden afgeleid dat het *het-c* gene vermoedelijk onder positieve selectie staat.

List of publications

Peer-reviewed scientific papers

Herman J. Pel, Johannes H. de Winde, David B. Archer, Paul S. Dyer, Gerald Hofmann, Peter J. Schaap, Geoffrey Turner, Ronald P. de Vries, Richard Albang, Kaj Albermann, Mikael R. Andersen, Jannick D. Bendtsen, Jacques A.E. Benen, Marco van den Berg, Stefaan Breestraat, Mark X. Caddick, Roland Contreras, Michael Cornell, Pedro M., Coutinho, Etienne G.J. Danchin, Alfons J.M. Debets, Peter Dekker, Piet W.M. van Dijck, Alard van Dijk, Lubbert Dijkhuizen, Arnold J.M. Driessen, Christophe d'Enfert, Steven Geysens, Coenie Goosen,Gert S.P. Groot, Piet W.J. de Groot, Thomas Guillemette, Bernard Henrissat, Marga Herweijer, Johannes P.T.W. van den Hombergh, Cees A. M. J.J. van den Hondel, Rene T.J.M.van der Heijden, Rachel M. van der Kaaij, Frans M. Klis, Harrie J. Kools, Christian P. Kubicek, Patricia A. van Kuyk, Jürgen Lauber, Xin Lu, Marc J.E.C. van der Maarel, Rogier Meulenberg, Hildegard Menke, A. Martin Mortimer, Jens Nielsen, Stephen G. Oliver, Maurien Olsthoorn, **Karoly Pal**, Noël N.M.E. van Peij, Arthur F.J. Ram, Ursula Rinas, Johannes A. Roubos, Cees M.J. Sagt, Monika Schmoll, Jibin Sun, David Ussery, Janos Varga, Wouter Vervecken, Peter J.I. van de Vondervoort, Holger Wedler, Han A.B. Wösten, An-Ping Zeng, Albert J.J. van Ooyen, Jaap Visser and Hein Stam, 2007. Genome sequence of Aspergillus niger strain CBS 513.88: a versatile cell factory. *Nature Biotechology* 25, 221-231.

János Varga, Krisztina Rigó, Sándor Kocsubé, **Károly Pál**, Beáta Tóth, Robert A. Samson, Zofia Kozakiewicz, 2006. Evolutionary relationships among economically important species of *Aspergillus* subgenera *Aspergillus* and *Fumigati*. Plant Genome series Vol 2 (Biodiversity and Evolution) Part B

Kesseru P, Kiss I, Bihari Z, **Pal K**, Portoro P, Polyak B, 2005. Nitrate-dependent salicylate degradation by *Pseudomonas butanovora* under anaerobic conditions. Bioresource Technology 96 (7) 779-794

Varga J, Rigo K, Kocsube S, Farkas B, **Pal K**, 2003. Diversity of polyketide synthase gene sequences in *Aspergillus* species. Research in Microbiology 154 (8) 593-600

Manuscripts

Károly Pál, Anne D. van Diepeningen, János Varga, Alfons J.M. Debets and Rolf F. Hoekstra, 2007. Differences in heterokaryon incompatibility (*het*) genes within *Aspergillus niger* and between species. Manuscript.

Károly Pál, Anne D. van Diepeningen, Bertha Koopmanschap, Rolf F. Hoekstra and Alfons J.M. Debets, 2007. The *het-c* gene in *Aspergillus niger*. Manuscript.

Károly Pál, Anne D. van Diepeningen, A. Bertha Koopmanschap, Rolf F. Hoekstra, and Alfons J.M. Debets, 2007. Analysis of *hch* and *het-c* in natural isolates of *Podospora anserina*. Manuscript.

Károly Pál, Sándor Kocsubé, Anne D. van Diepeningen, János Varga, Alfons J.M. Debets and Rolf F. Hoekstra, 2007. Sexual genes in the asexual filamentous fungus *Aspergillus niger* and relatives. Manuscript.

Oral presentations

13-15 April 2003 Siófok (Hungary): 5th Hungarian Genetic Congress

19-23 March 2005 Spring School in Wageningen(the Netherlands): Chemical communication: from gene to ecosystem.

12-13 January 2006 Noordwijkerhout (the Netherlands): DSM workshop on *Aspergillus niger*.

11-14 April 2007 Utrecht (the Netherlands): International workshop 'Aspergillus systematics in the genomics era'.

Poster presentations

8th European Conference on Fungal Genetics (ECFG) Wien (Austria), April 8-11, 2006: Vegetative incompatibility in *Aspergillus niger* (improved version) (**Károly Pál**, János Varga, Alfons J.M. Debets, Rolf Hoekstra)

10th congress of the European Society for Evolutionary Biology (ESEB) Krakow (Poland), August 15-20, 2005: Vegetative incompatibility in *Aspergillus niger* (**Károly Pál**, János Varga, Alfons J.M. Debets, Rolf Hoekstra)

Acknowledgements

In June 2001 I graduated as a biologist with a specialization in biotechnology and microbiology. After spending 23 years in Hungary I planned to study abroad but first I had to find a laboratory in my field. This dream of mine was realized by János Varga who later became my Hungarian supervisor at the Department of Microbiology in Szeged. In September 2001 I made my first trip to Wageningen and met my new group. Rolf Hoekstra, my promoter, was the first who I met from the Laboratory of Genetics; he took care of me since then and helped during the years of my PhD education. Dear Rolf, thank you for all the hospitality and kindness I got during my study and for the opportunity to be part of this research group, which became my second family. Your door was always opened when I faced problems and you always found a way to help me. As one of my co-promoters I was lucky to have Fons Debets, since he was always willing to help me; in Hungary János, my other co-promoter, was the person I could always find with my questions. Fons and János, I would like to thank you for your valuable advices and excellent scientific guidance. Whenever I got stuck with my research you always found a solution for my problems.

Maybe I would still work on this thesis without the help of Anne van Diepeningen. She came to my aid last year and did a wonderful job. Anne, I hardly can find words to thank you for the enormous help I got from you while working on my thesis.

Manual work in the laboratory was a prerequisite for getting publishable results. Marijke Slakhorst, Bertha Koopmanschap and Edu Holub were the angels around me in the fungal, molecular and – in case of Edu – both laboratories. Marijke taught how to do clean and precise work, thank you for that. Warm thanks to Bertha because I could always count on her when I needed some explanation on molecular techniques. Probably, I had the funniest protoplast transformations in my life with the assistance of Edu, who is both a cheerful person and a walking database at the same time, keeping all details of materials and methods in his head. He stored my stuff during my stays in Hungary and transported me with all my garbage when I moved back to Wageningen again. Thanks enormously, my old friend.

Thanks to Klaas Swart, who was my first Dutch teacher after arriving to Wageningen and gave me the practical basis for my work done in the Netherlands. Arjan de Visser supported me with statistical help which I needed to support some theories.

There is no institute without people who take care of finance and other official things. Many thanks to Aafke van der Kooi and Corrie Eekelder who were always very kind and smiling and did their best whenever I turned to them with any problems.

No one likes to work in unfriendly surroundings. I was in the lucky situation that all my ‘colleagues’ (also graduate students) were exceptionally nice persons. Judith Bruggeman, Siemen Schoustra, Marc Maas and Henk Dalstra graduated before me but I have to mention them, since they all helped me to adapt easily to the new surroundings. My best partner in debates, squash and Risk, and one of my closest friends was Michelle Habets, who I liked for her kindness, honesty and personality. Stefan Bosmans is a good example of a graduate student with never-ending optimism and enthusiasm, hopefully I could learn something from you. Thanks to Tamás Czárán, Daniel Rozen, Dominika Wloch and Anna Visser for the Risk parties, chats and barbecues.

Extra thanks to Merijn Salverda, the best partner in the lonely, dark weeks and months, who was always there when I needed someone to chat or have a beer at the Rijn with. It is not easy to find such a good friend like him, I count myself fortunate to know this exceptional guy.

I have to express my gratitude to Rob Samson and the community of Centraalbureau voor Schimmelcultures for the beautiful pictures that they provided me with.

Wageningen always had its own Hungarian community whose members founded WaMBaK (Wageningeni Magyarok Baráti Köre – The Hungarian Club of Wageningen). As a ‘wambaker’ I had countless wonderful trips and fun with these Hungarians. Thanks (in alphabetical order) to Árpád

Bokor, Erika Brath, Krisztina Gábor (thanks for technical help), Szilvia Gondár, Ditta Greguss, Zsófia Kádár (who is a really great friend), Oxana Kobzár, Tamás Máhr, Mónika Márkó, Annamária Nagy, Tibor Palugyai, Dénes Somogyi, Attila Stingli, Katalin Szalontay, Orsolya Tóth and Magdolna Trenovszki.
Many thanks to Petra Lengyel, Tamás Griff and Bernardo Suarez Cretton for their hospitality in the T42. Thanks to Dóra Szinay for the long chats and also thanks to Olga Elisseeva and Angelica for the dinners and nice talks.
During the Wageningen-years friends were coming and going but there was someone, who was permanently there. She became not only a friend but a family to me. It did not matter whether I was in a good mood or a bad one, Szvetlána was always around when I needed a friend, to share my joy or sadness. It is hard to appreciate her help as much as I would like to, at this time I can only say thank you.
Though our friendship did not start ideally, later on Anita Kurunczi became one of my best friends. Thanks for the healing mails and chats!
Since I have spent part of my PhD years in Hungary at the Department of Microbiology in the University of Szeged it is time to express my gratefulness to my Hungarian colleagues and friends who helped me in these hard years. Many thanks to Csaba Vágvölgyi, the head of the department, for allowing me to work at the department after my official PhD term ended. A nice working environment was ensured by László Galgóczy, Sándor Kocsubé, Beáta Linka, Gyöngyi Lukácsi, Ildikó Nyilasi, Zsanett Péteri and Miklós Takó.
Here I would like to thank my former colleagues and supervisors in the Bay Zoltán Foundation for Applied Research Insitute for Biotechnology in Szeged who sent me on my way: Péter Kesserű, Zoltán Bihari, István Kiss and Béla Polyák, Attila Szvetnik and Ágnes Finta.
I had a cheerful semester in Szeged with "Charlie's Angels", namely Orsolya Zöld and Anikó Göblös.
I have to thank to my 'landlords' and 'landlady' in Szeged: Balázs (Gyula) Fehér, Anita Lózsa and Gergő Maróti supported me with a place where I could stay whenever I stayed in Szeged. I am happy to have friends like them. I have to say thank you to Csaba Botos, a great friend of mine for more than 20 years, for the long chats and (sometimes acrobatic) bike tours. Also many thanks to János and Tamás Nagy for technical support and nice bike tours.
We can learn a lot from books, but some lessons in life are only taught by special persons. I believe these people appeared in my life when I had to meet them. They gave me so much that sometimes it took me years to realize and understand what they did teach to me about emotions, friendship and myself… Judit, Blanka and my travel mate on the bumpy road of life, thank you very much.
And last but not least I want to express my gratitude to my family, without whom I could never have finished my studies. Kedves Anyu, Apu és Laci! Köszönöm, hogy mindvégig mellettem álltatok és minden segítséget megadtatok, hogy elérhessem céljaimat! Nélkületek nem sikerült volna eljutnom idáig, és ezért örökké hálás leszek Nektek. ☺
Finally, many thanks to those ones who I could not mention here (due to lack of space) but contributed to my success.

Thanks to all of you!

Károly Pál

Curriculum Vitae

Károly Pál was born on July 16th, 1978 in Tiszafüred, Hungary. In 1996 he finished his secondary education at the Géza Gárdonyi Secondary School in Eger. In the same year he began his biology studies at University of Szeged (former József Attila University) in Szeged at the Faculty of Natural Sciences. He graduated in 2001 and got his MSc diploma in Biotechnology and Microbiology. From September 2001 he began his study as a PhD student in a Sandwich PhD program at Wageningen University and University of Szeged. The focus of his PhD project was on Comparative Analysis of Genetic Incompatibility in *Aspergillus niger* and *Podospora anserina*. During his PhD he followed the PhD education program in the C.T. de Wit Graduate School of Wageningen University.

Publisher

Grafisch Bedrijf Ponsen & Looijen BV, Wageningen

Cover design

Károly Pál