Comparative RNA-seq based transcriptomic analysis of *Aspergillus nidulans* recombinant strains overproducing heterologous glycoside hydrolases

Felipe Calzado[1,2], Gabriela F. Persinoti[1], Cesar Rafael Fanchini Terrasan[2], Mariane P. Zubieta[1,2], Marcelo V. Rubio[1,2], Fabiano J. Contesini[1], Fabio M. Squina[1], André R. L. Damásio[2*]

[1]Brazilian Bioethanol Science and Technology Laboratory (CTBE), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, SP, Brazil. 13083-970

[2]Department of Biochemistry and Tissue Biology, Institute of Biology, University of Campinas (UNICAMP), Campinas, SP, Brazil. 13083-862

***To whom correspondence should be addressed:**

André R. L. Damásio; Department of Biochemistry and Tissue Biology, Institute of Biology, University of Campinas (UNICAMP), Campinas-SP, Brazil Email: adamasio@unicamp.br

ABSTRACT

Filamentous fungi are important cell factories for the production and secretion of homologous and heterologous lignocellulolytic enzymes, however the regulation of protein secretion in these organisms need to be further explored. In order to investigate this regulation, *Aspergillus nidulans* recombinant strains were analyzed by transcriptomics. We designed three *A. nidulans* recombinant strains producing the heterologous proteins: alpha-arabinofuranosidase (*AbfA*), beta-glucosidase (*BglC*) and thermophilic mannanase (*1542*). The heterologous genes *abfA* and *bglC* were highly expressed, while the levels of *1542* mRNA were similar to the reference gene. There was an indirect relationship between mRNA and the protein secretion levels, suggesting that transcription is not a bottleneck in these systems. Based on the general RNA-seq analysis of the recombinant strains, it was possible to observe a predominant up-regulation response. Moreover, biological processes related to metabolism, protein with binding function and cellular transport were enriched. We also observed the unconventional splicing of *hacA* for the recombinant strains *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$, indicating some level of unfolded protein response. The global analysis showed mild stress after 2 h induction of heterologous protein production, which normalized after 8 h. Our results provide insights to understand how *A. nidulans* adapts to the overproduction of heterologous proteins.

Keywords: *Aspergillus*; RNA-seq; Heterologous expression; Secretion of client proteins

INTRODUCTION

Fungi belonging to the genera *Aspergillus* have a high capacity for producing large quantities of extracellular enzymes that provide nutrients from complex organic materials, featuring its saprophytic lifestyles [1]. The natural ability for efficient protein secretion has led to their biotechnological exploitation as cell factories for the production of industrially important enzymes. A broad range of homologous and heterologous enzymes such as amylases, xylanases, and cellulases can be noticeably produced by *Aspergilli*, thus contributing to the expansion and growth of the enzyme market [2]. If continuing an annual 8.2% increase rate over the 5-year forecast period, the enzyme market is expected to reach approximately $7.1 billion by 2018 [3]. *Aspergillus* is among the main enzymes-producing microorganisms, corresponding to 30% of the commercially produced enzymes, according to the Association of Manufacturers and Formulators of Enzyme Products list [4].

Despite the above-mentioned advantages, there are still many challenges to be overcome regarding the expression of enzymes in fungal systems. It has been shown that the main bottleneck in heterologous protein production is not related to low mRNA expression [5]. Instead, experimental evidence suggests that most target proteins produced in filamentous fungi are lost or stuck in the secretory pathway due to errors in processing, modification or misfolding, hence resulting in their elimination by the endoplasmic reticulum (ER) quality control [6,7]. Misfolded proteins alter cell homeostasis and proper ER function, resulting in ER stress. ER stress activates conserved signaling pathways termed *unfolded protein response* (UPR) and *ER-associated protein degradation* (ERAD), which upregulates genes responsible for restoring protein folding homeostasis in the cells and degrading misfolded proteins in the cytosol by the ubiquitin-proteasome system, respectively [8].

In fungi, the UPR sensor is the inositol-requiring protein 1 (IRE1), which is an ER-resident transmembrane protein that contains a luminal domain functioning as a sensor of the protein folding state, and also presents a site-specific endoribonuclease (RNase) domain at the cytoplasmic C-terminus [9]. In the presence of unfolded proteins, Kar2/BiP dissociates from IRE1, which subsequently oligomerizes for activation [10]. The IRE1 oligomer is activated by autophosphorylation and the RNAse domain becomes responsible for an unconventional splicing of the *hacA*^{u} (*hacA* uninduced form, non-spliced) mRNA [9]. The presence of an unconventional intron in the mRNA of the bZIP transcription factor *hacA*^{u} causes translational attenuation in the absence of ER stress [11]. The intron length is variable among different species. For example, *Aspergillus* and *Trichoderma* have an unconventional intron with 20-nucleotides (nt) in length, while it is 252 nt in *Saccharomyces cerevisiae*, 26 nt in mice and 23 nt in *Caenorhabditis elegans* [12,13]. In addition to the unconventional splicing, *hacA*^{i} (*hacA* induced form, spliced) from *Trichoderma* and *Aspergillus* are truncated at the 5' flanking region upon UPR induction by removal of an upstream ORF [14,15]. HacA^{i} then moves to the nucleus, recognizing and binding to unfolded protein response elements (UPREs) in the promoter region of the regulated

genes. UPR-responsive genes provide elevated protein folding capacity, secretion, ER expansion, and degradation of irreversibly misfolded proteins by the proteasome [16].

Manipulation of the UPR pathway and its components has been a common attempt to improve the production of heterologous proteins in filamentous fungi [17–19]. Many ER stress-induced UPR genes, including protein folding-related genes such as chaperones and foldases, have been co-expressed with a heterologous gene. However, the overexpression of chaperones usually does not increase the production of heterologous proteins. In *S. cerevisiae*, the overexpression of *bipA* increases the amount of extracellular prochymosin by over 20-fold, although the secretion of thaumatin is not significantly improved [20].

Genomic and transcriptomic approaches have been used to understand the overexpression and production of heterologous enzymes and their capacity to induce UPR, providing valuable information on *Aspergillus* genes involved with the secretion and coordination between UPR and ERAD. The induction of UPR in *Aspergillus niger* was investigated by transcriptomic analysis, comparing the UPR induced by chemicals and by overexpression of the tissue plasminogen activator (t-PA) [16]. Approximately 94 genes were commonly induced, most related to the functional categories protein folding, translocation/signal peptidase complex, glycosylation, vesicle trafficking, and lipid metabolism [16]. In addition, another study showed that UPR results in activation of approximately 400 genes in *S. cerevisiae* (7-8% of the genome) [21]. This reflects the complex network of interactions between UPR and other signaling pathways in the cell, mainly in relation to the secretory pathway. Transcriptome profiles can be particularly important to provide an overview of all genes and pathways regulated by UPR in the cell.

Previous works have shown that UPR is activated in recombinant strains and have mainly focused on the secretion pathway [6,22,23]. We performed a comparative transcriptome analysis of three *Aspergillus nidulans* recombinant strains producing the following heterologous proteins: 1) GH51 alpha-arabinofuranosidase (*A. nidulans*$_{\mathbf{AbfA}}$**)**; 2) GH3 beta-glucosidase (*A. nidulans*$_{\mathbf{BglC}}$**)** and 3) GH5 thermophilic mannanase (*A. nidulans*$_{\mathbf{1542}}$**)**. These proteins were produced and secreted at different levels (AbfA > BglC > 1542). The transcriptional profiles of the *A. nidulans* recombinant strains were determined at early and late periods of induction. Furthermore, 17 genes related to the secretion pathway were differentially expressed (DE) in all the recombinant strains. This study describes the genome-wide transcriptional response of *A. nidulans* for heterologous protein production, offering new perspectives by comparing three different heterologous proteins with variable secretion levels. Our results provided insights on understanding how *A. nidulans* adapts to the overproduction of heterologous proteins and highlighted new potential targets for genetic manipulation and improvement of the protein secretion yields in this microbial system.

RESULTS AND DISCUSSION

Recombinant proteins are secreted in different levels

In order to understand the heterologous protein production in *A. nidulans*, three different proteins were monitored: GH51 alpha-arabinofuranosidase (AbfA) and GH3 beta-glucosidase (BglC) both cloned from *Aspergillus fumigatus*, and a thermophilic GH5 mannanase (1542) cloned from the hyperthermophilic bacterium *Thermotoga petrophila*. The target sequences were cloned into a pEXPYR vector and transformed into *A. nidulans*$_{\mathbf{A773}}$ [24]. These recombinant strains were selected due to an increased secretion level of the target proteins, *i.e*, AbfA > BglC > 1542. The transformed strains were denominated *A. nidulans*$_{\mathbf{AbfA}}$, *A. nidulans*$_{\mathbf{BglC}}$ and *A. nidulans*$_{\mathbf{1542}}$.

Initially, the time course of heterologous enzymes secretion was evaluated by growing the recombinant strains in maltose to induce the *glaAp* promoter [24] and activate expression of the heterologous genes (**Figure 1**). Despite the differences in secretion of the target proteins, the three strains showed similar profiles of mycelium dry weight, extracellular proteins and pH. The similarity among these parameters indicates that the overexpression of heterologous genes may not imply major physiological changes of the strains.

AbfA was the most secreted recombinant protein with an activity peak at 48 h (9.53 U/ml), with a low level of activity detected even at 6 h (**Figure 1A;** *middle panel*). BglC was less secreted than AbfA and the peak activity was early at 24 h (**Figure 1B;** *middle panel*; 8.73 U/ml). In general, after 48 h of cultivation, a protease-degradation profile was observed (**Figure 1**; *bottom panel*) for the three recombinant strains. At this cultivation time, cells probably undergo disruption due to the absence of nutrients. Especially high beta-glucosidase activity was detected from 72 to 120 h by using pNPG (**Figure 1B;** *middle panel*), but this activity was due to a native intracellular beta-glucosidase (AN2828; corresponding to an intense protein band between 75 and 85 kDa) whose amino acid sequence/identity was further confirmed by LC-MS/MS (data not shown). In contrast, the enzyme 1542 could not be observed in the gel, however some residual activity (under 0.6 U/ml) was detected in similar amounts from 18 to 120 h, suggesting that this thermophilic enzyme was resistant to degradation by *A. nidulans* proteases (**Figure 1C;** *middle panel*). It is important to emphasize that the enzymatic activity of this mannanase was measured at high temperatures (87 ºC), eliminating the possibility for false positive results, *i.e*, the activity of a native enzyme. Secretion of proteins into the medium driven by strong promoters (such as *glaAp*) is often subjected to extracellular proteolysis. This is a very common phenomenon, occurring in filamentous fungi that secrete native proteases, or in response to glucose starvation or pH increase [24].

In order to determine whether the low BglC and 1542 secretion were caused by problems during protein secretion, SDS-PAGE and enzyme activity assays were carried out in the intracellular fraction (**Figure S1**). The overexpression of heterologous proteins may cause their intracellular

accumulation, suggesting breakdown or overloading of the secretory pathway [25]. Protein bands related to the recombinant proteins were not observed in the intracellular fraction (**Figure S1**), indicating that most of the target proteins were immediately secreted, with low levels of unfolded/misfolded proteins trapped inside the cell.

In general, we observed that the secretion levels depend on the protein sequence and overexpression does not affect the overall physiology of the strains. The levels of target protein secretion increased from 1542 to AbfA and an intermediate level was observed for BglC. Thus, we sought to understand at the transcriptomic level whether these differences in heterologous enzymes secretion were related to their mRNA expression.

The heterologous genes are highly expressed in *A. nidulans*

The production of recombinant proteins in filamentous fungi by using strong promoters activates some regulatory genes in pathways related to protein quality control and UPR [26–28]. The UPR represents an adaptive response that aims to restore cellular homeostasis triggered by ER stress [29]. In order to investigate if the overexpression of recombinant genes in *A. nidulans* triggers UPR, expression of the target genes *abfA*, *bglC* and *1542* were measured by qPCR. The *abfA* and *bglC* genes were highly expressed when compared to the endogenous reference gene (*tubC*), while the levels of *1542* mRNA were similar to the *tubC.* The genes were detected even after 72 h of maltose induction (**Figure 2A**).

UPR target genes such as *bipA* and *sec63* were quantified in function of time (**Figures 2B and C**) [22,25,30,31]. The *bipA* gene was overexpressed in the three recombinant strains at 12 h and exclusively in *A. nidulans*$_{\mathbf{AbfA}}$ at 8 h (**Figure 2**), probably due to the higher AbfA production compared to the other recombinant proteins. The *sec63* gene was overexpressed in *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$ at 12 h and in *A. nidulans*$_{\mathbf{1542}}$ only at 24 h. Late overexpression (48 and 72 h) of *bipA* and *sec63* is probably related to the absence of nutrients, considering that maltose was totally consumed (**Figure S2**), causing starvation, UPR and cell death [32]. In *Trichoderma reesei,* protein overproduction induced the overexpression of UPR target genes only at early cultivation stages (from 1 to 12 h) [33]. After 12 h, the quantity of UPR target genes returned to basal levels. Heterologous expression of the bacterial xylanase B (*xynB*) in *A. niger* also showed lower mRNA levels of some UPR target genes (*gla*, *bip*, and *hac1*) [34]. Here, we observed the overexpression of *bipA* and *sec63* primarily at 12 h after induction of the heterologous gene expression. Based on the time course of recombinant protein production, the fast maltose consumption and qPCR data, we defined 2 and 8 h of heterologous genes expression as standards to understand how the *A. nidulans* strains adapt to the production of heterologous proteins.

The indirect relationship between high mRNA levels (**Figure 2A**) and low levels of secreted proteins (**Figure 1B and C**) was also reported for cellobiohydrolases (CBHs) from *Aspergillus*

terreus and *T. reesei* heterologously expressed in *Aspergillus carbonarius* [35]. These authors speculated that proteolytic degradation could be one reason, possibly along with other factors such as incorrect folding, post-translational processing, and impairment of intracellular transport [35]. The indirect relationship between mRNA and protein levels is frequently reported, suggesting that transcription is not a bottleneck in these systems. In an attempt to establish a set of cellular processes altered during the adaptation of *A. nidulans* to the overexpression of recombinant genes, we analyzed the three recombinant strains by RNA-seq at 2 and 8 h.

Global transcriptional response to heterologous protein production

Differential expression analysis by RNA-seq was performed by pairwise comparisons of each recombinant strain compared to the control strain *A. nidulans*$_{\mathbf{A773}}$. We identified 476 differentially expressed (DE) genes by applying the pipeline and filters described in the Methods section (**Figure 3**). Except for the strains BglC and 1542 at 8 h, there was a predominant up-regulation response during the overexpression of recombinant genes in *A. nidulans.* Thirty out of 476 DE genes were common for the three recombinant strains. *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$ also showed 74 DE genes in common. This observation is supported by the Venn diagrams and PCA analysis (**Figure 3 and Figure S3**). This relative similarity between *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$ cannot be explained by mRNA titer, which was not similar for all strains (**Figure 3**). Protein features such as amino acid composition, the complexity of folding and post-translational modifications are probably crucial to their successful production and secretion. Optimal expression levels of recombinant proteins are known to be variable for different proteins, as folding rates can vary according to protein size and structure [36].

To better understand how the general cellular processes are regulated during the production of recombinant proteins in *A. nidulans*, the annotation of DE genes was performed with FunCat [37]. The most significant functional category was “metabolism”, followed by “protein with binding function or cofactor”, “cellular transport” and “cell defense”. These four categories were enriched in all recombinant strains. In a more detailed analysis of the functional subcategories (**Figure 4),** we found several altered processes which are related to cellular stress, such as the subcategories metabolism, transport routes and cellular defense.

These subcategories were also reported in other *Aspergillus* species overproducing recombinant proteins [27,38,39]. The overexpression of glucoamylase (GlaA) in *A. niger* resulted in the activation of “translocation”, “protein glycosylation”, “vesicle transport” and “ion homeostasis” processes [38]. The functional categories “cellular transport”, “amino acid metabolism”, “aminoacyl-tRNA biosynthesis” and “metabolism” were overrepresented in an *Aspergillus oryzae* strain expressing a constitutively active form of *hacA*, indicating its importance for UPR, fungal growth,

and physiology [39]. As previously mentioned, we found several altered processes related to cellular stress, such as metabolism, transport routes and cellular defense. This conclusion is supported by qPCR data that showed an upregulation of *bipA* in all the strains and sec63 in *A. nidulans*$_{\mathbf{1542}}$ (**Figure 2).**

The basic sensing pathway to detect ER stress or an increase in the folding load is highly conserved from yeast to humans. The increase in misfolded proteins stimulates the Ire1 autophosphorylation and dimerization, triggering the unconventional splicing of an intron in the *hacA*^{u} to create the transcriptionally active form *hacA*^{i} [29]. HacA is a conserved bZIP transcription factor in eukaryotic cells, regulating gene expression in response to various forms of secretion stress and as part of secretory cell differentiation [16]. In order to evaluate the canonical pathway of UPR, we analyzed the ratio of *hacA* splicing in the *A. nidulans* recombinant strains based on the RNAseq data (**Figure S4**). The levels of *hacA* splicing were higher for AbfA and 1542 than for BglC at 2 h, and after 8 h the levels of *hacA* splicing for all the recombinant strains were lower than the control strain, showing a normalization of the UPR canonical pathway. This result suggests that there is no accumulation of the BglC misfolded form or the presence of an alternative pathway of ER stress and UPR. On the other hand, for AbfA and 1542 it was possible to infer that there is some misfolded protein accumulation at 2 h. Wang et al (2014) analyzed *hac1* levels in the *T. reesei* Rut C30 and QM9414 strains. The transcript levels of total *hac1* increased earlier in Rut C30 (1 h of induction) and this response was maintained longer than that in QM9414 [33]. Compared to our data, during the expression of recombinant genes, the *hacA* expression appears to be up-regulated at earlier times, with a response similar to that from *T. reesei* [33].

Based on the general analysis of RNA-seq of the recombinant strains, it was possible to observe a predominant up-regulation response. Moreover, biological processes related to metabolism, protein with binding function and cellular transport were enriched. Although there are not many processes related to cell stress, we observed the unconventional splicing of *hacA* for *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$, indicating some level of ER stress. The global analysis showed mild stress at 2 h after inducing heterologous protein production, which was normalized after 8 h.

Recombinant strains adapt differently at early and late times of heterologous gene induction

To further evaluate how the recombinant strains adapt to heterologous protein production, a strain-specific expression profile was suggested. The DE genes previously identified were classified as early, continuous or late responsive genes. Early response corresponds to DE genes exclusively found at 2 h; continuous response, DE genes found at both 2 and 8 h; and late response, DE genes exclusively found at 8 h. The strain-specific profiles (**Figure S5**) showed a clear transcriptional difference between the early and late responsive genes, and a higher number of common genes

between *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$. These two strains presented a more evident early response, while *A. nidulans*$_{\mathbf{BglC}}$ presented a predominant late response.

It is interesting to note that there are fewer genes associated with the continuous response, reinforcing that there are distinct responses between 2 and 8 h of heterologous protein induction. *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$ presented a more evident early response, with 52 and 54% of the DE genes classified as early response genes, respectively. The opposite was observed for *A. nidulans*$_{\mathbf{BglC}}$, which presented 68% of the DE genes as late response genes (**Figure S5**). Functional annotation indicated that in general the DE genes were mainly involved in metabolism and transport **(Figure 5)**. Early response genes were also associated with transmembrane transport and amino acid metabolism, while late response genes were predominantly related to C-compound and carbohydrate metabolism. Only *A. nidulans*$_{\mathbf{AbfA}}$ presented a continuous response during this analysis with the predominance of fatty acid metabolism and C-compound subcategories.

Apart from processes related to metabolism, by analyzing the specific time response genes of each recombinant strain, it is clear that *A. nidulans*$_{\mathbf{1542}}$ and *A. nidulans*$_{\mathbf{BglC}}$ presented more functional categories associated with "32 Cell Rescue, Defense and Virulence" than *A. nidulans*$_{\mathbf{AbfA}}$, which on the other hand predominantly presented processes related to "02 Energy" (**Figure 5**). *A. nidulans*$_{\mathbf{AbfA}}$ is the strain which produces the highest level of recombinant protein, thus it is expected to require more energy to correctly fold and secret the overproduced protein[40]. *A. nidulans*$_{\mathbf{1542}}$ and *A. nidulans*$_{\mathbf{BglC}}$ were enriched for cell defense processes; which might be one of the explanations for the lower levels of secreted proteins. Filamentous fungi have multiple defense systems to maintain genome stability and prevent propagation of selfish DNA, such as that from transposable elements and viruses. These systems can also be deployed against transformed exogenous DNA intended for transgene expression [5]. In this perspective, defenses against foreign genes may impose significant barriers to the use of filamentous fungi as hosts for heterologous gene expression, which involves the transformation of foreign DNA [5].

These early, late and continuous specific responses enabled us to observe that each recombinant protein triggered a specific adaptation in *A. nidulans* during the production of heterologous proteins. In order to determine a specific profile for each recombinant strain over time, we analyzed three functional subcategories specifically related to protein production and secretion, such as "32.01.07 unfolded protein response (*e.g.* ER quality control)", "14.13 protein/peptide degradation" and "14.07.02.02 N-directed glycosylation, deglycosylation" (**Figure 6 andAdditional file 1**). These three functional subcategories are related to UPR, ERAD and glycosylation, respectively, important processes for monitoring cellular stress caused by recombinant protein production. During the overexpression of recombinant genes, large amounts of transcripts go to ER to be translated and later folded into fully functional proteins. If the heterologous protein does not

fold correctly, it will accumulate in the ER and generate a condition termed ER stress or protein secretion stress [5]. The cell counteracts ER stress with different response mechanisms, including UPR and ERAD [41].

First, by observing 75 genes related to "32.01.07 Unfolded protein response (*e.g.* ER quality control)" (**Additional file 1: Table S1**), 15 genes were DE (**Figure 6A; Additional file 1: Table S2**), and only one gene was common to all strains (AN1163). This gene encodes a chaperone with an orthologue in *S. cerevisiae* (Hsp78p), which is described as able to prevent aggregation of misfolded proteins as well as re-solubilize protein aggregates [42]. *A. nidulans*$_{\mathbf{BglC}}$ showed more specific DE genes related to UPR, and almost all were described as chaperones (AN2432, AN6010, AN8269, AN9124 and AN2646). The only exception was the late up-regulated gene AN1296, described as an RNA ligase (TRL1, orthologue in *S. cerevisiae*) required for tRNA splicing and for both splicing and translation of HAC1 mRNA in UPR [42]. The canonical biological marker of UPR in *A. nidulans*, the *hacA*, was down-regulated in *A. nidulans*$_{\mathbf{1542}}$ and *A. nidulans*$_{\mathbf{AbfA}}$ at 2 h. In spite of the down-regulation of *hacA*, which is contrary to that commonly described for other *Aspergillus* species producing recombinant proteins [5], this result corroborates to our *hacA* coverage data (**Figure S4**), showing discrete unconventional splicing in all strains.

The other functional subcategory analyzed was "14.13 Protein/peptide degradation" with 354 genes (**Additional file 1: Table S3**). From that, 38 genes were DE in the recombinant strains (**Figure 6B and Additional file 1: Table S4**). Eight genes were common in all recombinant strains and two of these are chaperones (AN0858 and AN1163), with orthologues in *S. cerevisiae* (HSP104 and HSP78) [42]. The other six genes were *pkfC* (AN3226) and AN6931, both encoding a protein (ENV9) proposed to be involved in vacuolar functions in *S. cerevisiae*; the predicted threonine-type endopeptidase (AN3493) and the *prtA* (AN5558), which encode a broad-specificity thermostable alkaline protease; the extracellular deuterolysin-type metalloproteinase (AN7962) and the gene AN11028 predicted as a nucleotide binding protein with oxidoreductase activity [31]. Another interesting observation was that *A. nidulans*$_{\mathbf{1542}}$**,** the strain with the lowest level of protein production, showed more DE genes in this functional subcategory, indicating that protein degradation could interfere with 1542 production. This observation is supported by the differential expression of ubiquitin-ligase genes (AN0444 and AN0927) and peptidases (AN5619 and AN5810) only in *A. nidulans*$_{\mathbf{1542}}$. *A. nidulans*$_{\mathbf{AbfA}}$ also presented many specific DE genes related to protein/peptide degradation, which is probably associated with the higher amount of heterologous protein production, as evidenced by ubiquitin (AN7491 and AN2000) and proteolysis-related (AN7201 and AN10900) genes. It is well established that high production of heterologous proteins is connected to different stress responses[40]; and ERAD is among the conserved mechanisms that filamentous fungi use to overcome ER stress caused by the accumulation of misfolded heterologous proteins [43].

Finally, 43 genes related to the functional subcategory "14.07.02.02 N-directed glycosylation, deglycosylation" were analyzed (**Additional file 1: Table S5**). From this analysis, only three genes were DE in the recombinant strains (**Figure 6C and Additional file 1: Table S6**). These genes were down-regulated and predicted as alpha-1,6-mannosyltransferases (AN8986, AN7534, and AN1620) with the OCH1 and HOC1 orthologues in *S. cerevisiae* [42]. The enzyme OCH1 is involved in the transfer of mannose from GDP-Man to the α-1,3 branch of the tri-mannose core resulting in a α-1,6 extension, which is the substrate for additional mannosyltransferases throughout the yeast secretory pathway [44]. These alpha-1,6-mannosyltransferases are interesting targets for genetic manipulation in order to understand if the glycosylation pattern of these enzymes might affect the levels of recombinant protein production in *A. nidulans*.

Autophagy and cell wall pathways were also analyzed. These processes are important to overcome cellular stress caused by the overproduction of heterologous proteins. In order to analyze both pathways, Kimura et al. (2011) defined lists of genes involved in the *A. oryzae* autophagy pathway [45] and cell wall pathway genes were defined by Groot et al. (2009) [46]. Based on these lists, we defined their homologous genes in *A. nidulans* by using the AspGD data, generating a set of 26 genes for autophagy and 95 for the cell wall (which includes chitin synthesis, 1,3-alpha-glucan synthesis, 1,3-beta-glucan synthesis and cell wall integrity) [46]. For the autophagy pathway, from 26 genes only two were DE in our RNA-seq data. The gene AN10728, up-regulated only in *A. nidulans*$_{\mathbf{1542}}$, has an ortholog in *S. cerevisiae* (*atg22*) that encodes a vacuolar integral membrane protein required for efflux of amino acids during autophagic body breakdown [42]. The other gene, AN2876 (uncharacterized), was up-regulated in *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{BglC}}$. It is well known that autophagy is an important process that delivers misfolded secretory proteins accumulated in ER to vacuoles, a mechanism previously described for *A. oryzae* [45]. In another perspective, autophagy was also reported to be dispensable to overcome ER stress in *A. niger* [47]. Despite this controversy, the *A. nidulans* recombinant strains showed only two DE genes, which is a limited number that does not permit for making a specific analysis with regards to this pathway.

For the cell wall pathway, 31 genes were DE. The transcriptomic response of cell wall genes was primarily of down-regulation, and six genes were common for all recombinant strains. Interestingly, AN1604 (*agnE*) and AN5885 (*agsA*), were down-regulated at 2 h and up-regulated at 8 h. The first is predicted as a putative glycosylphosphatidylinositol (GPI)-anchored alpha-1,3-glucanase. The second encodes the catalytic subunit of the alpha-1,3 glucan synthase complex and plays a minor role in alpha-1,3 glucan synthesis [31]. In general, *A. nidulans*$_{\mathbf{1542}}$ and *A. nidulans*$_{\mathbf{AbfA}}$ showed a very similar response, specifically at 2 h, as previously observed for UPR and ERAD related genes. *A. nidulans*$_{\mathbf{AbfA}}$ showed more DE genes in the cell wall pathway, what may indicate that this strain is more exposed to secretion stress, which can also influence cell wall homeostasis [48]. Fiedler

et al. (2014) showed that *A. niger* senses and responds to cell wall stress with at least three transcription factors, RlmA, MsnA and CrzA [49]. Our transcriptomic data showed down-regulation of *rlmA* both in *A. nidulans*$_{\mathbf{1542}}$ and *A. nidulans*$_{\mathbf{AbfA}}$, indicating that these recombinant strains were regulated by cell wall stress caused by heterologous protein production [49].

The classification of early, continuous and late responses allowed us to observe that there is a great transcriptomic difference between the analyzed time points (2 and 8h) as well as high transcriptomic similarity between *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$. These peculiarities lead us to determine specific profiles of the three recombinant strains over time. Five functional subcategories were applied to understand the specific profile of heterologous protein secretion stress in *A. nidulans*, with UPR, glycosylation, ERAD, autophagy and cell wall synthesis/integrity. We then performed a deeper analysis of the secretion pathway genes.

Differential expression of genes related to the secretion pathway

The production of heterologous proteins is not only impaired by low expression of heterologous genes, but it can also be hampered by problems in post-translational processing in the secretory pathway [50]. Many studies have attempted to comprehend the superior secretion capacity of *Aspergillus*, primarily at the transcriptomic level [6,16,25,27,38]. These studies identified important genes in different stages of protein production such as translocation, folding, cargo transport and exocytosis [51]. Liu et al. (2014) defined a set of genes involved in the secretion pathway of *A. oryzae* using the secretory model of *S. cerevisiae* as a scaffold. Based on this list, we defined the homologous or the best hit genes in *A. nidulans* by using the AspGD data, generating a set of 374 genes (**Additional file 2: Table S7**).

Seventeen genes were DE in the recombinant strains (**Additional file 2: Table S8**) and each recombinant strain showed a singular response. *A. nidulans*$_{\mathbf{1542}}$ showed a predominant down-regulation response, while *A. nidulans*$_{\mathbf{BglC}}$ had an up-regulation response (**Figure 7A)**. *A. nidulans*$_{\mathbf{AbfA}}$ also presented more genes in common with *A. nidulans*$_{\mathbf{1542}}$ than with *A. nidulans*$_{\mathbf{BglC}}$ (**Figure 7B**), reinforcing the idea that the cellular response for these two recombinant strains was very similar. When we classified these 17 secretion pathway genes into early, continuous and late response, we found 12, 2 and 4 genes, respectively, strengthening the idea that the early response becomes basal over time, which was previously observed in the global analysis (**Figure S5)**.

In order to understand the functional categories represented by the 17 DE genes of the secretion pathway, a functional categorization was performed using FunCat. The most represented categories were “stress response”, “protein folding and stabilization” and “unfolded protein response”, with 10, 8 and 8 genes, respectively (**Figure 7C**). The absence of representative categories directly related to ER stress in the transcriptome global annotation (**Figure 4**) shows that the

overexpression of heterologous genes does not influence only the secretion pathway, but also other biological processes such as transport facilities/routes, homeostasis of ions and the fatty acid metabolism. The overproduction of α-amylase in *A. oryzae* altered some biological processes such as "protein N-linked glycosylation", "ER translocation and folding", "signal peptide processing", "ER to Golgi" and others [27].

Based on the secretion pathway genes we designed a secretory machinery profile for the three recombinant strains, highlighting some process such as protein folding, glycosylation, sugar transporters, stress response and others (**Figure 8**). The similar pattern between *A. nidulans*$_{\mathbf{AbfA}}$ and *A. nidulans*$_{\mathbf{1542}}$ was evident, as previously observed in the general transcriptomic response. This conclusion was corroborated by the overrepresented biological processes: "protein folding and stabilization", "protein modification" and "stress response". We considered some explanations for this observation, first that the enzyme BglC is a complex protein with 797 amino acids, and according to NetNGlyc 1.0 it is predicted to have five N-glycosylation sites, while AbfA and 1542 are smaller proteins (656 and 667 amino acids, respectively) with three and one predicted N-glycosylation sites, respectively. *A. nidulans*$_{\mathbf{BglC}}$ showed six exclusively up-regulated genes related to the secretion pathway (**Figure 7**), with five chaperones genes (AN2731, AN6010, AN6089, AN8269 and AN9124) involved in protein folding, stabilization, targeting, sorting and translocation (**Additional file 2: Table S8**). This observation also indicates the hypothesis that *A. nidulans*$_{\mathbf{BglC}}$ is a complex protein and it needs to overcome the properly folding. The sixth DE gene in this strain was AN1296, found to be up-regulated at 2 and 8 h, and moderately up-regulated in *A. nidulans*$_{\mathbf{AbfA}}$. The orthologous of this gene in *S. cerevisiae* (TRL1) is required for tRNA splicing and for both splicing and translation of *hac1* mRNA in the UPR [42]. This result converges with our relative coverage data (**Figure S4**) that shows the unconventional splicing of *hacA*, suggesting the induction of UPR for heterologous protein production, as previously reported [15,22,29,39,52,53]. On the other hand, *hacA* was down-regulated in *A. nidulans*$_{\mathbf{1542}}$ and *A. nidulans*$_{\mathbf{AbfA}}$ in the early induction periods (**Additional file 2: Table S8**), demonstrating that the unconventional splicing of *hacA* is required for UPR activation, but does not necessarily lead to *hacA* overexpression [22]. Liu et al. (2014) also observed that mRNA levels of the *hacA* were not increased in *A. oryzae* recombinant strains, indicating that *hacA* splicing is not the only mechanism responsible for activating UPR in this organism [27].

A. nidulans$_{\mathbf{AbfA}}$ also presented two exclusives up-regulated genes, the AN4583 and AN6145. Both genes are predicted as peptidyl-prolyl cis-trans isomerases (PPIases) [31]. These proteins accelerate folding of proteins, and these two PPIases may be interesting targets for genetic manipulation in order to improve protein secretion [54].

A. nidulans$_{\mathbf{1542}}$ showed only two exclusive DE genes. The first is a predicted 1,2-alpha-mannosidase (AN2045) with an ortholog (gh47-3) in *Neurospora crassa* reported to be an ER-

associated ubiquitin-dependent protein with catabolic process and N-glycan processing (according to UniProt). The second was AN8153 down-regulated at 2 and 8 h, which was not observed for the other strains which showed an up-regulation response at 2 h. This gene is predicted as ATP binding, with nucleoside-triphosphatase activity and nucleotide binding activity [31]. In *S. cerevisiae* the best hit is CDC48, an ATPase; it is a subunit of a polyubiquitin-selective segregase complex involved in ERAD, with a role in mobilizing membrane-bound transcription factors by regulating ubiquitin/proteasome-dependent processing in macroautophagy [42]. Both down-regulated genes (AN2045 and AN8153) were related to ERAD, and *A. nidulans*$_{\mathbf{1542}}$ had several DE genes in this pathway, as previously observed in the FunCat subcategory "Protein/peptide degradation" (**Figure 6**). Moreover, DE genes related to ERAD in this strain could partially explain the low levels of 1542 secretion. In *A. niger*, the knock-out of one ERAD factor improved the production of a heterologous beta-glucuronidase [55]. DE genes related to ERAD indicated by the *A. nidulans* recombinant strains studied in this work seem to be potential targets for genetic manipulation in an attempt to improve heterologous protein secretion.

Tyo et al. (2012) identified biological mechanisms which alter the secretory pathway in response to the secretion of recombinant proteins of different sizes, *i.e*. a small protein (human insulin precursor) and a comparatively larger protein (α-amylase) in *S. cerevisiae* [36]. By perturbing the yeast secretory pathway with these proteins, they conclude that in order to maintain the optimal secretion ratio, protein folding rates must increase and oxidation rates must decrease (by repressing both ERAD and UPR). In this scenario, recombinant proteins would be folded slowly and without high cellular stresses. This would result in a longer/prolonged overall process which may be required due to the difficulty of protein folding [36].

Furthermore, several studies [26,28,37,47] indicated that different recombinant proteins have different secretion levels in fungi. By observing the secretory pathway profile (**Figure 7**), it is clear that *A. nidulans* does not respond equally for all overexpressed proteins, but instead it regulates different processes to correctly fold and secret each recombinant protein.

CONCLUSIONS

Despite the importance of *Aspergilli* as microbial cell factories with many industrial applications, only a limited number of genetic tools are currently available for efficient overproduction and secretion of heterologous proteins [41]. In our study, the secretion of heterologous proteins in *A. nidulans* was investigated. We performed a comparative RNA-seq based transcriptomic analysis of *A. nidulans* recombinant strains overproducing heterologous glycoside hydrolases.

The general response to heterologous protein production was primarily an up-regulated response. Transcriptomic data were classified as early, continuous and late responses, and it was

possible to observe a different profile at 2 and 8 h of induced protein production. At 2 h, we observed the unconventional splicing of *hacA*, which was normalized at 8 h. A transcriptomic profile for each strain was established based on the functional categorization of DE genes. From this analysis several proteins, including many chaperones, were DE and appeared to be interesting candidates for genetic manipulation, requiring a deeper understanding of their effects on heterologous protein secretion by *A. nidulans.*

Finally, by observing secretory pathway genes we found 17 DE genes distributed in the three recombinant strains. The most representative functional categories of these genes were used to build a secretory pathway profile of the transcriptional response. This study addressed the response to heterologous protein production in *A. nidulans* by associating transcriptomic and monitoring of three different recombinant proteins which present different secretion levels. Thus, we may provide insights to better comprehend how *A. nidulans* adapts to the overproduction of heterologous proteins and highlight potential targets for genetic manipulation in order to obtain higher protein secretion yields.

MATERIALS AND METHODS

Strains maintenance and cultivation conditions

A. nidulans$_{\mathbf{A773}}$ (pyrG89;wA3;pyroA4) was purchased from the Fungal Genetics Stock Center (FGSC, St Louis, MO). The strains *A. nidulans*$_{\mathbf{AbfA}}$, *A. nidulans*$_{\mathbf{BglC}}$ and *A. nidulans*$_{\mathbf{1542}}$ were previously transformed and tested by our research group, as described by Segato 2012 [24]. 5-fluorotic acid (5-FOA) was purchased from Oakwood Products Inc (NC9639762), and all other chemicals were from Sigma–Aldrich, Megazyme and Fisher Scientific.

Vegetative cultures and conidia production were prepared by cultivation in minimal medium as described by Clutterbuck [56] and Pontecorvo [57]. *A. nidulans* strains were cultivated in minimal medium containing 5% of 20× Clutterbuck salts, 0.1% of 1000× vitamins, 0.1% of 1000× trace elements, pH 6.5 and supplemented with pyridoxine (1 mg/L), uracil/uridine (2.5 mg/L each) whenever required. Minimal medium was supplemented with 1% of glucose for mycelium growth (growth medium) or with 2% maltose and 250 mM HEPES buffer (Sigma–Aldrich) for promoter activation and protein production (induction medium). Incubation was carried out at 37 °C.

Production and secretion of client proteins

Pre-cultures were prepared by inoculating fresh 10^8 spores/ml in 30 mL of minimal medium supplemented with 1% (w/v) glucose for the mycelium growth for 48 h at 150 rpm and 37 °C. Strains cultivation were performed in biological triplicates. The mycelium was washed with autoclaved distilled water, filtered with Miracloth, and then transferred to the induction medium. After defined

periods for protein induction, the mycelium was collected, Miracloth filtered, dried, frozen in liquid nitrogen and stored at -80 °C. The medium was collected, centrifuged at 10,000 × g for 10 min prior to concentration by ultrafiltration (10 kDa cutoff Amicon), quantified by the Bradford method [58], purity was assessed by SDS-PAGE [59] and then used for biochemical studies.

Analytical assays

Enzymatic activity for beta-glucosidase GH3 (BglC) was performed with 4-nitrophenyl-β-D-glucopyranoside (Sigma–Aldrich) and for alpha-arabinofuranosidase GH51 (AbfA) with 4-nitrophenyl-α-L-arabinofuranoside (Sigma–Aldrich). The enzymatic activity assays were performed by adding 40 µL of extracellular crude extract to 50 µL of each 5 mM substrate in 10 µL of 100 mM ammonium acetate buffer pH 5.5 followed by incubation for 45 min at 50 ºC. The enzymatic reactions were stopped with sodium carbonate. The absorbance was read at 405 nm with a Multimode Infinite M200 Reader (Tecan, SC), and activity was calculated using p-nitrophenol as standard.

Enzymatic activity of thermophilic mannanase GH5 (1542) was conducted with locust bean gum (Sigma–Aldrich) as substrate. The reaction was performed by adding 1 µg of proteins from concentrated extracellular crude extract to 50 µL of 1% (w/v) substrate in 50 mM ammonium acetate buffer pH 5.5 under incubation for 10 h at 87 ºC. The release of reducing sugars was measured with DNS acid [60], and absorbance was read at 540 nm with a Multimode Infinite M200 Reader (Tecan, SC) and compared to a glucose standard curve. Statistics calculation and plotting were carried out with GraphPad Prism v 5.0 Software (California, US).

The concentration of maltose was detected by high-performance liquid chromatography (HPLC) Agilent Infinity 1260 with IR detector 50C, Aminex column HPX-87P 300 mm x 7.8 mm at 50 °C and 0.5 mL/min of ultrapure Milli-Q water as eluent phase.

Intracellular protein analysis

Intracellular proteins were extracted from frozen mycelium after grounding in liquid nitrogen to a fine powder using a mortar and a pestle. The powder was gently suspended in 5 mL of extraction buffer (20 mM Tris-HCl pH 8.0; 0.05 % (w/v) Triton™ X-100 (Sigma-Aldrich), and 150 mM NaCl, containing protease inhibitors (2 mM PMSF and Protease Inhibitor Cocktail N221-1mL from Amresco), and centrifuged twice at 8000 rpm for 15 min at 4 ºC; then, the supernatant was collected for SDS-PAGE [59] analysis.

RNA extraction

Total RNA was extracted from frozen mycelium using the Quick-RNA™ MiniPrep kit (Zymo Research) according to the manufacturer's instructions. The integrity of extracted RNA was evaluated

(RIN ≥ 8) on the Agilent Bioanalyzer 2100 and quantified using an ND-1000 NanoDrop (Thermo Scientific) spectrophotometer. Total purified RNA was stored at -80 °C until further processing.

RNA-seq data analysis

For RNA-seq sample preparation, the total RNA from the four *A. nidulans* strains (three recombinant strains and the parental strain) cultivated for 2 and 8 h in biological triplicates, resulted in twenty-four samples for library preparation, using TruSeq Stranded mRNA Library Prep Kit v2 (Illumina) according to the manufacturer's instructions. Sequencing libraries were prepared and sequenced using Illumina HiSeq 2500 at CTBE NGS sequencing facility. Approximately 148 million 100 bp paired-end reads were obtained for *A. nidulans*$_{\mathbf{A773}}$, *A. nidulans*$_{\mathbf{AbfA}}$, *A. nidulans*$_{\mathbf{SBglC}}$ strains and, 130 million reads were obtained for *A. nidulans*$_{\mathbf{1542}}$ strain, representing 115 GB. Reads were quality checked and filtered using FASTQC and Trimmomatic [61], respectively, and rRNA contamination was assessed and removed using sortmeRNA [62]. The rate of rRNA contamination was lower than 16% in all samples, except in one replicate of AbfA 8h (**Additional file 3: Table S9**). QC reads were aligned to *Aspergillus nidulans* genome available at AspGD (http://www.aspgd.org/) using Tophat2 [63], and concordant pair alignment rate varied between 84 and 93%. An acceptable replicate agreement was obtained for all conditions, as showed by the PCA analysis and clustering of the samples data.

For differential expression analysis, pairwise comparisons between each strain and the control strain *A. nidulans*$_{\mathbf{A773}}$ were performed using DESeq2 R/Bioconductor package [64], applying a log2 fold change ≥ 2 or ≤ −2 and an adjusted p-value ≤ 0.01 as thresholds. For the secretion pathway analysis, it was used a log2 fold change ≥ 1 or ≤ −1 and an adjusted p-value ≤ 0.01 as thresholds. Graph constructions were performed using GraphPad Prism v 5.0 Software (California, US) and the online tool for Venn diagrams - Venny 2.1 (http://bioinfogp.cnb.csic.es/tools/venny/).

Functional annotation

For functional annotation of DE genes identified by RNA-seq analysis, lists of genes were generated by filtering the results with log 2 FC ≤ -1.0 or ≥ 1.0 and adjusted p-value ≤ 0.01. The gene expression values of the recombinant strains were compared to the control strain *A. nidulans*$_{\mathbf{A773}}$. The gene lists obtained were then used for functional annotation using FunCat (Functional Catalogue) [37].

The selected genes for annotation with FunCat were faced with the database of the reference genome of *A. nidulans* and applied statistical Fisher test with p-value ≤ 0.05 as threshold. The enriched functional categories and subcategories were used for comparisons and graphics generation. Data analysis was performed using GraphPad Prism v 5.0 Software (California, US).

Secretion pathway analysis

A list of 374 secretory pathway genes was generated based on Liu et al. (2014), which defined a list with the functional protein secretory components from *A. oryzae* using the secretory model of *S. cerevisiae* as scaffold [27]. This list was further adapted to filamentous fungi by adding *A. oryzae* orthologous of the secretory components reported for other *Aspergillus* species, such as *A. nidulans* and *A. niger*. The homologous genes in *A. nidulans* reported in AspGD resources were then used to generate a list with 374 genes.

DE genes of the secretory pathway were functionally annotated with FunCat, and the most representative functional categories were used to build a secretory pathway profile of three recombinant strains. To generate this profile, the FC values of the DE genes involved with secretion pathway were used to create heatmaps using R.

Real-Time PCR (qPCR) analysis

To quantify the expression of selected genes, quantitative real-time PCR (qRT-PCR) was performed. For this analysis, 2 µg of pure RNA was used for cDNA synthesis using the First Strand cDNA kit Maxima™ Synthesis (Thermo Scientific), according to the manufacturer's instructions. The cDNA was 1/50 diluted and used for analysis in the ViiA™ 7 Real-Time PCR System, using Maxima SYBR™ Green (Thermo Scientific) for signal detection, in accordance with the manufacturer's instructions. The gene encoding β-tubulin (AN6838) was used as endogenous control for being a well stable gene in filamentous fungi [65].

The PCRs conditions were: 95 °C for 30 s; followed by 40 cycles at 95 °C for 10 s and 60 °C for 60 s. The melting curve was analyzed with ViiA™ 7 Software (Thermo Scientific) to confirm the presence of only one amplicon, according to the Tm expected for each gene. Gene expression values were calculated according to the 2-ΔΔCT method [66]. Primers used in the qPCR experiments are described in Additional file 4 (**Additional file 4: Table S10**). Data analysis was performed using GraphPad Prism v 5.0 Software (California, US).

Data availability

The datasets generated during and/or analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository, GSE101522, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE101522.

ACKNOWLEDGMENTS

This research was supported in part by FAPESP (grant 2012/20549-4 to ARLD; 2014/06923-6 to FMS). We are grateful to the National Council for Scientific and Technological Development (CNPq) for the financial support (441912/2014-1 to ARLD). MVR, MPZ and FC received FAPESP fellowships (2013/24988-5; 2014/15403-6; 2014/23051-2). We thank the LNBio Mass Spectrometry staff for the assistance with LC-MS/MS and the CTBE High Throughput Sequencing and Robotics Laboratory (NGS) for the assistance with RNAseq.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHOR'S CONTRIBUTIONS

ARLD conceived and designed the experiments. FC, MVR, MPZ participated in design of the study and performed the RNAseq experiments. GFP filtered the RNAseq raw data and generated the differential expressed genes lists. FC, MVR, MPZ, FJC, GFP and ARLD analyzed the RNAseq data. FC, FMS and FJC helped to interpret the experimental data. FC and ARLD drafted the manuscript. ARLD and GFP revised the manuscript. All authors read and approved the final manuscript.

REFERENCES

1. Lowe, R. G. T. & Howlett, B. J. Indifferent, affectionate, or deceitful: Lifestyles and secretomes of fungi. *PLoS Pathog.* **8,** 1–3 (2012).
2. Lubertozzi, D. & Keasling, J. D. Developing Aspergillus as a host for heterologous expression. *Biotechnol. Adv.* **27,** 53–75 (2009).
3. Reseach, B. Global Markets for Enzymes in Industrial Applications. *Focus Catal.* **1,** 2 (2014).
4. AMFEP. List of commercially available enzymes. *Assoc. Manuf. Formul. Enzym. Prod.* **Amfep/15/0,** http://www.amfep.org/content/list-enzymes (2015).
5. Su, X., Schmitz, G., Zhang, M., Mackie, R. I. & Cann, I. K. O. *Heterologous Gene Expression in Filamentous Fungi*. *Advances in Applied Microbiology* **81,** (Elsevier, 2012).
6. Guillemette, T. *et al.* Genomic analysis of the secretion stress response in the enzyme-producing cell factory Aspergillus niger. *BMC Genomics* **8,** 158 (2007).
7. Pakula, T. M. *et al.* Genome wide analysis of protein production load in Trichoderma reesei. *Biotechnol. Biofuels* **9,** 132 (2016).
8. Heimel, K. Unfolded protein response in filamentous fungi—implications in biotechnology. *Appl. Microbiol. Biotechnol.* **99,** 121–132 (2014).
9. Sidrauski, C. & Walter, P. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell* **90,** 1031–1039 (1997).
10. Perrone, C. A. A Novel Dynein Light Intermediate Chain Colocalizes with the Retrograde Motor for Intraflagellar Transport at Sites of Axoneme Assembly in Chlamydomonas and Mammalian Cells. *Mol. Biol. Cell* **14,** 2041–2056 (2003).
11. Ruegsegger, U., Leber, J. H., Walter, P. & Francisco, S. Block of HAC1 mRNA translation by long-range base pairing is released by cytoplasmic splicing upon induction of the unfolded protein response. *Cell* **107,** 103–114 (2001).
12. Calfon, M. *et al.* IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature* **415,** 92–96 (2002).
13. Iwawaki, T., Akai, R., Kohno, K. & Miura, M. A transgenic mouse model for monitoring endoplasmic reticulum stress. *Nat. Med.* **10,** 98–102 (2004).
14. Mulder, H. J., Saloheimo, M., Penttilä, M. & Madrid, S. M. The transcription factor HACA mediates the unfolded protein response in Aspergillus niger, and up-regulates its own transcription. *Mol. Genet. Genomics* **271,** 130–140 (2004).
15. Saloheimo, M., Valkonen, M. & Penttilä, M. Activation mechanisms of the HAC1-mediated unfolded protein response in filamentous fungi. *Mol. Microbiol.* **47,** 1149–1161 (2003).
16. Carvalho, N. D. *et al.* Genome-wide expression analysis upon constitutive activation of the

HacA bZIP transcription factor in Aspergillus niger reveals a coordinated cellular response to counteract ER stress. *BMC Genomics* **13,** 2–17 (2012).

17. Hayano, T., Hirose, M. & Kikuchi, M. Protein disulfide isomerase mutant lacking its isomerase activity accelerates protein folding in the cell. *FEBS Lett.* **377,** 505–511 (1995).
18. Valkonen, M., Penttilä, M. & Saloheimo, M. The ire1 and ptc2 genes involved in the unfolded protein response pathway in the filamentous fungus Trichoderma reesei. *Mol. Genet. Genomics* **272,** 443–51 (2004).
19. Xu, P., Raden, D., Doyle, F. J. & Robinson, A. S. Analysis of unfolded protein response during single-chain antibody expression in Saccaromyces cerevisiae reveals different roles for BiP and PDI in folding. *Metab. Eng.* **7,** 269–279 (2005).
20. Harmsen, M. M., Bruyne, M. I., Raué, H. A. & Maat, J. Overexpression of binding protein and disruption of the PMR1 gene synergistically stimulate secretion of bovine prochymosin but not plant Thaumatin in yeast. *Appl. Microbiol. Biotechnol.* **46,** 365–370 (1996).
21. Travers, K. J. *et al.* Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* **101,** 249–258 (2000).
22. Heimel, K. Unfolded protein response in filamentous fungi—implications in biotechnology. *Appl. Microbiol. Biotechnol.* **99,** 121–132 (2015).
23. Nevalainen, K. M. H., Te'o, V. S. J. & Bergquist, P. L. Heterologous protein expression in filamentous fungi. *Trends Biotechnol.* **23,** 468–474 (2005).
24. Segato, F. *et al.* High-yield secretion of multiple client proteins in Aspergillus. *Enzyme Microb. Technol.* **51,** 100–106 (2012).
25. Sims, A. H. *et al.* Transcriptome Analysis of Recombinant Protein Secretion by Aspergillus nidulans and the Unfolded-Protein Response In Vivo. *Appl. Environ. Microbiol.* **71,** 2737–2747 (2005).
26. Pakula, T. M. *et al.* The Effects of Drugs Inhibiting Protein Secretion in the Filamentous Fungus Trichoderma reesei. *J. Biol. Chem.* **278,** 45011–45020 (2003).
27. Liu, L., Feizi, A., Österlund, T., Hjort, C. & Nielsen, J. Genome-scale analysis of the high-efficient protein secretion system of Aspergillus oryzae. *BMC Syst. Biol.* **8,** 73 (2014).
28. Nevalainen, H. & Peterson, R. Making recombinant proteins in filamentous fungi- are we expecting too much? *Front. Microbiol.* **5,** 75 (2014).
29. Guillemette, T. *et al.* Methods for Investigating the UPR in Filamentous Fungi. in *Methods in enzymology* **490,** 1–29 (2011).
30. Wood, V. *et al.* PomBase: a comprehensive online resource for fission yeast. *Nucleic Acids Res.* **40,** D695–D699 (2012).
31. Cerqueira, G. C. *et al.* The Aspergillus Genome Database: Multispecies curation and

incorporation of RNA-Seq data to improve structural gene annotations. *Nucleic Acids Res.* **42,** D705–D710 (2014).

32. Szilagyi, M. *et al.* Transcriptome changes initiated by carbon starvation in Aspergillus nidulans. *Microbiology* **159,** 176–190 (2013).
33. Wang, G., Zhang, D. & Chen, S. Effect of earlier unfolded protein response and efficient protein disposal system on cellulase production in Rut C30. *World J. Microbiol. Biotechnol.* **30,** 2587–2595 (2014).
34. Zhang, J. *et al.* Expression and high-yield production of extremely thermostable bacterial xylanaseB in Aspergillus niger. *Enzyme Microb. Technol.* **43,** 513–516 (2008).
35. Zoglowek, M., Lübeck, P. S., Ahring, B. K. & Lübeck, M. Heterologous expression of cellobiohydrolases in filamentous fungi - An update on the current challenges, achievements and perspectives. *Process Biochem.* **50,** 211–220 (2015).
36. Tyo, K. E., Liu, Z., Petranovic, D. & Nielsen, J. Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress. *BMC Biol.* **10,** 16 (2012).
37. Ruepp, A. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.* **32,** 5539–5545 (2004).
38. Kwon, M. J. *et al.* The transcriptomic fingerprint of glucoamylase over-expression in Aspergillus niger. *BMC Genomics* **13,** 701 (2012).
39. Zhou, B., Xie, J., Liu, X., Wang, B. & Pan, L. Functional and transcriptomic analysis of the key unfolded protein response transcription factor HacA in Aspergillus oryzae. *Gene* **593,** 143–153 (2016).
40. Mattanovich, D., Gasser, B., Hohenblum, H. & Sauer, M. Stress in recombinant protein producing yeasts. *J. Biotechnol.* **113,** 121–135 (2004).
41. Fleißner, A. & Dersch, P. Expression and export: recombinant protein production systems for Aspergillus. *Appl. Microbiol. Biotechnol.* **87,** 1255–1270 (2010).
42. Cherry, J. M. *et al.* Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.* **40,** D700–D705 (2012).
43. Hampton, R. Y. ER-associated degradation in protein quality control and cellular regulation. *Curr. Opin. Cell Biol.* **14,** 476–482 (2002).
44. Hamilton, S. R. & Gerngross, T. U. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Current Opinion in Biotechnology* **18,** 387–392 (2007).
45. Kimura, S., Maruyama, J. ichi, Kikuma, T., Arioka, M. & Kitamoto, K. Autophagy delivers misfolded secretory proteins accumulated in endoplasmic reticulum to vacuoles in the filamentous fungus Aspergillus oryzae. *Biochem. Biophys. Res. Commun.* **406,** 464–470 (2011).

46. de Groot, P. W. J. *et al.* Comprehensive genomic analysis of cell wall genes in Aspergillus nidulans. *Fungal Genet. Biol.* **46,** S72–S81 (2009).

47. Burggraaf, A.-M. M. & Ram, A. F. J. Autophagy is dispensable to overcome ER stress in the filamentous fungus Aspergillus niger. *Microbiologyopen* **5,** 647–658 (2016).

48. Malavazi, I., Goldman, G. H. & Brown, N. A. The importance of connections between the cell wall integrity pathway and the unfolded protein response in filamentous fungi. *Brief. Funct. Genomics* **13,** 456–470 (2014).

49. Fiedler, M. R. *et al.* The capacity of Aspergillus niger to sense and respond to cell wall stress requires at least three transcription factors: RlmA, MsnA and CrzA. *Fungal Biol. Biotechnol.* **1,** 5 (2014).

50. Yoon, J., Aishan, T., Maruyama, J. I. & Kitamoto, K. Enhanced production and secretion of heterologous proteins by the filamentous fungus Aspergillus oryzae via disruption of vacuolar protein sorting receptor gene AoVps10. *Appl. Environ. Microbiol.* **76,** 5718–5727 (2010).

51. Schalén, M., Anyaogu, D. C., Hoof, J. B. & Workman, M. Effect of secretory pathway gene overexpression on secretion of a fluorescent reporter protein in Aspergillus nidulans. *Fungal Biol. Biotechnol.* **3,** 3 (2016).

52. Al-Sheikh, H. *et al.* Endoplasmic reticulum stress leads to the selective transcriptional downregulation of the glucoamylase gene in Aspergillus niger. *Mol. Microbiol.* **53,** 1731–1742 (2004).

53. Krishnan, K. *et al.* Polysome profiling reveals broad translatome remodeling during endoplasmic reticulum (ER) stress in the pathogenic fungus Aspergillus fumigatus. *BMC Genomics* **15,** 159 (2014).

54. Pemberton, T. J. Identification and comparative analysis of sixteen fungal peptidyl-prolyl *cis/trans* isomerase repertoires. *BMC Genomics* **7,** 244 (2006).

55. Jacobs, D. I. *et al.* Effective lead selection for improved protein production in Aspergillus niger based on integrated genomics. *Fungal Genet. Biol.* **46,** 141–152 (2009).

56. Clutterbuck, A. J. Sexual and parasexual genetics of Aspergillus species. *Biotechnology* **23,** 3–18 (1992).

57. Pontecorvo, G., Roper, J. A., Chemmons, L. M., Macdonald, K. D. & Bufton, A. W. J. The Genetics of Aspergillus nidulans. in *Advances in Genetics* **5,** 141–238 (1953).

58. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–254 (1976).

59. Shapiro, a L. & Maizel, J. V. Molecular weight estimation of polypeptides by SDS-polyacrylamide gel electrophoresis: further data concerning resolving power and general considerations. *Anal. Biochem.* **29,** 505–514 (1969).

60. Miller, G. L. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal. Chem.* **31,** 426–428 (1959).

61. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30,** 2114–2120 (2014).

62. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28,** 3211–3217 (2012).

63. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36 (2013).

64. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).

65. Llanos, A., François, J. M. & Parrou, J. Tracking the best reference genes for RT-qPCR data normalization in filamentous fungi. *BMC Genomics* **16,** 71 (2015).

66. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25,** 402–408 (2001).

FIGURE CAPTIONS

Figure 1. Physiological parameters and heterologous protein production by *A. nidulans* recombinant strains. A) *A. nidulans*$_{\mathbf{AbfA}}$; **B)** *A. nidulans*$_{\mathbf{BglC}}$; **C)** *A. nidulans*$_{\mathbf{1542}}$. Physiological parameters (top): mycelium dry weight (g), final pH of the cultivation medium, extracellular proteins (mg/mL). Enzymatic activity of the *A. nidulans* strains in different cultivation periods (center). The highest enzymatic activity in each cultivation was considered 100%. SDS-PAGE of the heterologous proteins secreted by the recombinant *A. nidulans* strains in different cultivation periods (bottom). On each lane, 10 µg of proteins were applied. The target proteins are indicated by an asterisk. The red box indicates a native intracellular beta-glucosidase detected after 48 h of cultivation. The band corresponding to the 1542 protein was not detected in the SDS-PAGE (bottom at right). Error bars indicate the standard deviation of three replicates.

Figure 2. Gene expression profile of *A. nidulans* recombinant strains. Expression levels were determined by qPCR in different cultivation periods and expressed on a logarithmic scale as folding changed. **A)** Gene expression profile of the heterologous genes *abfA*, *bglC* and *1542*. **B)** Expression levels of *bipA* in *A. nidulans*$_{\mathbf{AbfA}}$, *A. nidulans*$_{\mathbf{BglC}}$ and *A. nidulans*$_{\mathbf{1542}}$. **C)** Expression levels of sec63 in *A. nidulans*$_{\mathbf{AbfA}}$, *A. nidulans*$_{\mathbf{BglC}}$ and *A. nidulans*$_{\mathbf{1542}}$. *tubC* was used as the reference gene and the strain *A. nidulans*$_{\mathbf{A773}}$ as the experimental control. The ΔΔCt method was used for gene expression calculation. Asterisks indicate statistically different results (two-way ANOVA and Bonferroni post-test p-value<0.05). The strains were cultivated as described in the Materials and Methods section. Results are expressed as mean values of three biological replicates and the error bars indicate standard deviation.

Figure 3. Global analysis of differentially expressed genes. A) Bar chart representing differentially expressed (DE) genes by the *A. nidulans* recombinant strains from transcriptomic data in different cultivation periods. Rectangles above the bars present the number of up and downregulated genes. RNA-seq data of each recombinant strain were compared to the control strain *A. nidulans*$_{\mathbf{A773}}$ and (log 2) fold change ≤ -2.0 or ≥ 2.0, and adjusted p-values ≤ 0.01 were used as thresholds. In general, 476 DE genes were identified in all strains and at the analyzed time points. **B)** Venn diagram representing DE genes analyzed by the RNA-seq at 2 and 8 h of cultivation for the three *A. nidulans* recombinant strains. The Venn diagram was created using the online tool Venny 2.1.

Figure 4. Functional annotation of the transcriptomic data generated by RNA-seq analysis of *A. nidulans* recombinant strains. MIPS FunCat term annotations of *A. nidulans*$_{\mathbf{1542}}$, *A. nidulans*$_{\mathbf{AbfA}}$

and *A. nidulans*$_{\mathbf{BglC}}$ show a broad variety of biological processes associated with the DE genes of the *A. nidulans* recombinant strains in different cultivation periods. Analysis of the functional categories from global transcriptomic data was performed with Funcat (The Functional Catalogue). Functional categories of the heatmap represent the number of DE genes found in the transcriptomic data based on the reference genome annotation in FunCat. The heatmap was created with the online software from the Broad Institute – Morpheus.

Figure 5. **Functional annotation of differentially expressed genes of early, late and continuous responses.** Functional annotation of differentially expressed (DE) genes in the recombinant *A. nidulans* strains. The early response corresponds to DE genes exclusively found at 2 h; continuous response, DE genes found at both 2 and 8 h; and late response, DE genes exclusively found at 8h. The analysis of functional categories was performed with FunCat (The Functional Catalogue). The functional categories graph values represent the number of altered genes found (abs set) in the transcriptomic data based on the reference genome annotation.

Figure 6. **Profile of differentially expressed genes related to protein production pathways.** Heatmaps and Venn diagrams of FunCat annotated genes in *A. nidulans* were compared to RNAseq data of the three recombinant strains for the functional subcategories: **A)** 32.01.07 Unfolded protein response (*e.g*. ER quality control). **B)** 14.13 Protein/peptide degradation. **C)** N-directed glycosylation, deglycosylation. Each group was annotated with FunCat (The Functional Catalogue). Venn diagrams were created with the online tool Venny 2.1. Heatmaps were created with online software from the Broad Institute – Morpheus.

Figure 7. Secretory pathway analysis of the differentially expressed genes obtained from the RNA-seq data. A) Heatmap representing the differentially expressed genes related to the secretion pathway in *A. nidulans*. Data from recombinant strains were compared to the control strain *A. nidulans*$_{\mathbf{A773}}$ and (log 2) Fold Change ≤-1.0 or ≥1.0, and adjusted p-values ≤ 0.01 were used as thresholds. **B)** Venn diagram representing the 17 DE genes related to the secretion pathway. No gene was found in common to all the recombinant strains. The Venn diagram was created using the online tool Venny 2.1. **C)** Functional annotations of DE genes related to the secretory pathway of recombinant *A. nidulans*$_{\mathbf{AbfA}}$, *A. nidulans*$_{\mathbf{BglC}}$ and *A. nidulans*$_{\mathbf{1542}}$. The analysis of functional categories was performed with FunCat (The Functional Catalogue). The functional category graph values represent the number of altered genes found (abs set) in the transcriptomic data based on values of

the reference genome annotation. All functional subcategories with at least five DE genes are represented. Heatmaps were created with online software from the Broad Institute – Morpheus.

Figure 8. Illustrative hyphae tip as a representative model of the secretion pathway profile based on the transcriptomic response of the recombinant *A. nidulans* strains. The most representative functional categories representing the 17 DE genes were used to build a secretory pathway profile of the *A. nidulans*$_{\mathbf{AbfA}}$, *A. nidulans*$_{\mathbf{BglC}}$ and *A. nidulans*$_{\mathbf{1542}}$ recombinant strains. Heatmaps were constructed based on fold change values of the differentially expressed genes using R.

Figure 1

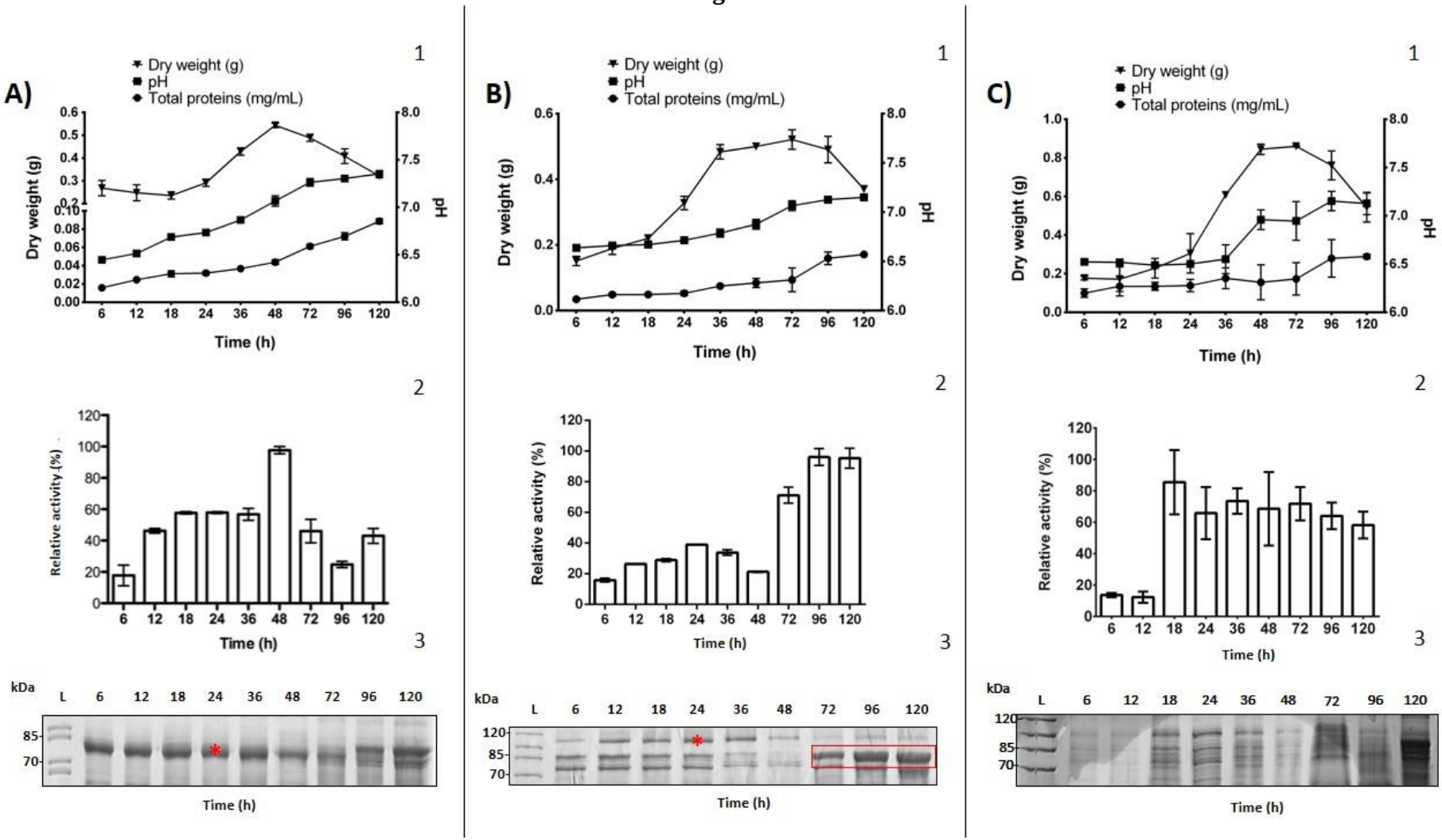

A)
B)
C)
1
2
3
Dry weight (g)
pH
Total proteins (mg/mL)
Time (h)
Relative activity (%)
kDa
L
6 12 18 24 36 48 72 96 120
85
70
120

Figure 2

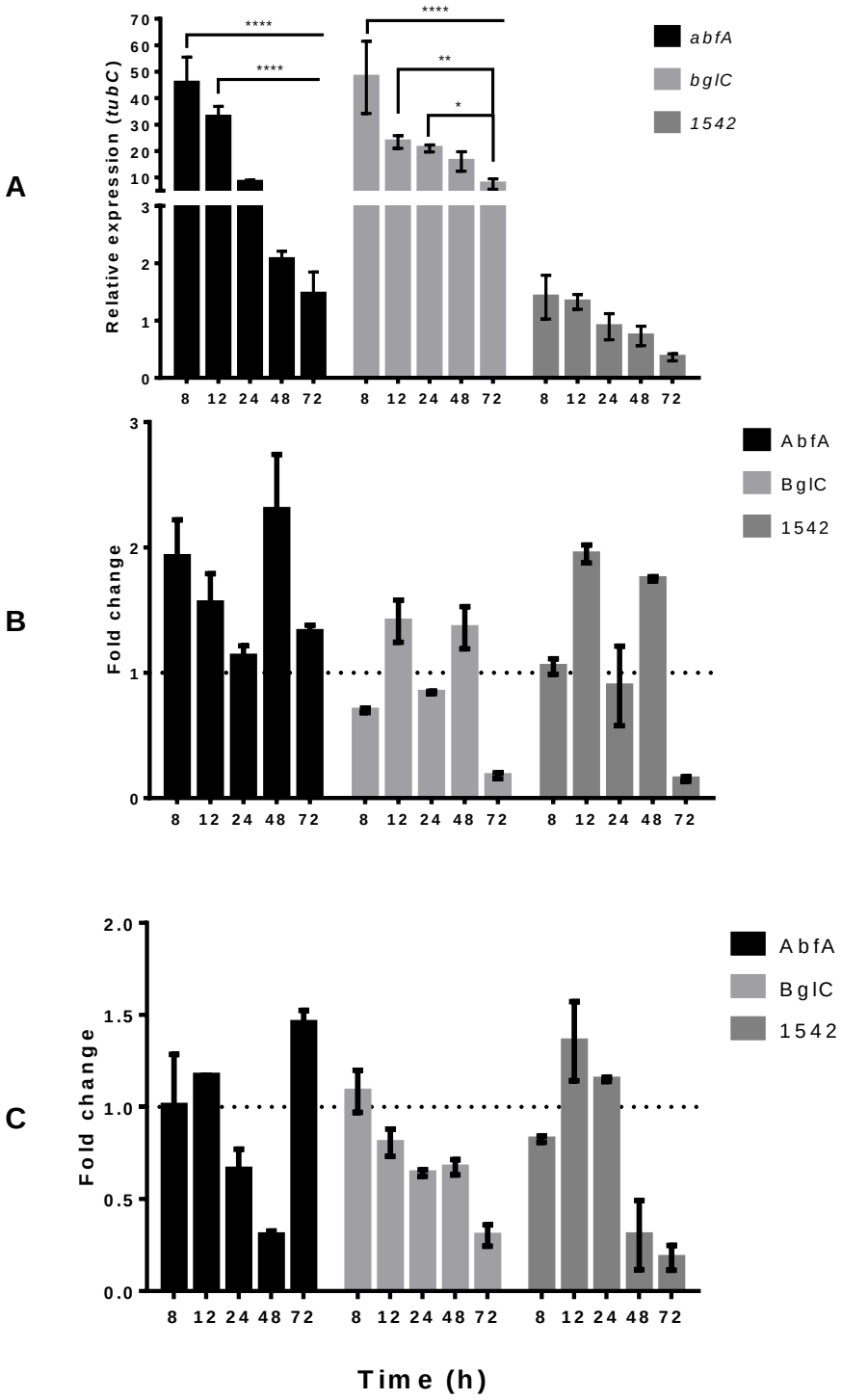

A
Relative expression (tubC)
abfA
bglC
1542
****
****
****
**
*
B
Fold change
AbfA
BglC
1542
C
Fold change
AbfA
BglC
1542
8 12 24 48 72
Time (h)

Figure 3

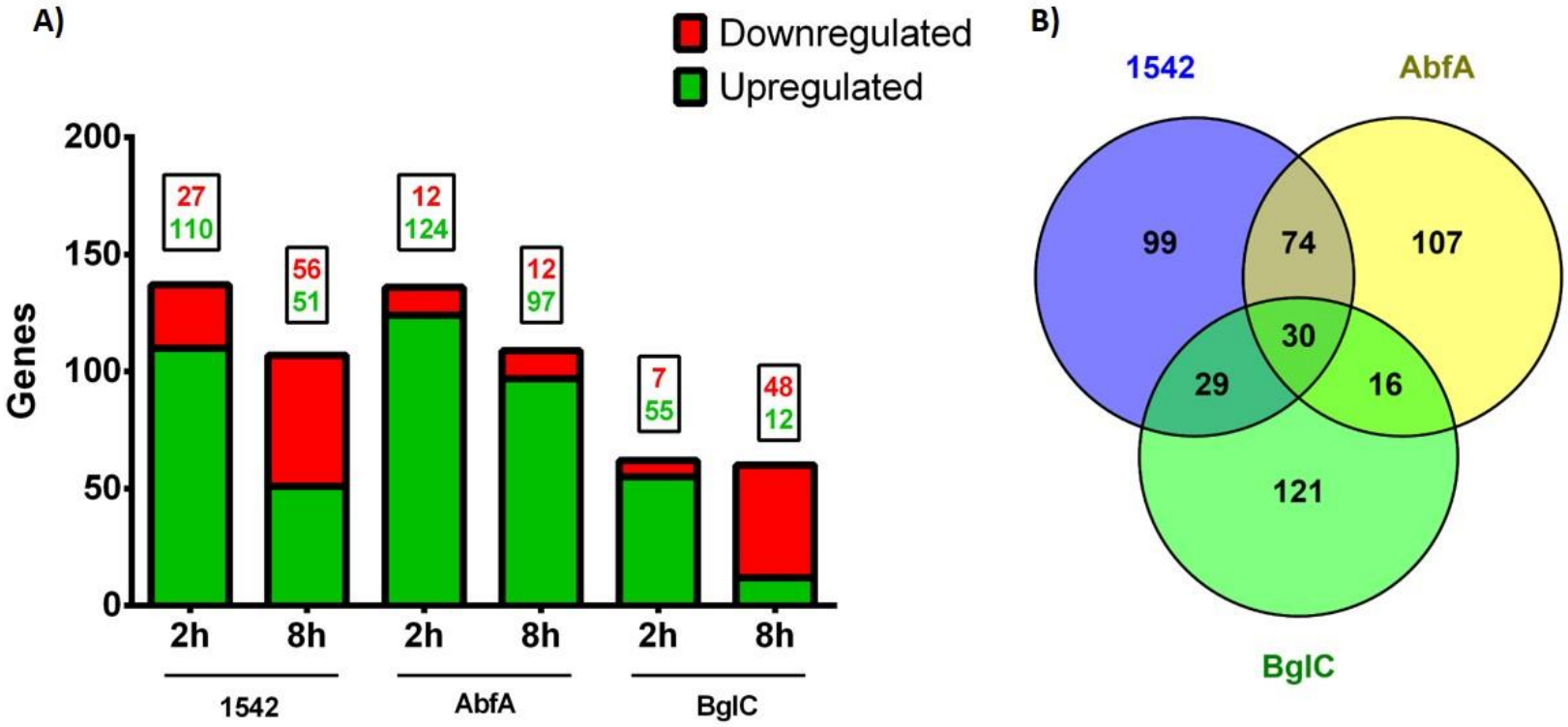

A)
Downregulated
Upregulated
Genes
200
150
100
50
0
27
110
56
51
12
124
12
97
7
55
48
12
2h
8h
2h
8h
2h
8h
1542
AbfA
BglC
B)
1542
AbfA
BglC
99
74
107
30
29
16
121

Figure 4

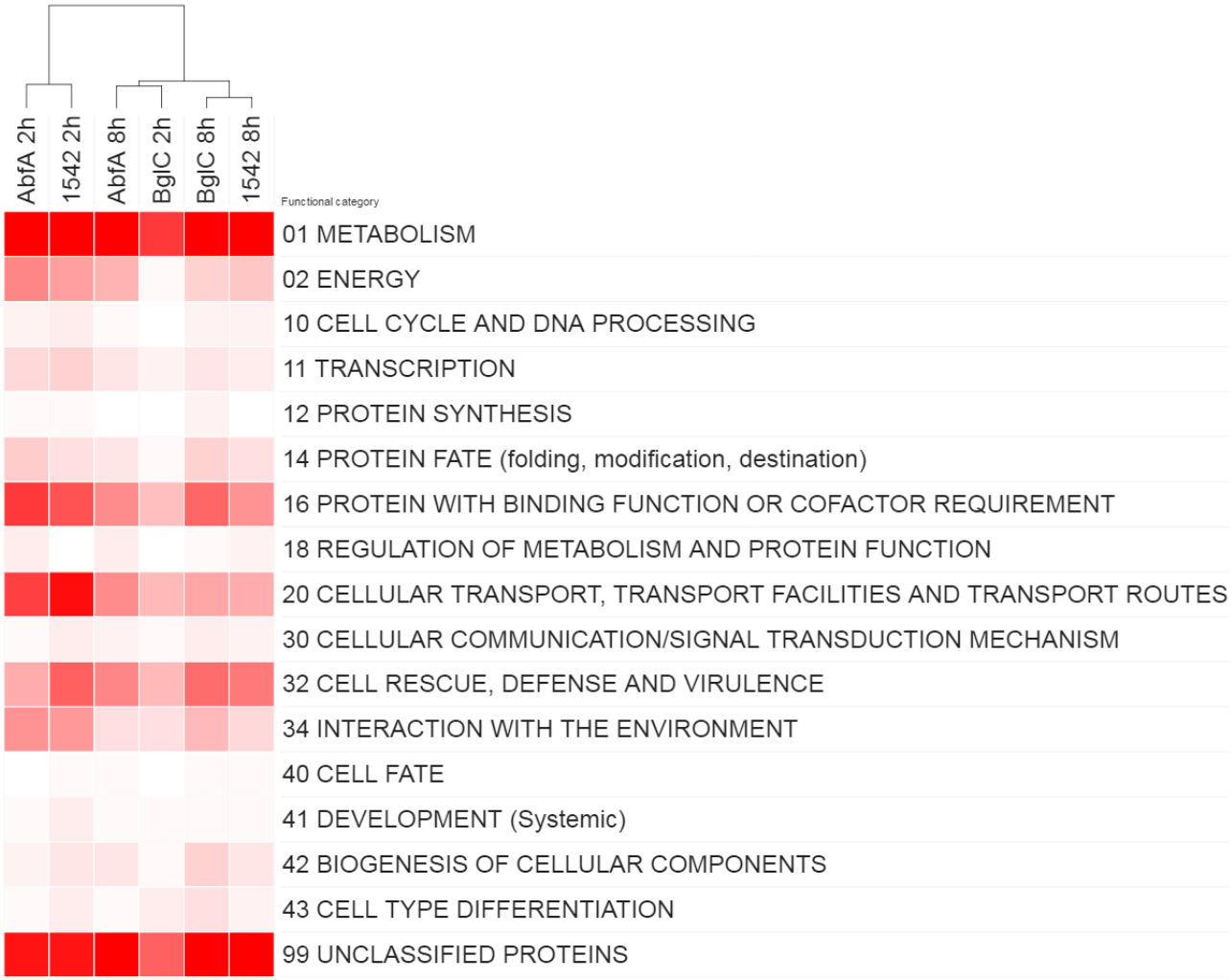

AbfA 2h
1542 2h
AbfA 8h
BglC 2h
BglC 8h
1542 8h
Functional category
01 METABOLISM
02 ENERGY
10 CELL CYCLE AND DNA PROCESSING
11 TRANSCRIPTION
12 PROTEIN SYNTHESIS
14 PROTEIN FATE (folding, modification, destination)
16 PROTEIN WITH BINDING FUNCTION OR COFACTOR REQUIREMENT
18 REGULATION OF METABOLISM AND PROTEIN FUNCTION
20 CELLULAR TRANSPORT, TRANSPORT FACILITIES AND TRANSPORT ROUTES
30 CELLULAR COMMUNICATION/SIGNAL TRANSDUCTION MECHANISM
32 CELL RESCUE, DEFENSE AND VIRULENCE
34 INTERACTION WITH THE ENVIRONMENT
40 CELL FATE
41 DEVELOPMENT (Systemic)
42 BIOGENESIS OF CELLULAR COMPONENTS
43 CELL TYPE DIFFERENTIATION
99 UNCLASSIFIED PROTEINS

Figure 5

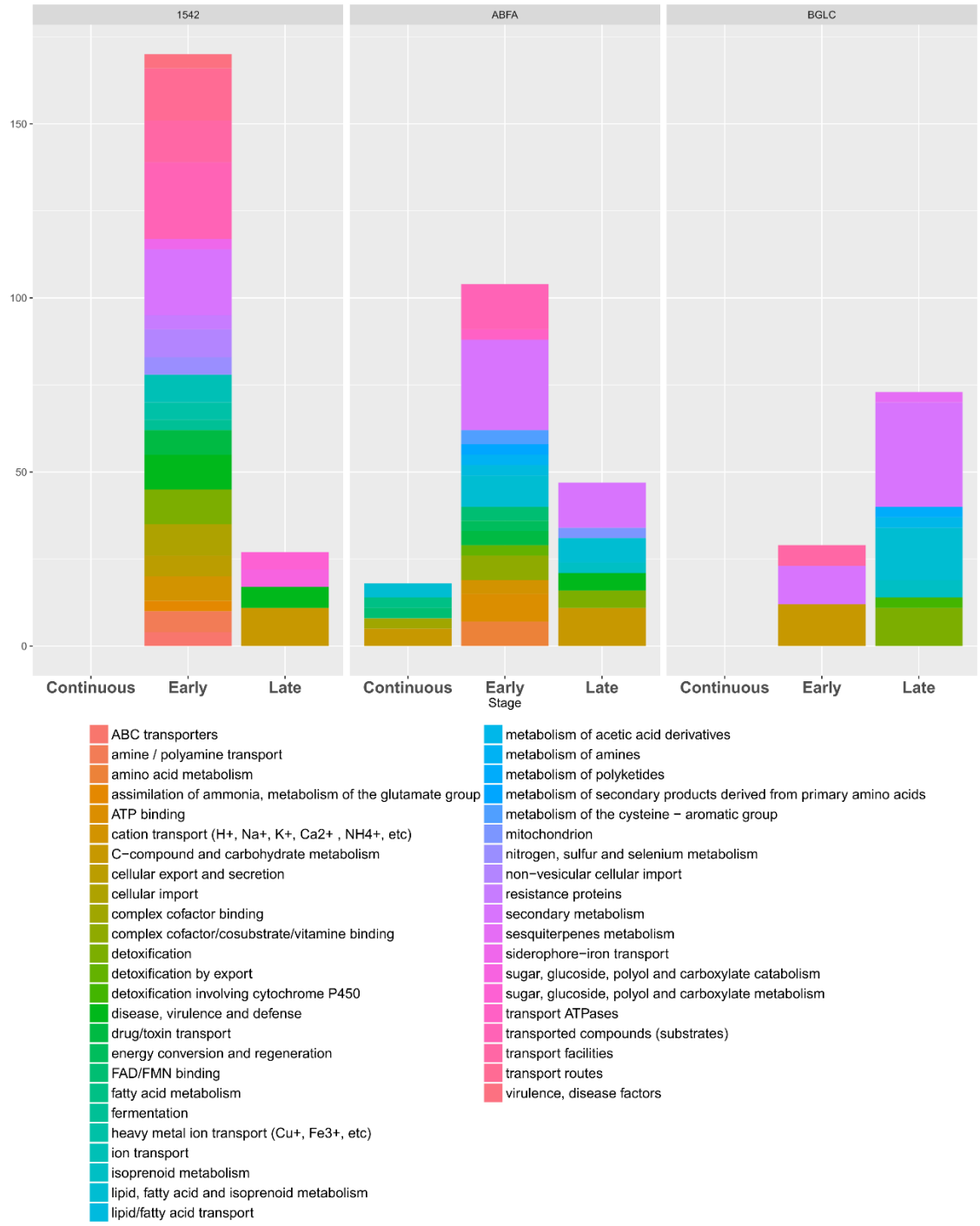

1542
ABFA
BGLC
150
100
50
0
Continuous
Early
Late
Stage
ABC transporters
amine / polyamine transport
amino acid metabolism
assimilation of ammonia, metabolism of the glutamate group
ATP binding
cation transport (H+, Na+, K+, Ca2+ , NH4+, etc)
C−compound and carbohydrate metabolism
cellular export and secretion
cellular import
complex cofactor binding
complex cofactor/cosubstrate/vitamine binding
detoxification
detoxification by export
detoxification involving cytochrome P450
disease, virulence and defense
drug/toxin transport
energy conversion and regeneration
FAD/FMN binding
fatty acid metabolism
fermentation
heavy metal ion transport (Cu+, Fe3+, etc)
ion transport
isoprenoid metabolism
lipid, fatty acid and isoprenoid metabolism
lipid/fatty acid transport
metabolism of acetic acid derivatives
metabolism of amines
metabolism of polyketides
metabolism of secondary products derived from primary amino acids
metabolism of the cysteine − aromatic group
mitochondrion
nitrogen, sulfur and selenium metabolism
non−vesicular cellular import
resistance proteins
secondary metabolism
sesquiterpenes metabolism
siderophore−iron transport
sugar, glucoside, polyol and carboxylate catabolism
sugar, glucoside, polyol and carboxylate metabolism
transport ATPases
transported compounds (substrates)
transport facilities
transport routes
virulence, disease factors

Figure 6

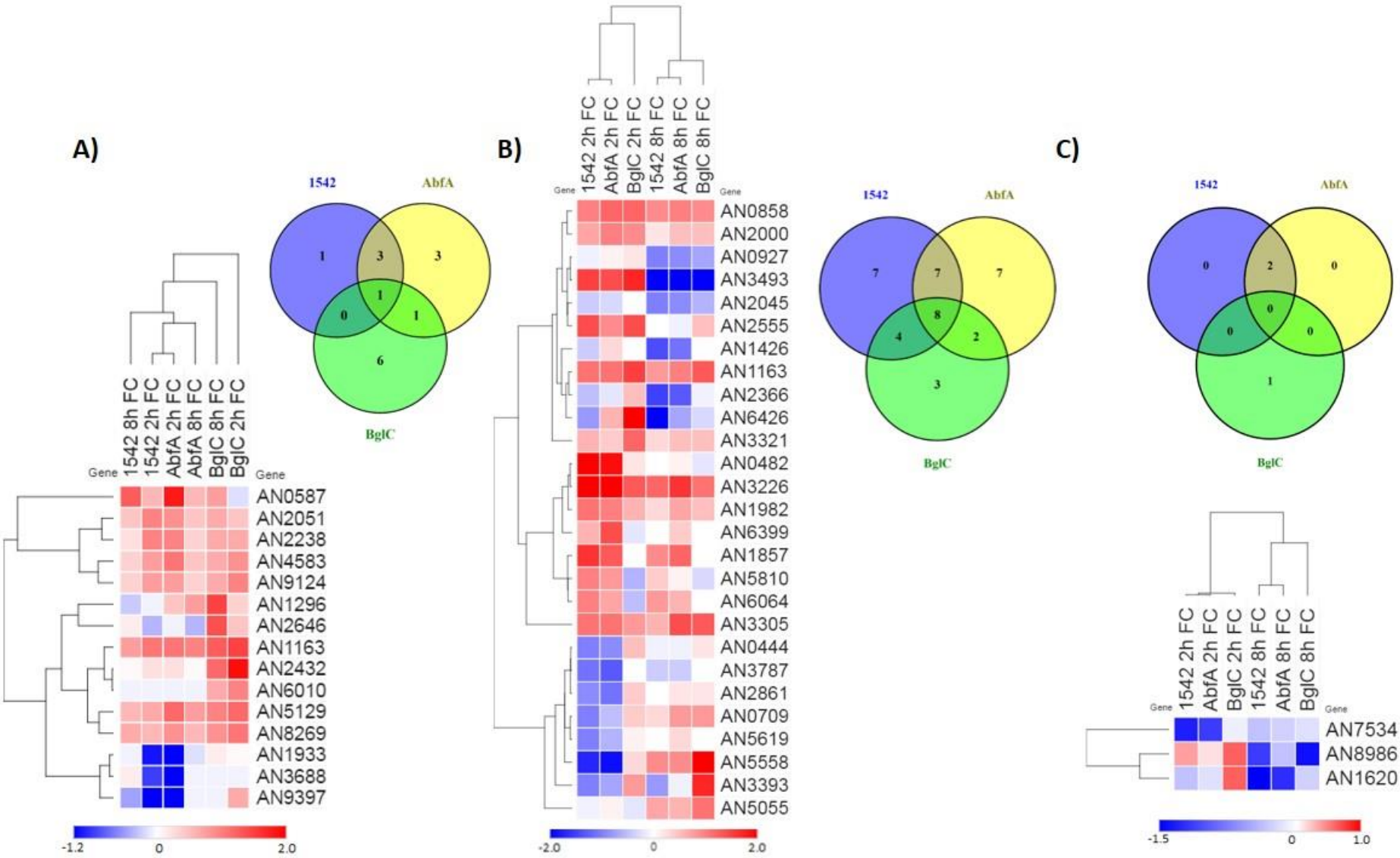


Figure 7

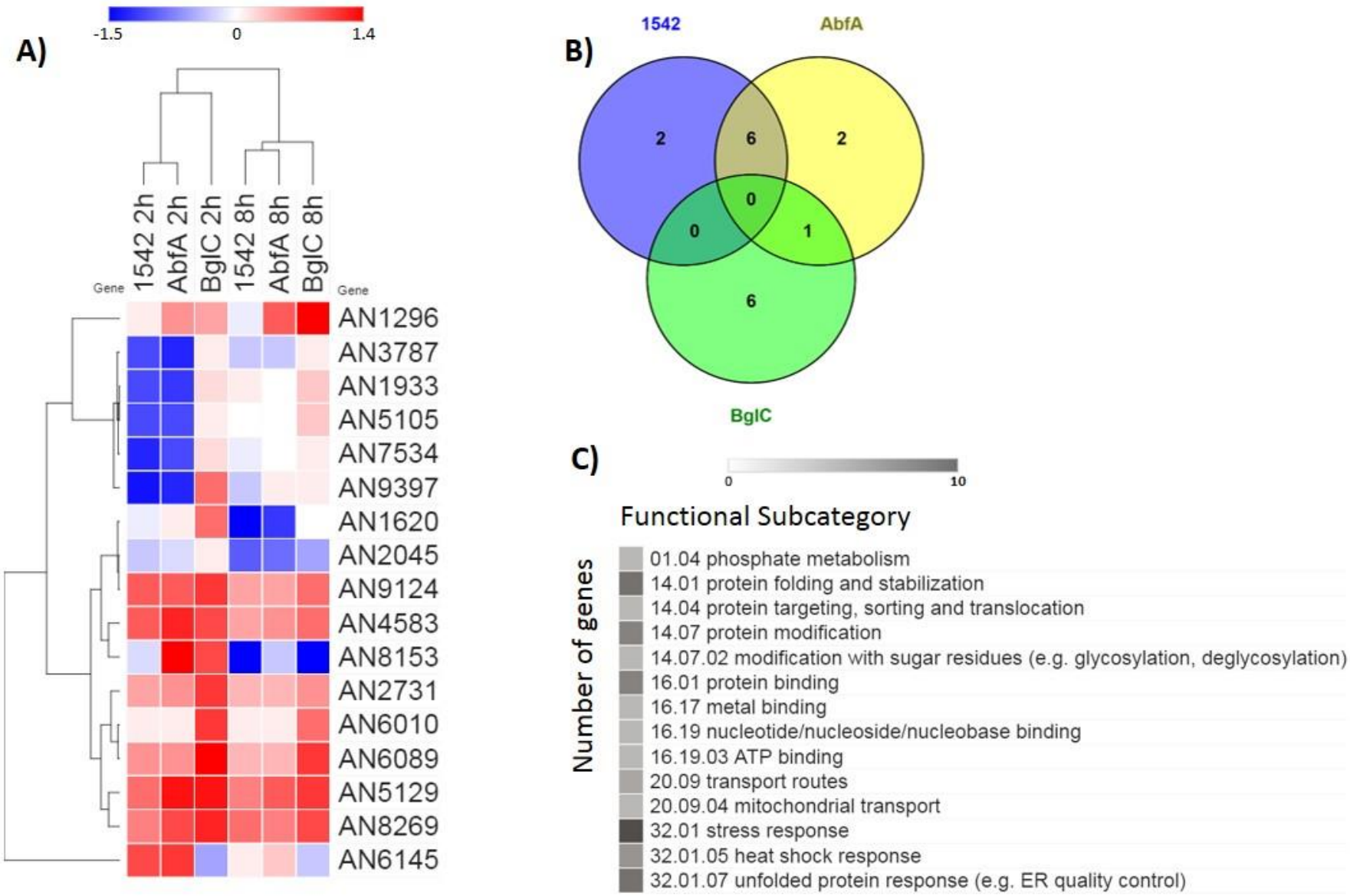
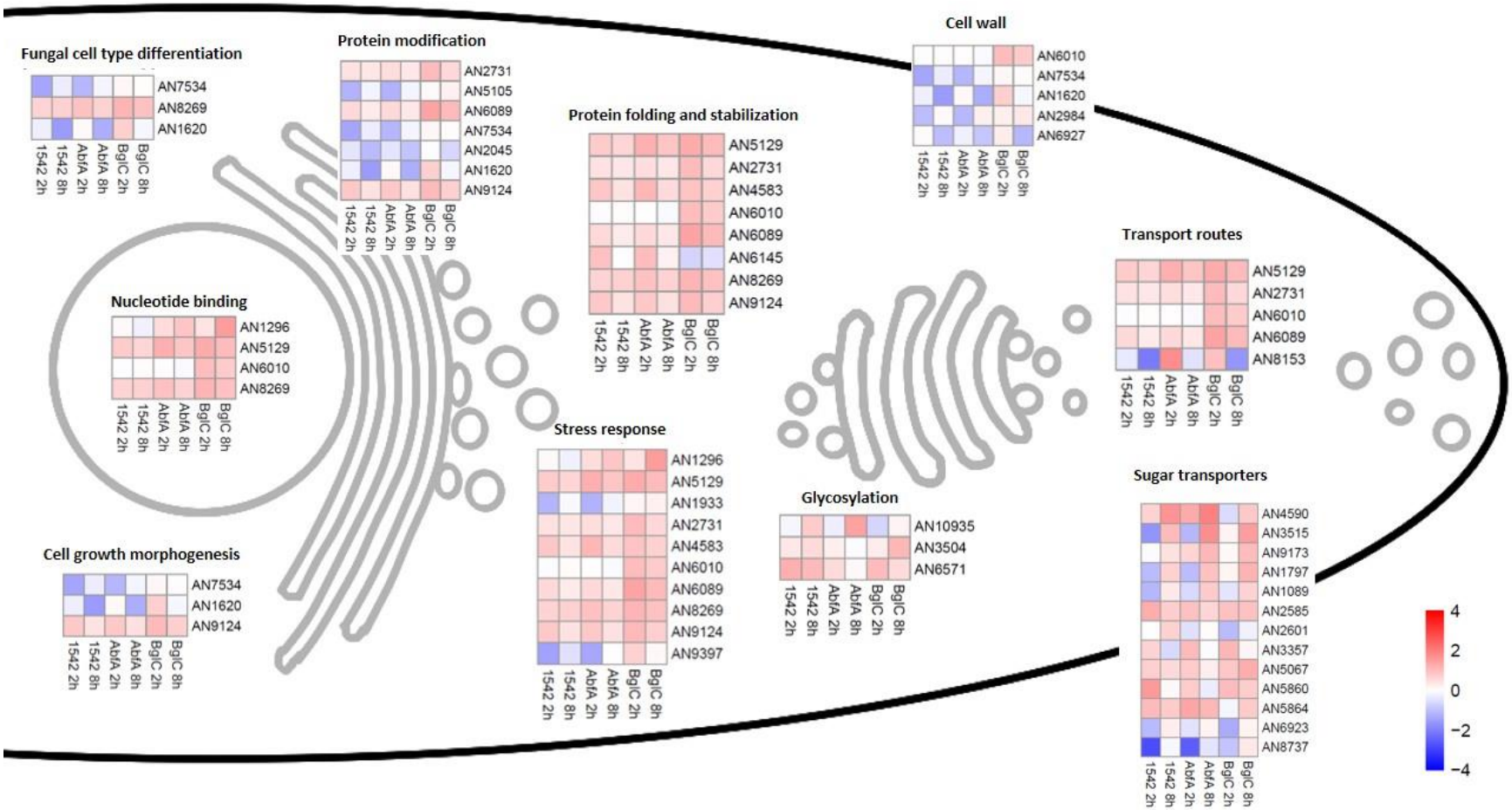

A)
-1.5
0
1.4
1542 2h
AbfA 2h
BglC 2h
1542 8h
AbfA 8h
BglC 8h
Gene
Gene
AN1296
AN3787
AN1933
AN5105
AN7534
AN9397
AN1620
AN2045
AN9124
AN4583
AN8153
AN2731
AN6010
AN6089
AN5129
AN8269
AN6145
B)
1542
AbfA
BglC
2
6
2
0
0
1
6
C)
0
10
Functional Subcategory
Number of genes
01.04 phosphate metabolism
14.01 protein folding and stabilization
14.04 protein targeting, sorting and translocation
14.07 protein modification
14.07.02 modification with sugar residues (e.g. glycosylation, deglycosylation)
16.01 protein binding
16.17 metal binding
16.19 nucleotide/nucleoside/nucleobase binding
16.19.03 ATP binding
20.09 transport routes
20.09.04 mitochondrial transport
32.01 stress response
32.01.05 heat shock response
32.01.07 unfolded protein response (e.g. ER quality control)

Figure 8


Fungal cell type differentiation
AN7534
AN8269
AN1620
1542 2h
1542 8h
AbfA 2h
AbfA 8h
BglC 2h
BglC 8h
Protein modification
AN2731
AN5105
AN6089
AN7534
AN2045
AN1620
AN9124
Protein folding and stabilization
AN5129
AN2731
AN4583
AN6010
AN6089
AN6145
AN8269
AN9124
Cell wall
AN6010
AN7534
AN1620
AN2984
AN6927
Nucleotide binding
AN1296
AN5129
AN6010
AN8269
Transport routes
AN5129
AN2731
AN6010
AN6089
AN8153
Stress response
AN1296
AN5129
AN1933
AN2731
AN4583
AN6010
AN6089
AN8269
AN9124
AN9397
Glycosylation
AN10935
AN3504
AN6571
Sugar transporters
AN4590
AN3515
AN9173
AN1797
AN1089
AN2585
AN2601
AN3357
AN5067
AN5860
AN5864
AN6923
AN8737
Cell growth morphogenesis
AN7534
AN1620
AN9124
4
2
0
−2
−4