WESTERDIJK FUNGAL BIO DIVERSITY INSTITUTE

# Duplications and losses of genes encoding known elements of the stress defence system of the Aspergilli contribute to the evolution of these filamentous fungi but do not directly influence their environmental stress tolerance

T. Emri[1*], K. Antal[2], R. Riley[3,9], Z. Karányi[4], M. Miskei[1,5], E. Orosz[1], S.E. Baker[6], A. Wiebenga[7,8], R.P. de Vries[7,8], and I. Pócsi[1]

[1]*Department of Biotechnology and Microbiology, Faculty of Science and Technology, University of Debrecen, Egyetem tér 1, H-4032, Debrecen, Hungary;* [2]*Department of Zoology, Faculty of Sciences, Eszterházy Károly University, Eszterházy tér 1., H-3300, Eger, Hungary;* [3]*US Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA;* [4]*Department of Medicine, Faculty of Medicine, University of Debrecen, Nagyerdei krt. 98, H-4032, Debrecen, Hungary;* [5]*MTA-DE Momentum, Laboratory of Protein Dynamics, Department of Biochemistry and Molecular Biology, University of Debrecen, Nagyerdei krt. 98, H-4032, Debrecen, Hungary;* [6]*Environmental Molecular Sciences Division, Earth and Biological Sciences, Pacific Northwest National Laboratory, Richland, Washington, 99352, USA;* [7]*Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT, Utrecht, the Netherlands;* [8]*Fungal Molecular Physiology, Utrecht University, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands*

**Correspondence*: T. Emri, emri.tamas@science.unideb.hu

[9]Current address: Amyris, Inc., 5885 Hollis St Ste. 100, Emeryville, CA 94608, USA.

**Abstract:** The contribution of stress protein duplication and deletion events to the evolution of the Aspergilli was studied. We performed a large-scale homology analysis of stress proteins and generated and analysed three stress defence system models based on *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and *Aspergillus nidulans*. Although both yeast-based and *A. nidulans*-based models were suitable to trace evolutionary changes, the *A. nidulans*-based model performed better in mapping stress protein radiations. The strong Mantel correlation found between the positions of species in the phylogenetic tree on the one hand and either in the *A. nidulans*-based or *S. cerevisiae*-based models on the other hand demonstrated that stress protein expansions and reductions contributed significantly to the evolution of the Aspergilli. Interestingly, stress tolerance attributes correlated well with the number of orthologs only for a few stress proteins. Notable examples are Ftr1 iron permease and Fet3 ferro-$O_2$-oxidoreductase, elements of the reductive iron assimilation pathway, in the *S. cerevisiae*-based model, as well as MpkC, a HogA-like mitogen activated protein kinase in the *A. nidulans*-based model. In the case of the iron assimilation proteins, the number of orthologs showed a positive correlation with $H_2O_2$-induced stress tolerance while the number of MpkC orthologs correlated positively with Congo Red induced cell wall stress, sorbitol induced osmotic stress and $H_2O_2$ induced oxidative stress tolerances. For most stress proteins, changes in the number of orthologs did not correlate well with any stress tolerance attributes. As a consequence, stress tolerance patterns of the studied Aspergilli did not correlate with either the sets of stress response proteins in general or with the phylogeny of the species studied. These observations suggest that stress protein duplication and deletion events significantly contributed to the evolution of stress tolerance attributes of Aspergilli. In contrast, there are other processes, which may counterbalance the effects of stress gene duplications or deletions including (i) alterations in the structures of stress proteins leading to changes in their biological activities, (ii) varying biosynthesis of stress proteins, (iii) rewiring stress response regulatory networks or even (iv) acquiring new stress response genes by horizontal gene transfer. All these multilevel changes are indispensable for the successful adaptation of filamentous fungi to altering environmental conditions, especially when these organisms are entering new ecological niches.

**Key words:** *Aspergillus* phylogeny, Environmental stress, Evolution of the Aspergilli, Fungal stress defence system, Gene deletion, Gene duplication, Stress protein radiation.

Available online 11 October 2018; https://doi.org/10.1016/j.simyco.2018.10.003.

## INTRODUCTION

The Kingdom of Fungi is a large and diversified taxon with an estimated 2.2–3.8 million species (Lücking & Hawksworth 2018) occupying a breadth of ecological niches. Extensive fungal genome sequencing has led the construction of MycoCosm, a fungal genomics portal (https://genome.jgi.doe.gov/programs/fungi/index.jsf), which allows mycologists to gain a deeper and unique insight into the evolution of these organisms as new genome sequences continue to fill gaps in the Fungal Tree of Life (Grigoriev *et al.* 2014). These comparative genomics research projects are fuelled by the fact that the role of fungi in future bioeconomy including fermentation industry, biorefineries and agriculture cannot be overestimated (Baker *et al.* 2008, Grigoriev *et al.* 2011, Martin *et al.* 2011, Lange 2014, Meyer *et al.* 2016).

Among filamentous fungi, the ascomycetous genus *Aspergillus* includes several hundreds of cosmopolitan asexual species with world-wide distribution. Although these fungi seem to occupy various soil habitats with preference (*e.g.* the black Aspergilli *A. aculeatus*, *A. brasiliensis*, *A. niger*; Supplementary Table S1; Samson *et al.* 2007) some *Aspergillus* species are also well-known opportunistic colonisers of animals or even humans (*e.g. A. fumigatus*, *A. flavus*, *A. niger*, *A. terreus*; Sugui *et al.* 2014), and some others are indispensable production hosts for a wide spectrum of industrial fermentation and biotechnological processes (*e.g. A. niger*, *A. oryzae*, *A. terreus*; Park *et al.* 2017). Most Aspergilli have outstanding capabilities for biomass deconstruction with high efficiency due to their unique hydrolytic enzyme repertoire (*e.g. A. aculeatus*, *A. niger*, *A. oryzae*, *A. tubingensis*; Benoit *et al.* 2015, Park *et al.* 2017, Souza Guimarães & da Costa Souza 2017). Additionally, these fungi

Peer review under responsibility of Westerdijk Fungal Biodiversity Institute.




Studies in Mycology

are also known to spoil corn, fruits as well as animal feed causing significant economic losses (*e.g. A. carbonarius*, *A. flavus*, *A. niger*; Perrone & Gallo 2016). The Aspergilli are ubiquitously present in indoor environments causing deterioration of artworks and also versatile health complications like asthma (*e.g. A. clavatus*, *A. fumigatus*, *A. niger*, *A. versicolor*; Egbuta *et al.* 2017, Mallo *et al.* 2017).

Not surprisingly, *Aspergillus* spp. have remarkable oxidative, osmotic, heavy metal and cell wall integrity stress tolerances, which help to explain the plethora of ecological niches these fungi occupy (de Vries *et al.* 2017, Orosz *et al.* 2018). As demonstrated by Orosz *et al.* (2018), some species grow remarkably fast at 37 °C (*A. fisheri*, *A. acidus* and *A. nidulans*) while others can tolerate osmolytes added at high concentrations or are even osmophilic in the presence of either non-ionic (sorbitol; *A. glaucus*, *A. wentii*, *A. versicolor*, *A. oryzae*) or ionic (NaCl; *A. glaucus*, *A. sydowii*, A. versicolor, *A. wentii*) osmolytes (Supplementary Table S2). Other species are surprisingly tolerant to other types of deleterious environmental stress like oxidative stress (*A. nidulans*, *A. niger*, *A. oryzae* to $H_2O_2$; *A. brasiliensis*, *A. aculeatus* to menadione sodium bisulfite), heavy metal stress (*A. sydowii*, *A. fumigatus*, *A. terreus*, *A. versicolor* and *A. wentii* to $CdCl_2$) and cell wall integrity stress (*A. niger* and *A. glaucus* to Congo Red) (de Vries *et al.* 2017, Orosz *et al.* 2018).

Previous work sheds light on the importance of both segmental and whole genome gene duplication events in the evolution of fungi (Wapinski *et al.* 2007). Gene duplications are important elements of evolutionary adaptation processes (Ames *et al.* 2010) and the duplicants produced by these events may undergo neofunctionalisation or subfunctionalisation processes (Levasseur & Pontarotti 2011) to avoid the disadvantageous consequences of increased and imbalanced gene dosages (Papp *et al.* 2003).

Gene duplication, diversification and differential gene loss processes also contributed significantly to the evolution of opportunistic human pathogenic fungi such as *A. fumigatus* (Fedorova *et al.* 2008). The rapid expansion and evolution of certain gene families functioning in the invasion of the host organisms by fungi typically takes place in genomic islands located at sub-telomeric regions, and which are also known as “gene factories” or “gene dumps” (Fedorova *et al.* 2008). Expansion of protein families, *e.g.* cell surface proteins and hydrolytic enzymes, was also reported in the near-obligate nematode endoparasitic fungus *Drechmeria coniospora* with the concomitant increase in the number of the orthologs of the *S. pombe* Mak1/2/3-type oxidative stress sensor kinases and also in that of the *A. nidulans* HogA-type mitogen activated protein kinases (MAPKs; Zhang *et al.* 2016). Importantly, the number of stress sensor proteins and stress response-related transcriptional regulators decreased, which indicated certain simplifications in the stress defence system of this endoparasite (Zhang *et al.* 2016). While core elements of stress signalling pathways seem to be evolutionarily well-conserved in fungi in general, upstream stress sensor proteins and down-stream transcriptional regulators evolve rapidly presumably as a way for these eukaryotes to tailor and fine-tune their stress defence systems for an ecological niche (Nikolaou *et al.* 2009).

Previously we collected and classified a large group of fungal stress response proteins with verified physiological functions, in order to generate the Fungal Stress Response Database version 2 (Karányi *et al.* 2013, Zhang *et al.* 2016, de Vries *et al.* 2017, http://internal.med.unideb.hu/fsrd2/default.aspx?p=consortium). Moreover, the Fungal Stress Database was also set up by us, and currently incorporates *Aspergillus* stress tolerance data recorded in a number of agar plate experiments performed under various types of stress conditions (oxidative stress, high-osmolarity stress, cell wall stress and heavy metal stress) as well as at different incubation temperatures (25 and 37 °C) (de Vries *et al.* 2017, Orosz *et al.* 2018; http://www.fung-stress.org/ ). Based on the plethora of fungal stress data accommodated mainly by these two databases, we set the following aims in this study: (i) To find any correlation between gene duplication, diversification and differential gene loss processes concerning stress response genes/proteins and the evolution of *Aspergillus* species. (ii) To assess whether evolutionary changes in the *Aspergillus* stress defence systems affect directly or indirectly the environmental stress tolerances of these important ascomycetes. (iii) To estimate the applicability of *Saccharomyces cerevisiae*-based, *Schizosaccharomyces pombe*-based and *Aspergillus nidulans*-based stress defence system models to describe the stress defence systems operating in the Aspergilli.

## MATERIALS AND METHODS

### Homology search and counting *Aspergillus* orthologs of fungal stress proteins

In stress protein homology search, our fungal stress protein (FSP) collection was utilised. Our FSP collection contains 2 150 proteins with known/verified physiological functions (Karányi *et al.* 2013; Zhang *et al.* 2016). The distribution of the proteins among fungal species was the following: *A. flavus*: 1; *A. fumigatus*: 83; *A. nidulans*: 145; *A. oryzae*: 13; *Candida glabrata*: 31; *C. neoformans*: 79; *F. graminearum*: 13; *F. oxysporum*: 14; *F. verticillioides*: 4; *N. crassa*: 78; *N. fischeri*: 2; *C. albicans*: 210; *S. cerevisiae*: 921; *S. pombe*: 534; *U. maydis*: 22 (see “Stress Database” in Supplementary Table S3). Within the scope of this study, the set of FSPs was manually curated, increasing the reliability of and making any background literature search easier in the “Stress Database”.

Homology searches were performed in the fully sequenced genomes of 25 *Aspergillus* and *Penicillium* strains representing 22 species (de Vries *et al.* 2017). The following species were included in the study: *A. aculeatus*, *A. brasiliensis*, *A. carbonarius*, *A. clavatus*, *A. fischeri*, *A. flavus*, *A. fumigatus*, *A. glaucus*, *A. kawachii*, *A. luchuensis*, *A. nidulans*, *A. niger* represented by three strains (CBS 113.46/ATCC 1015, CBS 513.88 and NRRL3), *A. oryzae*, *A. sydowii*, *A. terreus*, *A. tubingensis*, *A. versicolor*, *A. wentii*, *A. zonatus*, *Eurotium rubrum*, *P. chrysogenum*, *P. digitatum* and *P. rubens* (see “Stress Protein Orthologs” in Supplementary Table S3). After clicking on “Links to genome sequences” in Supplementary Table S3, the list of links to the appropriate genome sequence resources will appear.

In the identification and counting of stress homologs of FSPs, the protocol of Miskei *et al.* (2009) and Karányi *et al.* (2013) was used with modifications. Briefly, (i) the set of FSPs was blasted against the selected 25 *Aspergillus* and *Penicillium* species using Blastp (protein-protein BLAST), (ii) the set of potential homologs was reverse blasted *versus* the set of FSPs with Blastp, (iii) for each protein A from FSP, the list of $b_0$, $b_1$, ..., $b_N$ was gained from each of the 25 *Aspergillus* and *Penicillium* species, ranked by e-

value, (iv) for the lowest e-value protein (or multiple proteins if there are ties) $b_0$, the ranked list of hits $a_0$, a1, ..., $a_M$, was obtained for the full protein sets of each organism in the FSP set, and (v) if $a_i$ had best hit $b_j$, and $b_j$ had best hit $a_i$, then we considered them to be best bidirectional BLAST hits, and $a_i$ and $b_j$ stress protein pairs were treated as putative "orthologs" in subsequent analyses. Although we use the expression of "ortholog" for clarity it is worth emphasising that the applied method does not differentiate between orthologs and paralogs. The complete list of orthologs is found in Supplementary Table S3.

To simplify stress defence system modelling in the Aspergilli, three groups of stress proteins were selected with the following features: Group 1 (*Saccharomyces cerevisiae*-based model): *S. cerevisiae* stress proteins with functionally characterised ortholog(s) in at least in one more fungal (*e.g. Schizosaccharomyces*, *Candida*, *Neurospora*, *Aspergillus*, etc.) species. Group 2 (*Schizosaccharomyces pombe*-based model): *S. pombe* stress proteins with functionally characterised ortholog(s) in at least in one more fungal species. Group 3 (*Aspergillus nidulans*-based model): all *A. nidulans* stress proteins functionally characterised before finalising the construction of the set of FSPs in this study (see the "*S. cerevisiae*-based model", "*S. pombe*-based model" and "*A. nidulans*-based model" in Supplementary Table S3). In the case of the *A. nidulans*-based model, the relatively low number of known *A. nidulans* stress proteins did not allow us to make any further simplifications like we did in the yeast-based models.

In this study, we also collected stress protein orthologs for *S. pombe* and *S. cerevisiae* from the PomBase database (ftp://ftp.pombase.org/pombe/orthologs/cerevisiae-orthologs.txt). We collated this independently generated ortholog list by BLASTing *S. cerevisiae* proteins in *S. pombe* and *vice versa.* We regarded best bidirectional BLAST hits as confirmed orthologous stress protein pairs in this study but in addition, we also present PomBase orthologs for *S. cerevisiae* and *S. pombe* on worksheets "*S. cerevisiae*-based model" and "*S. pombe*-based model" without bidirectional BLAST confirmation.

Before starting deeper bioinformatic analyses, some more simplifications were done on both the yeast-based and *A. nidulans*-based models. At first, ortholog protein sets, which came up with *S. cerevisiae* or *S. pombe* paralog pairs in the Aspergilli (*e.g.* orthologs of Ptc3/Ptc2, Grx1/Grx2, Trr1/Trr2 in baker's yeast and those of Win1/Wis4, Pyp1/Pyp2, Ptc3/Ptc2, in fission yeast), were combined (see the *S. cerevisiae*-based and *S. pombe*-based models in Supplementary Table S4). Apparently highly expanding stress proteins with more than four putative orthologs in at least one *Aspergillus* species were excluded from further analyses. In these cases, it became highly uncertain whether or not all stress protein functions had been retained or diverged considerably in the corresponding paralogs. Moreover, these stress proteins with a high number of putative orthologs would have caused a very biased statistical analysis although the number of such proteins was relatively low. Similarly, stress proteins with exceptionally high Expect (E) value orthologs only (E typically higher than $10^{-10}$) were not analysed further.

The number of the studied *Aspergillus* species was limited to 18 in this part of the project (Supplementary Table S4). Please note that only data gained with the strain CBS 513.88 was processed for *A. niger*. The gene models for *A. carbonarius* have not been recently updated and were therefore excluded due to artefacts related to the use of an early generation gene modelling pipeline. Finally, all *Aspergillus* stress protein orthologs identified in the three models were manually checked and curated using the genome analysis tools available in the Aspergillus Genome Database (AspGD; http://www.aspgd.org/).

## Construction of phylogenetic tree

Phylogeny of the *Aspergillus* spp. included in the study of de Vries *et al.* (2017) was inferred from 149 conserved protein sequences. After clustering proteins using the Markov Cluster (MCL) algorithm (Enright *et al.* 2002), clusters with only one protein from each species were selected (149 such clusters in total). Protein sequences were aligned with MAFFT multiple sequence alignment software (Nakamura *et al.* 2018), trimmed to well-aligned regions using Gblocks (Castresana 2000) with default parameters. The phylogenetic tree was constructed with the RAxML phylogenetic analysis program (Stamatakis 2014) using the PROTMIXWAG model, *Batrachochytrium dendrobatidis* set as an outgroup, and 100 rounds of bootstrapping.

## Cluster analysis of stress tolerances

To estimate and present the differences in the stress tolerances of the studied *Aspergillus* species, growth data available in the Fungal Stress Database (FSD, Orosz *et al.* 2018; http://www.fungstress.org/) for the following strains in the presence of various stressors were processed and analysed: *A. aculeatus* (CBS 172.66), *A. brasiliensis* (CBS 101740), *A. carbonarius* (CBS 141172 = DTO 115-B6), *A. clavatus* (CBS 513.65 = NRRL1), *A. fischeri* (CBS 544.65), *A. flavus* (CBS 128202 = NRRL 3357), *A. fumigatus* (CBS 126847 = Af293), *A. glaucus* (CBS 516.65), *A. luchuensis* (CBS 106.47), *A. nidulans* (FGSCA4), *A. niger* represented by two strains (CBS 113.46 and N402), *A. oryzae* (Rib40), *A. sydowii* (CBS 593.65), *A. terreus* (NIH2624), *A. tubingensis* (CBS 134.48), *A. versicolor* (CBS 795.97), *A. wentii* (CBS 141173 = DTO 134-E9). Growth data recorded in cultures exposed to $H_2O_2$ (oxidative stress, increases intracellular peroxide concentration), menadione sodium bisulfite (MSB, oxidative stress, elevates intracellular superoxide level) and $CdCl_2$ (heavy metal stress) as functions of stressor concentrations were fit by second order polynomials to calculate $MIC_{50}$ and $MIC_{90}$ values. $MIC_{50}$ and $MIC_{90}$ were defined as the lowest concentration of a given stress initiating agent, which caused 50 % or 90 % decreases in colony growth, respectively. For sorbitol (non-ionic hyperosmotic stress), NaCl (ionic hyperosmotic stress) and Congo Red (cell wall integrity stress) treatments, relative growth values (% of those recorded in unstressed control cultures) measured at 2.0 mol/L, 1.0 mol/L and 108 µmol/L concentrations, respectively, were calculated and taken into consideration in further analyses. Please note that only growth datasets recorded on nitrate minimal medium at 25 °C (5 and 10 days of incubation) and at 37 °C (5 d of incubation) were chosen and processed for each species. Conidia were always produced on malt extract––mycological peptone sporulation agar medium (1.5 % agar, 25 °C in the dark, 6 days) and, in the case of the osmophilic fungus *A. glaucus*, all sporulation and culture media were supplemented with 1.0 mol/L NaCl.

After mathematical standardisation (with "scale" function of R project) of growth values recorded in unstressed cultures as well as $MIC_{50}$, $MIC_{90}$ and relative growth data calculated for stress exposed cultures, the Euclidean distances of the species were determined in a multidimensional space ("dist" function of R

project). These distances were used to perform cluster analysis and construct a dendrogram ("hclust" functions of R project) to demonstrate similarities and differences in the overall stress behaviours of the studied *Aspergillus* species. Multidimensional scale (MDS) plots were also generated to demonstrate the versatility of the stress tolerances of these Aspergilli (with "cmdscale" function of R project). Normalised data are available in Supplementary Table S5.

## Gene onthology term (GO) enrichment analysis

Significant enriched GO terms within a subset of the *S. cerevisiae*-based, *S. pombe*-based and *A. nidulans*-based models were identified with the SGD Gene Ontology Term Finder (*Saccharomyces* Genome Database; https://www.yeastgenome.org/), Generic Gene Onthology Term Finder of Princeton University (http://go.princeton.edu/cgi-bin/GOTermFinder) and AspGD Gene Ontology Term Finder (*Aspergillus* Genome Database; http://www.aspergillusgenome.org), respectively. Default settings were employed with the following exceptions: The *p* value was set to 0.05 as well as genes from the appropriate model were used as background gene set in each case.

## Determination of interacting partners of stress proteins

The String Database (https://string-db.org/) was used to count the number of interacting partners of stress proteins in the three, *S. cerevisiae*-based, *S. pombe*-based and *A. nidulans*-based stress defence system models. Only interactions determined experimentally and scored with at least 0.8 confidence values were taken into consideration.

## Gene duplication and gene deletion based multidimensional scaling and cluster analysis of the *Aspergillus* species

The numbers of ortholog genes were used to calculate Manhattan distances between the studied *Aspergillus* species in each stress defence system model. These distance matrices were used to perform cluster analysis and construct dendrograms using complete linkage. Distance matrices and dendrograms were constructed with the "dist" and "hclust" functions of the R project, respectively. To show similarities and differences between the *Aspergillus* species, distance matrices were also used to generate multidimensional scale (MDS) plots. MDS plots were created with the "cmdscale" function of the R project.

## Mantel test

Pairwise Mantel tests {implemented in the R package "ade4" by Dray & Dufour (2007)} were applied to rank-transformed versions of the distance matrices. The number of permutations for the tests of significance was 9 999. Holm corrected one-tailed p-values, associated with the Mantel correlations were calculated to test the alternative hypothesis that the Mantel correlation under test is positive. Only Mantel correlations with Holm adjusted $p < 0.05$ were regarded as significant positive correlation. The origins of the compared distance matrices were as follows:

- Manhattan distance matrices of the normalised stress-related physiological features of 16 *Aspergillus* species. (*A. carbonarius* and the strain *A. niger* N402 were omitted from these analyses.)
- Manhattan distance matrices of the *S. cerevisiae*-based, *S. pombe*-based and *A. nidulans*-based stress defence system models, which covered the same 16 *Aspergillus* species presented in the above mentioned stress physiology matrices.
- Cophenetic distance matrices originated from the dendrogram representing phylogenetic relationships of the 16 *Aspergillus* species (de Vries *et al.* 2017).

# RESULTS

Three stress protein groups were selected based on *S. cerevisiae*, *S. pombe* or *A. nidulans* stress proteins with experimentally verified functions (Supplementary Table S4). By definition, the physiological functions of selected *S. cerevisiae* and *S. pombe* stress proteins were also confirmed in at least one more fungal species through the functional characterisation of their appropriate orthologs. Because the number of already characterised *A. nidulans* stress proteins was dwarfed by known stress proteins in both *S. cerevisiae* and *S. pombe* this restriction was not employed for this species. The selected stress protein groups were regarded as elements of *S. cerevisiae*-based, *S. pombe*-based and *A. nidulans*-based stress defence system models. The uniqueness (specificity) of the stress models is summarised in Table 1. According to these data, the overlap between the two yeast-based models was much larger than those calculated between the *A. nidulans*-based model and any of the two yeast-based stress defence system models. The overlap between the *S. cerevisiae*-based and *S. pombe*-based models came from the fact that the stress defence systems of baker's yeast and fission yeast are by far the most extensively studied among fungi and, as a consequence, orthologous stress protein pairs are most often characterised in these two species (Supplementary Tables S3 and S4).

**Table 1.** Specificities of the *S. cerevisiae*-based, *S. pombe*-based and *A. nidulans*-based stress response system models.

| Compared models[1] | Number of proteins in the compared two models | | Number and per cent ratio of unique stress proteins in the | |
|---|---|---|---|---|
| | former model | latter model | former model | latter model |
| *S. cerevisiae* vs. *S. pombe* | 301 | 248 | 89 (30 %) | 39 (16 %) |
| *S. cerevisiae* vs. *A. nidulans* | 301 | 133 | 238 (79 %) | 70 (53 %) |
| *S. pombe* vs. *A. nidulans* | 248 | 133 | 195 (79 %) | 79 (59 %) |

[1] – Note that direct relationships between the models cannot be calculated because one protein present in a given stress defence system model may have more than one ortholog in the other models. Therefore, only the number of unique proteins present merely in one of the two compared stress defence models can be counted. By definition, a unique protein has no ortholog in the other analysed stress defence system model.

Orthologs of the selected *S. cerevisiae*, *S. pombe* and *A. nidulans* stress response proteins were identified in 18 *Aspergillus* species (Supplementary Table S3). The stress proteins in the stress defence system models were grouped into three subsets referred to as "deleted", "duplicated" and "conserved" proteins in accordance with the number of their orthologs found in the Aspergilli. These stress protein subgroups were defined as follows:

"Deleted" protein – a stress protein in the employed stress defence system model, which has no ortholog in at least one *Aspergillus* species. It can be the consequence of losing the gene encoding the protein in some species or gaining a gene in other species. Note that a "deleted" stress protein may arise as the consequence of either a gene deletion event or substantial changes in its gene sequence or even fusion with another gene, which made the identification of the orthologous stress protein impossible. It can also not be fully excluded that some genes are missing due to gaps in the genome sequence.

"Duplicated" protein – a stress defence system model protein, which has more than one ortholog in at least one *Aspergillus* species.

"Conserved" protein – a stress model protein, which has exactly one ortholog in each *Aspergillus* species.

It is important to note that genes coding for either "duplicated" proteins or "conserved" proteins may also have paralog(s) in the genomes of the model yeasts *S. cerevisiae* and *S. pombe*.

## Characterisation of the "deleted", "duplicated" and "conserved" stress proteins

Interestingly, the ratio of the "deleted" proteins (20–26 %) did not differ significantly in the three studied stress defence system models (Table 2). In contrast, the ratio of the "duplicated" proteins varied between 4 and 18 %, with the highest ratio of "duplicated" proteins observed in the *A. nidulans*-based model (Table 2).

Only few GO terms showed significant enrichment within the "deleted", "duplicated" and "conserved" protein subgroups (Table 3, Supplementary Table S6). It is noteworthy that in the case of the *S. cerevisiae*-based model the set of "deleted" proteins was enriched in "transcription factors", while the set of "duplicated" proteins was enriched in "ion transporters" and also in proteins involved in "carbohydrate metabolism" (Table 3). In the *S. pombe*-based and *A. nidulans*-based models, the GO terms "phagophore assembly site membrane" and "intracellular organelle" were significantly enriched in the "deleted" and "conserved" proteins, respectively.

Stress proteins with five or more known interacting partners were enriched within the "conserved" proteins subgroup in the *S. cerevisiae*-based model and were depleted within the "duplicated" and "deleted" proteins subgroups in the *S. cerevisiae*-based and *S. pombe*-based models, respectively (Table 4; Supplementary Fig. S1). This means that stress proteins located at network nodes tend to preserve their doses, *i.e.* they are

**Table 2.** Number and ratio of "conserved", "deleted" and "duplicated" proteins in the studied stress defence system models.

| Model | Number of proteins in the model | Number (percentage) of | | |
|---|---|---|---|---|
| | | "conserved" proteins | "deleted" proteins | "duplicated" proteins |
| *S. cerevisiae* | 301 | 197 (65 %)[1] | 78 (26 %)[1] | 29 (10 %)[1] |
| *S. pombe* | 248 | 184 (74 %)[2] | 54 (22 %)[1] | 11 (4 %)[2] |
| *A. nidulans* | 133 | 88 (66 %)[1,2] | 27 (20 %)[1] | 24 (18 %)[3] |

[1-3] – Ratios marked with the same superscript within a column are not significantly different, as indicated by the Fisher's exact test ($p < 0.05$).

**Table 3.** Enriched GO terms in the sets of "conserved", "deleted" and "duplicated" proteins.

| Model | Significantly ($p < 0.05$) enriched GO terms[1] in the set of | | |
|---|---|---|---|
| | "conserved" proteins | "deleted" proteins | "duplicated" proteins |
| *S. cerevisiae* | – | "transcription regulatory region DNA binding" | "ion transmembrane transporter activity", "carbohydrate metabolic process" |
| *S. pombe* | – | "phagophore assembly site membrane" | – |
| *A. nidulans* | "intracellular organelle" | – | – |

[1] - The full data set is available in Supplementary Table S6.

**Table 4.** Duplications and deletions of important stress proteins locating at the nodes of the stress defence system networks.

| Model | Number of proteins with at least 5 known interacting partners[1] in the protein sets of | | | |
|---|---|---|---|---|
| | "conserved" proteins | "deleted" proteins | "duplicated" proteins | whole stress defence systems |
| *S. cerevisiae* | 105[2] | 33[5] | 11[3] | 149 |
| *S. pombe* | 54[4] | 8[3] | 4 | 66 |
| *A. nidulans* | 16 | 5 | 4 | 25 |

[1] The number of interacting partners was identified using the String database (https://string-db.org/). Only interacting partners identified experimentally and scored with at least 0.8 confidence values were counted.
[2] Significant enrichment within the protein set (Fisher's exact test, $p < 0.05$).
[3] Significant depletion within the protein set (Fisher's exact test, $p < 0.05$).
[4] Enrichment within the protein set (Fisher's exact test, $0.05 < p < 0.10$).
[5] Depletion within the protein set (Fisher's exact test, $0.05 < p < 0.10$).

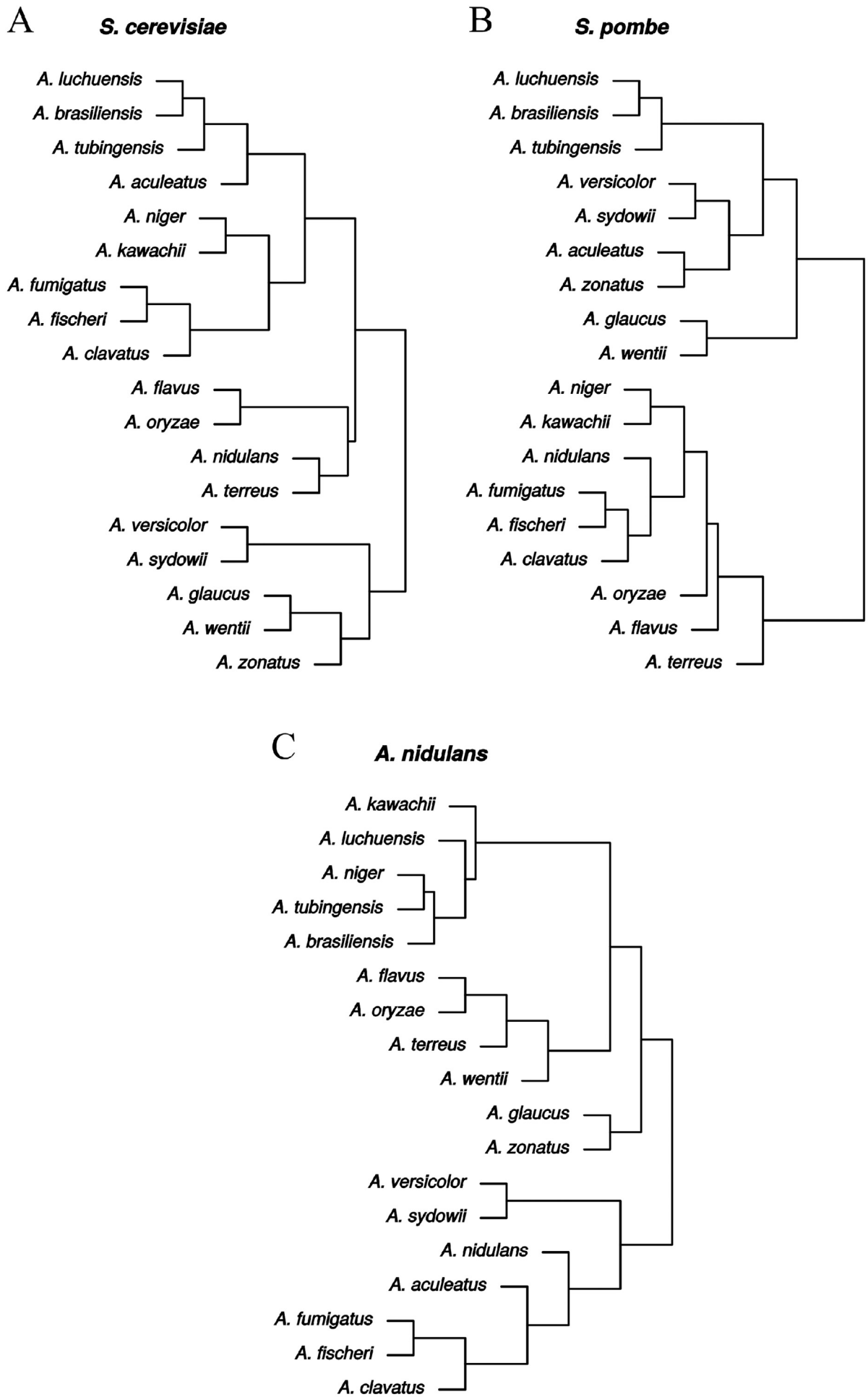


**Fig. 1.** Hierarchical cluster analysis of the *Aspergillus* species using complete linkage and Manhattan distances calculated from the features (numbers of ortholog genes) of the three stress defence system models. Part A: *S. cerevisiae*-based model, Part B: *S. pombe*-based model and Part C: *A. nidulans*-based model.

unlikely to go through deletions or duplications, which would jeopardise their distinguished and delicate positions in the stress response network.

## Characterisation of the evolution of stress proteins and stress tolerances in selected *Aspergillus* species

Hierarchical cluster analysis of the 18 studied Aspergilli was carried out to map the distances of the species, based on protein duplications and protein deletions. The similarities and differences between the species are visualised in Figs 1 and 2. Stress tolerances and phylogenetic relationships of the selected *Aspergillus* species were also summarised in dendrograms and, in the case of stress tolerances, in an MDS plot as well (Fig. 3).

Pairwise Mantel tests were applied to compare the results coming from the *S. cerevisiae*-based, *S. pombe*-based and *A. nidulans*-based stress response system models as well as to elucidate whether the stress protein deletions and duplications found in these models reflect phylogenetic relationships or physiological (stress tolerance) similarities between the studied *Aspergillus* species. The strongest Mantel correlation was observed between the two yeast models (0.74, $p = 0.001$) (Fig. 4), which can be explained well with the high overlap between the stress protein sets of these models (Table 1). A strong correlation was also found between the positions occupied in the phylogenetic tree and in the *A. nidulans*-based model (0.6, $p = 0.001$) (Fig. 4). Positions in the *S. cerevisiae*-based model correlated well with those in the *A. nidulans*-based model (0.57, $p = 0.001$) and in the phylogenetic tree (0.51, $p = 0.0014$) (Fig. 4). The positions of the Aspergilli in the dendrogram (Fig. 3B) constructed from the stress physiological data accommodated by FSD (http://www.fung-stress.org/) did not correlate at all with the observed positions of the same species in the yeast-based or *A. nidulans*-based stress response system models (Figs 1 and 2) or on the

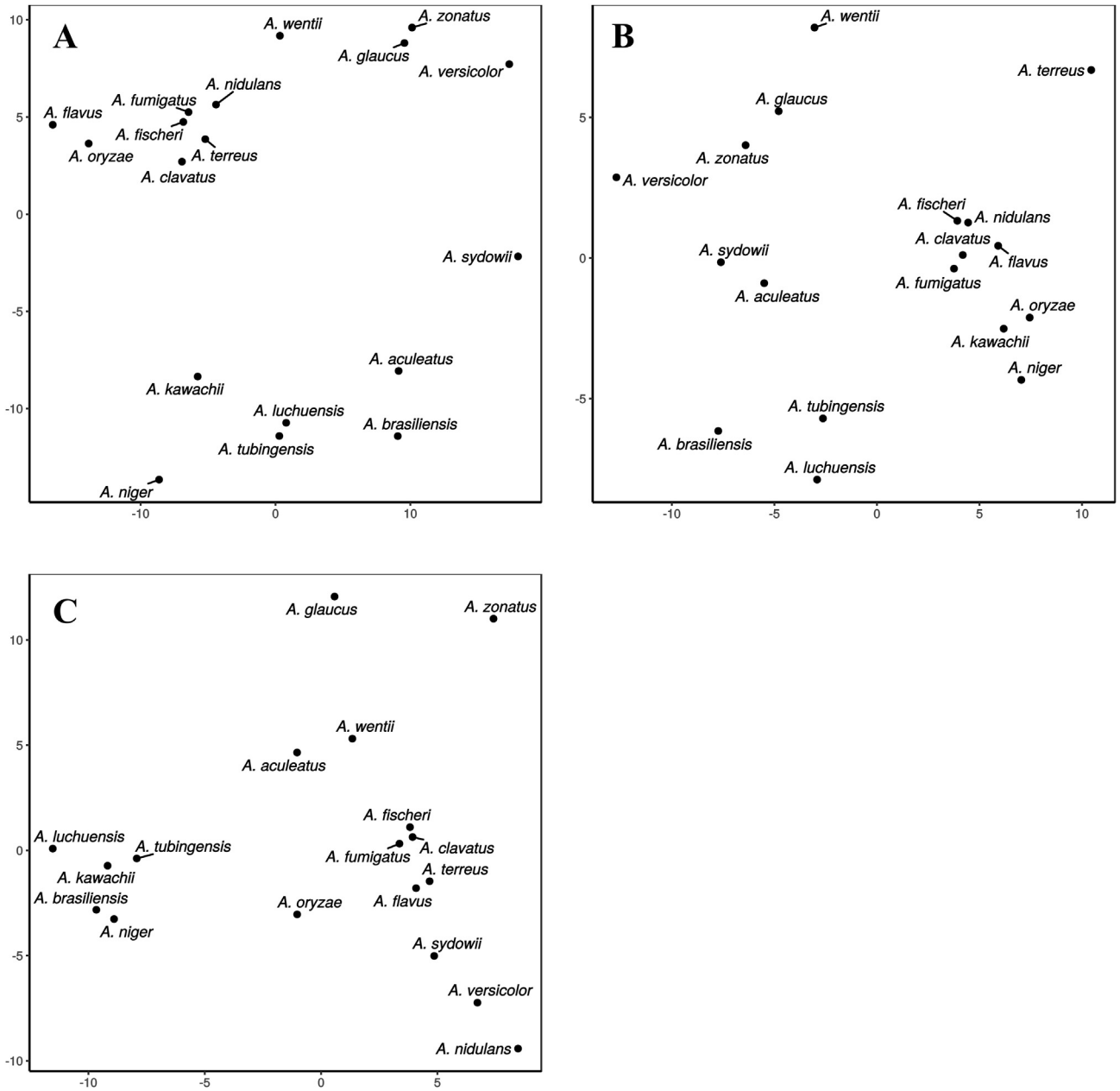


**Fig. 2.** Multidimensional scaling of Manhattan distances between *Aspergillus* species, calculated from the features (numbers of ortholog genes) of the three models. Part A: *S. cerevisiae*-based model, Part B: *S. pombe*-based model and Part C: *A. nidulans*-based model.

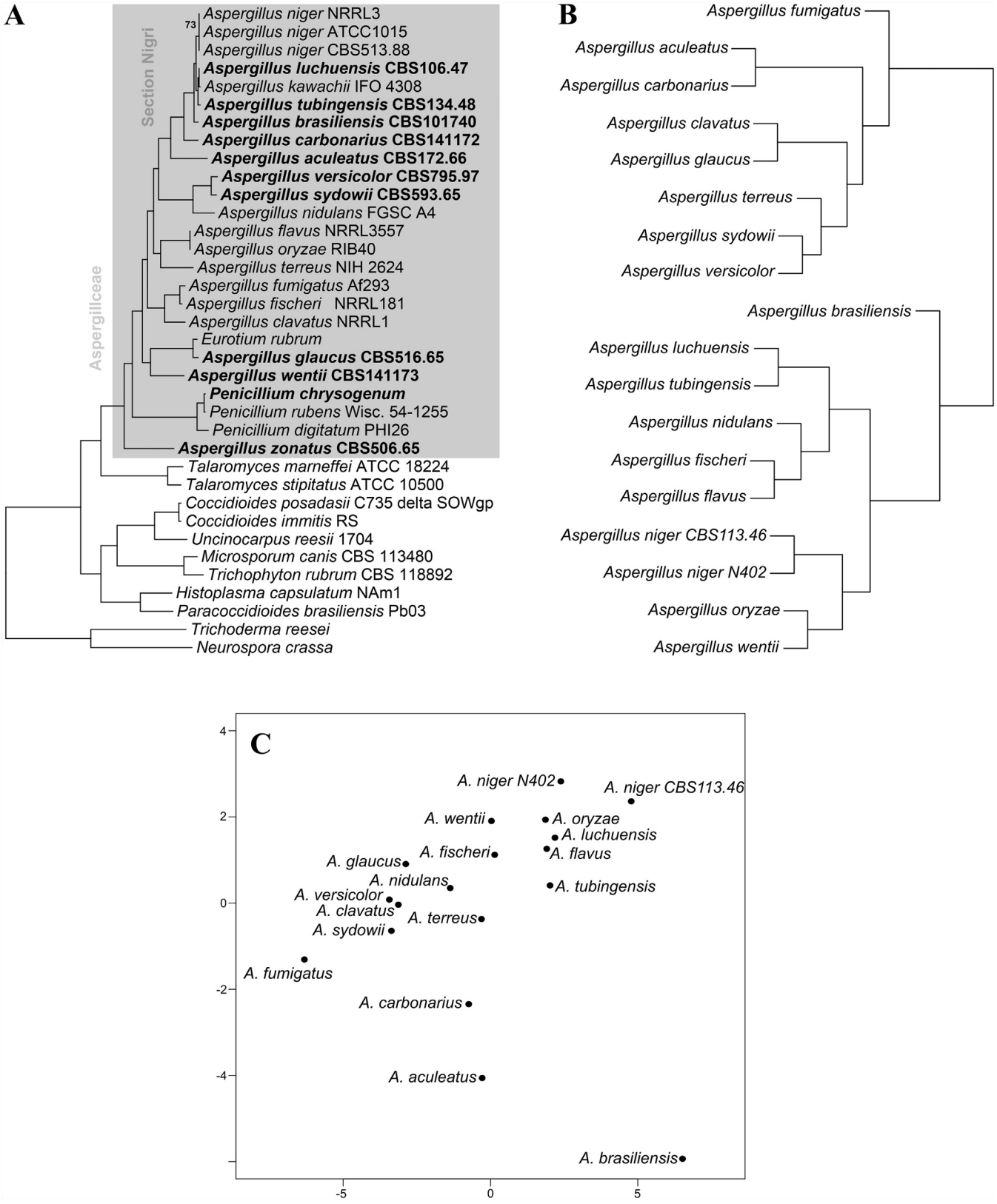


**Fig. 3.** Comparison of the phylogenetic positions and the stress tolerances of the Aspergilli. Part A: Maximum likelihood phylogeny of the *Aspergillus* spp. as published in the study of de Vries *et al.* (2017). The phylogenetic tree was deduced from 149 conserved protein sequences. Newly sequenced species are shown in **bold**. Note, *A. niger* ATCC 1015 is identical to CBS 113.46. Part B: Cluster analysis dendrogram constructed on the stress tolerance data reposited in the Fungal Stress Database (Orosz *et al.* 2018; **URL**: http://www.fung-stress.org/). Part C: Multidimensional scale plot presentation of the stress tolerance variability of the *Aspergillus* species tested (de Vries *et al.* 2017, Orosz *et al.* 2018).

phylogenetic tree in general (Figs 3 and 4). Nevertheless, some closely related species like *A. aculeatus* – *A. carbonarius*, *A. sydowii* – *A. versicolor*, *A. luchuensis* – *A. tubingensis* as well as the two *A. niger* strains (CBS 113.46 and N402) showed remarkably similar overall stress tolerances (Fig. 3B). On the other hand, some other species like *A. brasiliensis* and *A. fumigatus* showed highly anomalous stress tolerance patterns (Fig. 3B; de Vries *et al.* 2017, Orosz *et al.* 2018).

Importantly, Mantel correlations did not show any significant correlations between the positions occupied by the *Aspergillus*

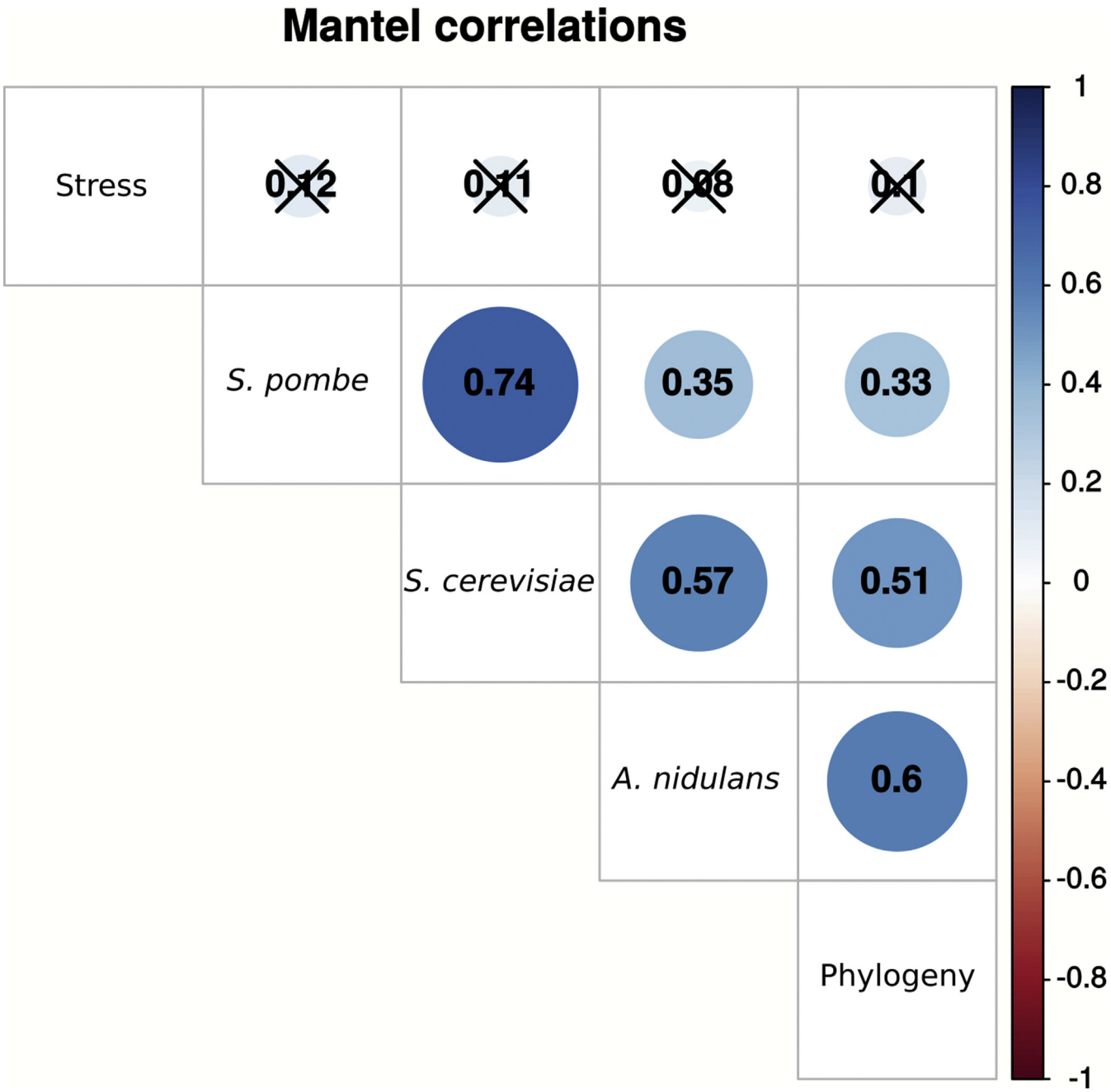


**Fig. 4.** Mantel correlations between the distance matrices of the studied *Aspergillus* species calculated on the *S. cerevisiae*-based, *S. pombe*-based and *A. nidulans*-based stress defence system models (Figs 1 and 2), distance matrix based on stress tolerance data ("Stress"; Fig. 3B) and cophenetic distance matrix of the phylogenetic tree ("Phylogeny"; Fig. 3A). Non-significant correlations ($p \geq 0.05$) are indicated by crosses (X).

spp. studied in the stress tolerance-based dendrogram and their positions either in the phylogenetic tree or in any of the yeast-based and *A. nidulans*-based stress response system models. Therefore, we examined individual stress proteins in the latter models to determine whether some of them would go along with any of the stress tolerance property of the studied *Aspergillus* spp. For most stress proteins we found that the copy number of genes coding for them did not show any strong correlation with any stress parameters tested (Supplementary Fig. S2, Supplementary Table S7). Those proteins which showed significant correlation with a certain stress condition generally also correlated well with other stress conditions as well (Supplementary Fig. S2, Supplementary Table S7). Remarkably, changes in the copy numbers of only few proteins could be linked specifically to one tested stress condition. For example, 10 proteins in the *A. nidulans*-based model and 17 proteins in the *S. cerevisiae*-based model formed groups, in which the copy numbers indicated a positive correlation with cadmium stress only (Supplementary Fig. S2, Supplementary Table S7). However, the correlation coefficients usually did not reach the value of 0.6. Out of these stress proteins, only Vma1 (locus ID S000002344, subunit of the vacuolar membrane ATPase) has been demonstrated to be involved in cadmium stress response (Ruotolo *et al.* 2008) so far, to the best of our knowledge. Considering the few stress proteins, which were linkable to certain stress conditions with high correlation coefficients (Supplementary Fig. S2, Supplementary Table S7) the following were especially interesting:

The number of

- Ftr1 {high affinity iron permease (Stearman *et al.* 1996); locus ID S000000947; *S. cerevisiae* model} orthologs showed positive correlation (correlation coefficients > 0.6,

*p*-values < 0.006) with sorbitol and $H_2O_2$ induced stress tolerance;
- Fet3 {ferro ion-$O_2$-oxidoreductase (Askwith *et al.* 1994); locus ID S000004662; *S. cerevisiae* model} orthologs showed positive correlation (correlation coefficients > 0.67, *p*-values < 0.004) with $H_2O_2$ induced stress tolerance at 25 °C;
- Gpp1 {glycerol-3-phosphate phosphatase (Norbeck *et al.* 1996); locus ID S000001315; *S. cerevisiae* model} orthologs correlated (correlation coefficients > 0.63, *p*-values < 0.009) with Congo Red induced cell wall stress tolerance or sorbitol induced osmotic stress tolerance at 25 °C;
- Ena1 {P-type ATPase sodium pump (Haro *et al.* 1991); locus ID S000002447; *S. cerevisiae* model} orthologs showed positive correlation (correlation coefficients > 0.66, *p*-values < 0.005) with $CdCl_2$ tolerance characterised with $MIC_{50}$ values;
- Dis2 {serine/threonine protein phosphatase PP1 (Ohkura *et al.* 1988, Grallert *et al.* 2015); locus ID SPBC776.02c; *S. pombe* model} orthologs positively correlated (correlation coefficients > 0.54, *p*-values < 0.031) with Congo Red induced cell wall stress tolerance;
- MpkC {putative HogA-like mitogen activated protein kinase (Furukawa *et al.* 2005, Jaimes-Arroyo *et al.* 2015, Bruder Nascimento *et al.* 2016, Pereira Silva *et al.* 2017); locus ID AN4668; *A. nidulans* model} orthologs correlated positively (correlation coefficients > 0.53, *p*-values < 0.036) with Congo Red induced cell wall stress tolerance or sorbitol induced osmotic stress tolerance and also showed a some positive correlation (correlation coefficients between 0.54 and 0.60, p < 0.03 with three out of the six attributes) with $H_2O_2$ induced stress tolerance;
- CatB {hyphal catalase (Kawasaki *et al.* 1997); locus ID AN9339; *A. nidulans* model} orthologs also showed positive correlation (correlation coefficients > 0.57, *p*-values < 0.021) with Congo Red induced cell wall stress tolerance and or sorbitol induced osmotic stress tolerance at 25 °C;
- NikA {putative histidine-specific protein kinase (Hagiwara *et al.* 2007); locus ID AN4479; *A. nidulans* model} orthologs correlated (correlation coefficients > 0.53, *p*-values < 0.034) with $H_2O_2$ stress tolerance detected at 25 °C.

Among the above-mentioned proteins, the involvement of MpkC in the cell wall and osmotic stress responses of *A. fumigatus* (Pereira Silva *et al.* 2017) and the contribution of NikA to the regulation of oxidative stress response in *A. nidulans* (Hayashi *et al.* 2014) have been experimentally verified. In the case of *A. fumigatus frtA* (an *frt1* ortholog) and *fetC* (an *fet3* ortholog) Kurucz *et al.* (2018a) found that a combined iron starvation – oxidative stress treatment up-regulated both genes markedly, however, iron starvation alone had no significant effect on the expression of these genes. These data suggest the potential involvement of FrtA and FetC (proteins in the reductive iron assimilation pathway) in oxidative stress response under iron deprivation. Skamnioti *et al.* (2007) demonstrated that CatB of *Magnaporthe grisea* (an ortholog of *A. nidulans* CatB) contributed to strengthening the cell wall during cell wall stress. Interestingly, similar physiological function for CatB orthologs has not been described in the Aspergilli.

It is worth noting that due to the huge number of investigated proteins, the Holm-corrected *p* value of the correlation coefficients were always higher than 0.05 with the exception of the copy number of AN0554 (*aldA*; aldehyde dehydrogenase gene) orthologs which correlated significantly (correlation coefficient: −0,87; $p = 0.0129$) with the normalised relative growth values of Congo Red treated cultures (37 °C 5 d) (Supplementary Table S7).

## DISCUSSION

Gene duplications – originating from either segmental or whole genome duplications – are common events in fungi and represent one of the most effective evolutionary driving forces operating in these wide-spread eukaryotes (Wapinski *et al.* 2007, Ames *et al.* 2010, Levasseur & Pontarotti 2011). Of course, imbalanced gene doses can be harmful for the organism (Papp *et al.* 2003) but neofunctionalisation or subfunctionalisation processes may stabilise duplicated genes (Shertz *et al.* 2010, Giacometti *et al.* 2011, Levasseur & Pontarotti 2011). The radiation of certain gene families likely plays a role in how fungi change life styles, *e.g.* towards parasitism (Fedorova *et al.* 2008, Jackson *et al.* 2009, Moran *et al.* 2011, Gabaldón *et al.* 2013, Maguire *et al.* 2013, Zhang *et al.* 2016). Comparative genomics is an outstanding tool to track evolutionary changes proceeding *via* gene duplication events (Fedorova *et al.* 2008, Jackson *et al.* 2009, Moran *et al.* 2011, Gabaldón *et al.* 2013). Erosion, pseudogenisation and eventual loss of superfluous gene copies are also among the predominant processes reshaping fungal genomes (Fedorova *et al.* 2008, Jackson *et al.* 2009, Shertz *et al.* 2010).

Considering the Aspergilli, an exceptionally important group of filamentous ascomycetous fungi (de Vries & Visser 2001, Sugui *et al.* 2014, Perrone & Gallo 2016, Egbuta *et al.* 2017, Mallo *et al.* 2017, Park *et al.* 2017, Souza Guimarães & da Costa Souza 2017), gene duplication and gene deletion events have contributed significantly to the evolution of this taxon (Table 2, Supplementary Tables S3 and S4). Importantly, the positions occupied by the Aspergilli on dendrograms and MDS plots constructed taking into consideration stress gene duplications and stress gene deletions (Figs 1 and 2) correlated well with the positions of the same species on the phylogenetic tree deduced from 149 conserved protein sequences (Figs 3 and 4). This correlation demonstrates the importance of the radiation and reduction of stress protein families in the evolution of the Aspergilli. Similar results were found by Zhang *et al.* (2016) for the nematode endoparasitic fungus *D. coniospora*. Moreover, some recent laboratory evolution experiments have also demonstrated the importance of gene amplification and gene deletion events in stress adaptation processes. In the case of the baker's yeast *S. cerevisiae*, amplification of *hxt5* and *hxt6* genes (encoding high affinity hexose transporters) and *sul1* (coding for a high affinity sulfate permease) was observed under glucose- and sulfate-limited conditions, respectively (Gresham *et al.* 2008). Meanwhile both amplification and complete deletion of the *pho5* (extracellular acid phosphatase) gene were reported in phosphate limited cultures (Gresham *et al.* 2008).

The number of duplicated *Aspergillus* stress proteins revealed by *S. cerevisiae*-based, *S. pombe*-based and *A. nidulans*-based stress defence system models varied significantly (10, 4 and 18 % of total stress proteins analysed; Table 2), which was attributed to two factors: (i) The *S. cerevisiae*-based and *S. pombe*-based models contained more conserved stress proteins, because only stress proteins with functional ortholog(s) in at least one additional fungal species were taken into

consideration in setting up the models (Supplementary Tables S3 and S4). (ii) Paralog genes (proteins) are quite difficult to identify using evolutionarily distant model species like *S. cerevisiae* or *S. pombe* (Wang *et al.* 2009), especially when one of the paralogs has been eroded *e.g.* in "gene dumps" (Fedorova *et al.* 2008). This observation strongly supports the urgent need for good-quality *A. nidulans*-based models in future evolutionary studies, which will be carried out in filamentous ascomycete taxa (Miskei *et al.* 2009).

Considering the evolution of *Aspergillus* stress proteins, conserved proteins located at network nodes with at least 5 known interacting partners in the *S. cerevisiae*-based model (Table 4, Supplementary Fig. S1) appear more recalcitrant to either duplications or deletions. This observation seems reasonable as deletions would result in serious functional disorders in any kind of environmental stress defence system while duplications could significantly disturb the stoichometry of a number of protein complexes (protein "dosage imbalances"; Papp *et al.* 2003). GO term enrichment analyses indicate that stress proteins with functions in "intracellular organelle" are especially "conserved" (Table 3). Not surprisingly, transcriptional regulators with "transcription regulatory region DNA binding" are among the most often deleted genes, which is in line with previous observations (Nikolaou *et al.* 2009, Zhang *et al.* 2016). The evolution of transcription factors is rapid in fungi, similar to that of stress stimuli sensors (Nikolaou *et al.* 2009, Zhang *et al.* 2016). Further studies are needed to shed light on the significance of the enrichment of those "deleted" stress proteins, which are functionally coupled to "phagophore assembly site membrane" ("a cellular membrane associated with the pre-autophagosomal structure"; https://www.yeastgenome.org/go/34045) in the Aspergilli, because autophagy is a crucial element of stress defence in both baker's and fission yeasts (Thorpe *et al.* 2004, Su *et al.* 2015). The enrichment of "duplicated" stress proteins with "ion transmembrane transporter activity" or with function in "carbohydrate metabolic process", which includes proteins (*e.g.* Tdh3, Gpp1, Gpd2, Dak2, Tps1 and Tps3) involved in the metabolism of glycerol or trehalose as well known "stress metabolites", is not surprising and obviously will contribute to the evolution of the *Aspergillus* stress defence system (Table 3). These duplicated proteins are likely selected to reach/maintain higher ion and metabolic fluxes (Papp *et al.* 2004).

Although some closely related *Aspergillus* species like *A. aculeatus* – *A. carbonarius*, *A. sydowii* – *A. versicolor*, *A. luchuensis* – *A. tubingensis* as well as the two tested *A. niger* strains (CBS 113.46 and N402) showed similar stress tolerance patterns (Fig. 3) there was no significant correlation between the positions of the studied *Aspergillus* species in the dendrogram constructed from the stress physiological data (Fig. 3) and their positions in the yeast-based or *A. nidulans*-based stress response system models (Figs 1 and 2) or in the phylogenetic tree (Figs 3 and 4). Nevertheless, we cannot rule out that either addition of further stress proteins (when their involvement in stress tolerance is justified) or studying and including additional stress tolerance data to the analysis will modify these results in the future. However, according to Supplementary Fig. 2, which demonstrates that gene duplications and deletions do not show any strong correlation (correlation coefficients lower than −0.6 or higher than 0.6) with any of the studied stress tolerances for the vast majority of stress genes, addition of few new stress proteins or new stress conditions are unlikely to result in significantly different conclusions. The lack of significance could indicate that the deletion or alteration of a single stress response protein causes profound changes in the stress tolerance of an *Aspergillus* species, and as a consequence, tolerances against various specific types of environmental stress may correlate well with the copy numbers of few selected stress proteins rather than the total number of stress protein duplications and deletions. *E.g.* the number of MpkC orthologs correlated positively with cell wall stress tolerance, osmotic stress tolerance and oxidative stress tolerance (Supplementary Table S7, Supplementary Fig. S2). The osmophility of *A. glaucus* and *A. wentii* could be attributed simply to the lack of their *A. nidulans* GfdB NAD-dependent glycerol-3-phosphate dehydrogenase orthologs while the remarkable superoxide tolerance of *A. brasiliensis* was related to the appearance of two new *sod* genes in the genome of the fungus (de Vries *et al.* 2017, Orosz *et al.* 2018).

Processes other than gene duplications or deletions contribute to stress adaptation and may substitute or even counterbalance the effects of the altered copy numbers of a stress gene which may also explain the lack of correlations we found in this study. These processes likely include the following:

(i) Changes in the stress protein structure/activity can increase resistance.

Several examples have been reported in the literature that mutations in genes encoding stress defence proteins, *e.g.* against an antifungal agent, led to acquired resistance (Revie *et al.* 2018). As an example, Bódi *et al.* (2017) found in laboratory-based evolution experiments performed with *S. cerevisiae* that increased fluconazole resistance was frequently accompanied with mutations in the *pdr5* gene encoding a fluconazole efflux pump.

(ii) Changes in the expression level of a stress gene can be an alternative of gene duplication/deletion events.

In the above-mentioned experiment by Bódi *et al.* (2017), another frequently mutated gene was the *rox1*, which encodes a repressor of hypoxic genes (including ergosterol biosynthesis genes; Montañés *et al.* 2011) suggesting that beside altered protein activities, altered expression levels were also important in the adaptation to fluconazole in these experiments. (In fact, in laboratory evolution experiments, most of the analysed endpoint mutants had mutations in a regulatory gene (Conrad *et al.* 2009, 2011). Pca1 is another good example for the importance of the transcriptional activity of a stress gene : *S. cerevisiae* Pca1 is a cadmium efflux pump (Adle *et al.* 2007) and its ortholog in *A. fumigatus* (PcaA) is also involved in cadmium tolerance (Bakti *et al.* 2018). Genome analysis studies demonstrated that the most cadmium tolerant *Aspergillus* spp. (*A. fumigatus*, *A. versicolor*, *A. sydowii*) have a *pca1* ortholog (in the case of *A. sydowii* two orthologs), while the most Cd sensitive species (*A. carbonarius*, *A. aculeatus*, *A. glaucus*) have no *pca1* ortholog (de Vries *et al.* 2017). However, in spite of the exceptions, *A. niger*, having no *pca1* ortholog, generally showed higher cadmium tolerance than *A. flavus*, harboring one *pca1* ortholog (de Vries *et al.* 2017). Not surprisingly, the Kruskal-Wallis test did not show any significant difference in the cadmium tolerance among the *Aspergillus* species possessing two, one or even zero *pcaA* orthologs (Kurucz *et al.* 2018b). Moreover, the $MIC_{50,cd}$ values of 10 *A. fumigatus* isolates varied within a wide range (0.25 mM–>2 mM) and showed strong positive correlation with

the relative transcription of the *pcaA* gene (Kurucz *et al.* 2018b). All these data demonstrated that harbouring a *pca1* ortholog is not necessarily accompanied with an outstandingly high cadmium tolerance and, *vice versa*, losing the *pca1* ortholog not necessarily increase cadmium sensitivity.

(iii) Rewiring stress response regulatory networks can also be an inherent part of genetic adaptation to stress.

Evolution of stress responding pathways in fungi demonstrates that – in addition to gene expansion/gene loss events as well as gene structure diversification – functional changes in signalling pathways is an important element of the adaptation to changing environment (Pusztahelyi & Pócsi 2013, Hagiwara *et al.* 2016, Xu *et al.* 2017). In contrast to mitogen-activated protein kinases which seem to be evolutionarily stable, up-stream (*e.g.* sensors of various environmental stress stimuli) and down-stream (*e.g.* transcriptional regulators) elements of the same stress sensing, signal transduction and stress response regulatory pathways tend to undergo rapid changes (Nikolaou *et al.* 2009, Zhang *et al.* 2016).

Positive feedback mechanisms commonly cause phenotypic heterogeneity (Becskei *et al.* 2001), which can be important in physiological adaptation to stress and can also contribute to genetic adaptation to a stress condition *via* compensating possible negative effects of mutations on fitness (Mustonen & Lässig 2010, Sánchez-Romero & Casadesús 2014, Bódi *et al.* 2017). Development or cessation of these autoregulatory mechanisms may therefore significantly influence the (evolution of) stress tolerance of microbes.

(iv) Acquiring new genes *via* horizontal gene transfer can be an efficient alternative of gene duplication and subsequent neofunctionalisation/subfunctionalisation events.

It is well documented that horizontal gene transfer significantly contributes to the evolution of fungi (Fitzpatrick 2012). As an example, the genome of the commercial wine yeast *S. cerevisiae* EC118 contains 34 genes that have been transferred horizontally from other fungal species (Novo *et al.* 2009). These genes contribute to the adaptation to various types of stress induced by high-osmolarity, nitrogen starvation or high ethanol concentrations (Novo *et al.* 2009).

Taking into consideration the outcomes of the comparative stress protein and stress tolerance analyses, we suggest the following:

Gene duplication events support successful stress adaptation in new habitats by modulating the dose of important genes (Wapinski *et al.* 2007, Fedorova *et al.* 2008, Ames *et al.* 2010, Shertz *et al.* 2010, Giacometti *et al.* 2011, Levasseur & Pontarotti 2011). Changing environmental conditions may trigger the erosion (*e.g.* pseudogenisation) and even the loss of certain genes, which lost their significance, with the concomitant arise of new ones, *e.g. via* new gene duplications (Fedorova *et al.* 2008, Jackson *et al.* 2009, Moran *et al.* 2011). These changes contribute to the expansion or reduction of certain stress response gene families (Nikolaou *et al.* 2009, Zhang *et al.* 2016) and, as a consequence, to reshaping the genome of the organism, which may lead to the evolution of new species (Fig. 4). Beside gene duplication and deletion events, other processes can also significantly affect the adaptation of fungi to new habitats which generally prevents any significant correlation between the copy numbers of proteins and stress tolerances (Fig. 4). These processes include (i) alterations in the structure of stress proteins, which has an impact on their activity, (ii) varying biosynthesis of these proteins, (iii) rewiring stress response regulatory networks, or (iv) acquiring new genes through horizontal gene transfer. All these multilevel changes seem to be indispensable for the successful adaptation of filamentous fungi when they are entering new ecological niches.

Considering the future applications of yeast-based models to gain a deeper insight in the stress defence systems of the Aspergilli we can conclude that both *S. cerevisiae*-based and *S. pombe*-based models are suitable to identify *Aspergillus* stress response proteins and also to reconstruct their interactions. Nevertheless, because of the significant evolutionary distances between ascomycetous yeasts and filamentous Aspergilli (Wang *et al.* 2009), stress protein paralog pairs are often difficult to identify using yeast-based models. Not surprisingly, *Aspergillus* orthologs of yeast stress proteins may even have modified or altered physiological functions (Balázs *et al.* 2010). In addition, *Aspergillus* genomes harbour significantly more genes than those of *S. cerevisiae* and *S. pombe* and, hence, many important *Aspergillus* stress proteins do not have any yeast counterparts (Supplementary Tables S3 and S4). Therefore, further efforts are needed to speed up the functional characterisation of *Aspergillus* (especially *A. nidulans*) stress response proteins and set up a detailed stress defence system model preferentially based on *A. nidulans*, which will be at least equivalent to or even better than today's frequently used yeast-based models in either quality or applicability (Miskei *et al.* 2009, Karányi *et al.* 2013).

## ACKNOWLEDGEMENTS

The Authors thank the following researchers at the University of Debrecen for their valuable contributions to this project: A Gazdag, R Mohácsi, and LG Tóth (for participation in the experimental work in setting up FSD), V Szabó and T Koszó (for literature search during the construction of the Stress Protein Database), and B Nyul (for contributing to the discussion of MDS methods). This work was supported by the European Union and the European Social Fund through the project EFOP-3.6.1-16-2016-00022, the Higher Education Institutional Excellence Programme of the Ministry of Human Capacities in Hungary, within the framework of the Biotechnology thematic programme of the University of Debrecen, and the National Research, Development and Innovation Office (Hungary) with the grants NKFIH K100464, K112181 and K119494.

## APPENDIX A. SUPPLEMENTARY DATA

Supplementary data to this article can be found online at https://doi.org/10.1016/j.simyco.2018.10.003.

## REFERENCES

Adle DJ, Sinani D, Kim H, *et al.* (2007). A cadmium-transporting P1B-type ATPase in yeast *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **282**: 947–955.

Ames RM, Rash BM, Hentges KE, *et al.* (2010). Gene duplication and environmental adaptation within yeast populations. *Genome Biology Evolution* **2**: 591–601.

Askwith C, Eide D, Van Ho A, *et al.* (1994). The FET3 gene of *S. cerevisiae* encodes a multicopper oxidase required for ferrous iron uptake. *Cell* **76**: 403–410.

Baker SE, Thykaer J, Adney WS, *et al*. (2008). Fungal genome sequencing and bioenergy. *Fungal Biology Reviews* **22**: 1–5.

Bakti F, Sasse C, Heinekamp T, *et al*. (2018). Heavy metal-induced expression of PcaA provides cadmium tolerance to *Aspergillus fumigatus* and supports its virulence in the *Galleria mellonella* model. *Frontiers in Microbiology* **9**: 744.

Balázs A, Pócsi I, Hamari Zs, *et al*. (2010). AtfA bZIP-type transcription factor regulates oxidative and osmotic stress responses in *Aspergillus nidulans*. *Molecular Genetics and Genomics* **283**: 289–303.

Becskei A, Séraphin B, Serrano L (2001). Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *The EMBO Journal* **20**: 2528–2535.

Benoit I, Culleton H, Zhou M, *et al*. (2015). Closely related fungi employ diverse enzymatic strategies to degrade plant biomass. *Biotechnology for Biofuels* **8**: 107.

Bódi Z, Farkas Z, Nevozhay D, *et al*. (2017). Phenotypic heterogeneity promotes adaptive evolution. *PLoS Biology* **15**: e2000644.

Bruder Nascimento AC, Dos Reis TF, de Castro PA, *et al*. (2016). Mitogen activated protein kinases $SakA^{HOG1}$ and MpkC collaborate for *Aspergillus fumigatus* virulence. *Molecular Microbiology* **100**: 841–859.

Castresana J (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**: 540–552.

Conrad TM, Joyce AR, Kenyon Applebee M, *et al*. (2009). Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biology* **10**: R118.

Conrad TM, Lewis NE, Palsson BØ (2011). Microbial laboratory evolution in the era of genome-scale science. *Molecular System Biology* **7**: 509.

de Vries RP, Visser J (2001). *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiology and Molecular Biology Reviews* **65**: 497–522.

de Vries RP, Riley R, Wiebenga A, *et al*. (2017). Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. *Genome Biology* **18**: 28.

Dray S, Dufour AB (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* **22**: 1–20.

Egbuta MA, Mwanza M, Babalola OO (2017). Health risks associated with exposure to filamentous fungi. *International Journal of Environmental Research and Public Health* **14**: E719.

Enright AJ, Van Dongen S, Ouzounis CA (2002). An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research* **30**: 1575–1584.

Fedorova ND, Khaldi N, Joardar VS, *et al*. (2008). Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. *PLoS Genetics* **4**: e1000046.

Fitzpatrick DA (2012). Horizontal gene transfer in fungi. *FEMS Microbiology Letters* **329**: 1–8.

Furukawa K, Hoshi Y, Maeda T, *et al*. (2005). *Aspergillus nidulans* HOG pathway is activated only by two-component signalling pathway in response to osmotic stress. *Molecular Microbiology* **56**: 1246–1261.

Gabaldón T, Martin T, Marcet-Houben M, *et al*. (2013). Comparative genomics of emerging pathogens in the *Candida glabrata* clade. *BMC Genomics* **14**: 623.

Giacometti R, Kronberg F, Biondi RM, *et al*. (2011). *Candida albicans* Tpk1p and Tpk2p isoforms differentially regulate pseudohyphal development, biofilm structure, cell aggregation and adhesins expression. *Yeast* **28**: 293–308.

Grallert Á, Boke E, Hagting A, *et al*. (2015). A PP1-PP2A phosphatase relay controls mitotic progression. *Nature* **517**: 94–98.

Gresham D, Desai MM, Tucker CM, *et al*. (2008). The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genetics* **4**: e1000303.

Grigoriev IV, Cullen D, Goodwin SB, *et al*. (2011). Fueling the future with fungal genomics. *Mycology* **2**: 192–209.

Grigoriev IV, Nikitin R, Haridas S, *et al*. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research* **42**: D699–D704.

Hagiwara D, Matsubayashi Y, Marui J, *et al*. (2007). Characterization of the NikA histidine kinase implicated in the phosphorelay signal transduction of *Aspergillus nidulans*, with special reference to fungicide responses. *Bioscience, Biotechnology, and Biochemistry* **71**: 844–847.

Hagiwara D, Sakamoto K, Abe K, *et al*. (2016). Signaling pathways for stress responses and adaptation in *Aspergillus* species: stress biology in the post-genomic era. *Bioscience, Biotechnology, and Biochemistry* **80**: 1667–1680.

Haro R, Garciadeblas B, Rodríguez-Navarro A (1991). A novel P-type ATPase from yeast involved in sodium transport. *FEBS Letters* **291**: 189–191.

Hayashi S, Yoshioka M, Matsui T, *et al*. (2014). Control of reactive oxygen species (ROS) production through histidine kinases in *Aspergillus nidulans* under different growth conditions. *FEBS Open Biology* **4**: 90–95.

Jackson AP, Gamble JA, Yeomans T, *et al*. (2009). Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*. *Genome Research* **19**: 2231–2244.

Jaimes-Arroyo R, Lara-Rojas F, Bayram Ö, *et al*. (2015). The SrkA kinase is part of the SakA mitogen-activated protein kinase interactome and regulates stress responses and development in *Aspergillus nidulans*. *Eukaryotic Cell* **14**: 495–510.

Karányi Z, Holb I, Hornok L, *et al*. (2013). FSRD: fungal stress response database. *Database (Oxford)* **2013**: bat037.

Kawasaki L, Wysong D, Diamond R, *et al*. (1997). Two divergent catalase genes are differentially regulated during *Aspergillus nidulans* development and oxidative stress. *Journal of Bacteriology* **179**: 3284–3292.

Kurucz V, Krüger T, Antal K, *et al*. (2018a). Additional oxidative stress reroutes the global response of *Aspergillus fumigatus* to iron depletion. *BMC Genomics* **19**: 357.

Kurucz V, Kiss B, Szigeti ZsM, *et al*. (2018b). Physiological background of the remarkably high $Cd^{2+}$ tolerance of the *Aspergillus fumigatus* Af293 strain. *Journal of Basic Microbiology*: in press.

Lange L (2014). The importance of fungi and mycology for addressing major global challenges. *IMA Fungus* **5**: 463–471.

Levasseur A, Pontarotti P (2011). The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics. *Biology Direct* **6**: 11.

Lücking R, Hawksworth DL (2018). Formal description of sequence-based voucherless *Fungi*: promises and pitfalls, and how to resolve them. *IMA Fungus* **9**: 143–166.

Maguire SL, ÓhÉigeartaigh SS, Byrne KP, *et al*. (2013). Comparative genome analysis and gene finding in *Candida* species using CGOB. *Molecular Biology and Evolution* **30**: 1281–1291.

Mallo AC, Nitiu DS, Elíades LA, *et al*. (2017). Fungal degradation of cellulosic materials used as support for cultural heritage. *International Journal of Conservation Science* **8**: 619–632.

Martin F, Cullen D, Hibbett D, *et al*. (2011). Sequencing the fungal tree of life. *New Phytologist* **190**: 818–821.

Meyer V, Andersen MR, Brakhage AA, *et al*. (2016). Current challenges of research on filamentous fungi in relation to human welfare and a sustainable bio-economy: a white paper. *Fungal Biology and Biotechnology* **3**: 6.

Miskei M, Karányi Z, Pócsi I (2009). Annotation of stress-response proteins in the aspergilli. *Fungal Genetics and Biology* **46**: S105–S120.

Montañés FM, Pascual-Ahuir A, Proft M (2011). Repression of ergosterol biosynthesis is essential for stress resistance and is mediated by the Hog1 MAP kinase and the Mot3 and Rox1 transcription factors. *Molecular Microbiology* **79**: 1008–1023.

Moran GP, Coleman DC, Sullivan DJ (2011). Comparative genomics and the evolution of pathogenicity in human pathogenic fungi. *Eukaryotic Cell* **10**: 34–42.

Mustonen V, Lässig M (2010). Fitness flux and ubiquity of adaptive evolution. *Proceedings of the National Academy of Sciences* **107**: 4248–4253.

Nakamura T, Yamada KD, Tomii K, *et al*. (2018). Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/bty121.

Nikolaou E, Agrafioti I, Stumpf M, *et al*. (2009). Phylogenetic diversity of stress signalling pathways in fungi. *BMC Evolutionary Biology* **9**: 44.

Norbeck J, Påhlman AK, Akhtar N, *et al*. (1996). Purification and characterization of two isoenzymes of DL-glycerol-3-phosphatase from *Saccharomyces cerevisiae*. Identification of the corresponding *GPP1* and *GPP2* genes and evidence for osmotic regulation of Gpp2p expression by the osmosensing mitogen-activated protein kinase signal transduction pathway. *Journal of Biological Chemistry* **271**: 13875–13881.

Novo M, Bigey F, Beyne E, *et al*. (2009). Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *Proceedings of the National Academy of Sciences* **106**: 16333–16338.

Ohkura H, Adachi Y, Kinoshita N, *et al*. (1988). Cold-sensitive and caffeine-supersensitive mutants of the *Schizosaccharomyces pombe dis* genes implicated in sister chromatid separation during mitosis. *The EMBO Journal* **7**: 1465–1473.

Orosz E, van de Wiele N, Emri T, *et al*. (2018). Fungal Stress Database (FSD) - a repository of fungal stress physiological data. *Database (Oxford)* **2018**: bay009.

Papp B, Pál C, Hurst LD (2003). Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**: 194–197.

Papp B, Pál C, Hurst LD (2004). Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* **429**: 661–664.

Park HS, Jun SC, Han KH, *et al*. (2017). Diversity, application, and synthetic biology of industrially important *Aspergillus* fungi. *Advances in Applied Microbiology* **100**: 161–202.

Pereira Silva L, Alves de Castro P, Dos Reis TF, *et al*. (2017). Genome-wide transcriptome analysis of *Aspergillus fumigatus* exposed to osmotic stress reveals regulators of osmotic and cell wall stresses that are $SakA^{HOG1}$ and MpkC dependent. *Cellular Microbiology* **19**: cmi12681.

Perrone G, Gallo A (2016). *Aspergillus* species and their associated mycotoxins. In: *Mycotoxigenic Fungi: Methods and Protocols* (Moretti A, Susca A, eds). Springer Science+Business Media, New York: 33–49.

Pusztahelyi T, Pócsi I (2013). Functions, cooperation and interplays of the vegetative growth signaling patwhays in the aspergilli. *Journal of Mycology* **2013**: 832521.

Revie NM, Iyer KR, Robbins N, *et al*. (2018). Antifungal drug resistance: evolution, mechanisms and impact. *Current Opinion in Microbiology* **45**: 70–76.

Ruotolo R, Marchini G, Ottonello S (2008). Membrane transporters and protein traffic networks differentially affecting metal tolerance: a genomic phenotyping study in yeast. *Genome Biology* **9**: R67.

Samson RA, Noonim P, Meijer M, *et al*. (2007). Diagnostic tools to identify black aspergilli. *Studies in Mycology* **59**: 129–145.

Sánchez-Romero MA, Casadesús J (2014). Contribution of phenotypic heterogeneity to adaptive antibiotic resistance. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 355–360.

Shertz CA, Bastidas RJ, Li W, *et al*. (2010). Conservation, duplication, and loss of the Tor signaling pathway in the fungal kingdom. *BMC Genomics* **11**: 510.

Skamnioti P, Henderson C, Zhang Z, *et al*. (2007). A novel role for catalase B in the maintenance of fungal cell-wall integrity during host invasion in the rice blast fungus *Magnaporthe grisea*. *Molecular Plant-Microbe Interactions* **20**: 568–580.

Souza Guimarães LH, da Costa Souza PN (2017). *Aspergillus* biotechnology: an overview on the production of hydrolases and secondary metabolites. *Current Biotechnology* **6**: 283–294.

Stamatakis A (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Stearman R, Yuan DS, Yamaguchi-Iwai Y, *et al*. (1996). A permease-oxidase complex involved in high-affinity iron uptake in yeast. *Science* **271**: 1552–1557.

Su Y, Chen C, Huang L, *et al*. (2015). *Schizosaccharomyces pombe* homologs of human DJ-1 are stationary phase-associated proteins that are involved in autophagy and oxidative stress resistance. *PLoS One* **10**: e0143888.

Sugui JA, Kwon-Chung KJ, Juvvadi PR, *et al*. (2014). *Aspergillus fumigatus* and related species. *Cold Spring Harbor Perspectives in Medicine* **5**: a019786.

Thorpe GW, Fong CS, Alic N, *et al*. (2004). Cells have distinct mechanisms to maintain protection against different reactive oxygen species: oxidative-stress-response genes. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 6564–6569.

Wang H, Xu Z, Gao L, *et al*. (2009). A fungal phylogeny based on 82 complete genomes using the composition vector method. *BMC Evolutionary Biology* **9**: 195.

Wapinski I, Pfeffer A, Friedman N, *et al*. (2007). Natural history and evolutionary principles of gene duplication in fungi. *Nature* **449**: 54–61.

Xu C, Liu R, Zhang Q, *et al*. (2017). The diversification of evolutionarily conserved MAPK cascades correlates with the evolution of fungal species and development of lifestyles. *Genome Biology and Evolution* **9**: 311–322.

Zhang L, Zhou Z, Guo Q, *et al*. (2016). Insights into Adaptations to a near-obligate nematode endoparasitic lifestyle from the finished genome of *Drechmeria coniospora*. *Scientific Reports* **6**: 23122.