Published in final edited form as:

J Proteome Res. 2012 December 7; 11(12): 5827–5835. doi:10.1021/pr300604s.

In-depth Analysis of the *Magnaporthe oryzae* Conidial Proteome

Emine Gokce^{1,*}, William L. Franck^{2,*}, Yeonyee Oh², Ralph A. Dean², and David C. Muddiman^{1,*}

¹W.M. Keck FT-ICR Mass Spectrometry Laboratory, Department of Chemistry, North Carolina State University, Raleigh, North Carolina, 27695

²Center for Integrated Fungal Research, North Carolina State University, Raleigh, North Carolina, 27695

Abstract

The filamentous fungus *Magnaporthe oryzae* (*M. oryzae*) is the causative agent of rice blast disease and presents a significant threat to worldwide rice production. To establish the groundwork for future research on the pathogenic development of *M. oryzae*, a global proteomic study of conidia was performed. The filter aided sample preparation method (FASP) and anion StageTip fractionation combined with long, optimized shallow 210 min nanoLC gradients prior to mass spectrometry analysis on an Orbitrap XL was applied, which resulted in a doubling of protein identifications in comparison to our previous GeLC analysis. Herein, we report the identification of 2912 conidial proteins at a 1% protein false discovery rate (FDR) and we present the most extensive study performed on *M. oryzae* conidia to date. A similar distribution between identified proteins and the predicted proteome was observed when subcellular localization analysis was performed, suggesting the detected proteins build a representative portion of the predicted proteome. A higher percentage of cytoplasmic proteins (associated with translation, energy and metabolism) were observed in the conidial proteome relative to the whole predicted proteome. Conversely, nuclear and extracellular proteins were less well represented in the conidial proteome. Further analysis by gene ontology revealed biological insights into identified proteins important for central metabolic processes and the physiology of conidia.

Keywords

proteomics; mass spectrometry; *Magnaporthe oryzae*; FASP; statistical tests; differential protein expression

1. Introduction

The filamentous fungus *Magnaporthe oryzae* (*M. oryzae*) is responsible for the destruction of millions of hectares of rice every year and impacts all rice growing regions of the world¹⁻³. In many Asian countries rice is a primary source of nutrition and yield losses associated with rice blast disease pose significant current and future risks to rice production. The lack of durable resistance genes in rice necessitates more effective control strategies to prevent fungal infection of rice crops. More in depth studies of the pathosystem, including

* Author for Correspondence W.M. Keck FT-ICR Mass Spectrometry Laboratory Department of Chemistry North Carolina State University Raleigh, North Carolina 27695 Phone: 919-513-0084 Fax: 919-513-7910 david_muddiman@ncsu.edu.

* Both authors equally contributed to this research.

Supporting Material: Tandem mass spectra for proteins with single peptide identification (MSIonsData.rar) and Tables S1-S3. This material is available free of charge via the Internet at <http://pubs.acs.org>.

application of proteomics will not only provide deeper insight in the biology of the pathogen, but may reveal new opportunities for disease control.

Asexual conidia are the primary source of inocula initiating pathogenesis. *M. oryzae* conidia typically contain three cells in a pyriform structure which adhere strongly to rice plant surfaces via a spore tip mucilage that is generated upon hydration of the conidium⁴. Following attachment, a germ tube emerges from the conidium and a dome shaped, melanized appressorium is formed⁵. Turgor pressure generated within the appressorium facilitates invasion of the leaf via a penetration peg from which invasive hyphae develop⁶. As a hemibiotrophic fungus, *M. oryzae* colonizes the outer cell layers of the leaf in its biotrophic phase and eventually proceeds to a necrotrophic phase characterized by lesion development. Aerial hyphae emerge from these lesions and produce conidiophores that release conidia to initiate additional infections⁷.

The genome of *M. oryzae* was recently sequenced by Dean *et al.*⁸ and is composed of 7 chromosomes, 41.7 million bases (Mb) and 11043 annotated genes. The latest genome annotation (MG8, Broad Institute) predicts the presence of 12991 proteins derived from 12827 protein coding genes. The availability of a high-quality genome sequence and annotation provides a reliable foundation for proteomic analyses.

Global proteomic analyses provide a framework towards an in depth understanding of an organism's cellular processes and organization but are subject to numerous challenges. The complexity of the proteome far exceeds that of the genome when protein isoforms, alternative splicing variants and post translational modifications are considered. Although genetic information encoded by the genome remains constant across different tissue types, the composition of the cellular proteome can vary amongst tissues, time and space. Analysis of the transcriptome via DNA microarray analysis or RNA sequencing captures a portion of this variation. However, measurement of mRNA levels does not capture regulation at the levels of protein translation and degradation and as a result mRNA levels show limited correlation to protein levels⁹. Furthermore, the complexity of protein and peptide chemistry complicates global proteome analyses, and no one technology is currently able to capture the complete diversity of a proteome.

Mass spectrometry (MS) is to date the most versatile and comprehensive tool for characterization of proteomes. Despite the fast duty cycles of modern hybrid mass spectrometers the complexity of proteomes requires fractionation methods to be applied prior to tandem MS analysis. One dimensional sodium dodecyl sulfate polyacrylamide (1D-SDS-PAGE) gels are a common way of performing protein level fractionation prior to proteomic analyses^{10,11}. However, for large scale analyses encompassing multiple biological conditions and replicates this method proves to be very time-consuming and labor-intensive. The filter aided sample preparation method (FASP)^{12,13} combined with stop and go extraction tip (StageTip) anion exchange fractionation was recently published¹⁴⁻¹⁷ as an alternative peptide level fractionation strategy. Utilizing this approach Mann and coworkers were able to identify 4206 proteins of the hippocampal membrane proteome¹⁴. The ability of this approach to facilitate the sample processing is impressive.

Previous proteomic analyses of *M. oryzae* were limited in scope. Kim *et al.* reported the identification of four proteins induced during appressorium formation¹⁸. Two additional studies targeted proteins secreted in response to nitrogen starvation¹⁹ (89 differentially expressed proteins) or appressorium formation on artificial surfaces²⁰ (53 secreted proteins identified). Bhadauria *et al.* found 31 proteins to be differentially regulated in the conidial proteome of $\Delta com1$ mutant when compared to the wild type²¹. Furthermore, in a previous study we reported identification of 1500 proteins from *M. oryzae* conidia using a GeLC

based approach¹¹. Herein we extend our analysis of *M. oryzae* conidial proteome and report the identification and analysis of 2912 proteins identified at 1% false discovery rate (FDR) from conidia of *M. oryzae* using FASP digestion and StageTip fractionation followed by nanoLC-MS/MS analysis.

2. Experimental Section

2.1 Sample preparation and digestion

All chemicals were purchased from Sigma Aldrich, St. Louis, MO, unless otherwise stated.

Conidia were harvested from 8 day old minimal medium agar as published in Gokce *et al.*¹¹ Briefly, mycelia were scraped from the surface of the medium and gently agitated in water to release conidia from the conidiophores prior to filtration through miracloth (EMD Biosciences, La Jolla, CA) to remove mycelium. Conidia were pelleted by centrifugation at $3,000 \times g$ for 10 min at 4°C and the resulting pellets were flash frozen in liquid nitrogen and stored at -80°C. Lysis buffer (100 µl per 2 million conidia) containing 2M urea, 1X phosphate buffered saline and 0.1% SDS was used to extract the proteins by bead beating with 100 mg of 0.5 mm Zirconia/Silica beads (BioSpec Products, Inc., Bartlesville, OK) followed by boiling for 5 min. Debris and beads were spun down by centrifugation at $10,000 \times g$ for 10 min at 4°C. Protein concentrations were determined by Bicinchoninic acid assay. One sample was prepared for preliminary analysis (gradient study) and two samples (same biological replicate) of 50 µg protein each were prepared for the reproducibility experiment. The samples were digested and fractionated on different days following the filter aided sample preparation (FASP) method protocol published by Mann and coworkers¹³ with minor changes. Samples were dried down to 27 µl and 3 µl of 50 mM Dithiothreitol (DTT) (Thermo Fischer Scientific, Rockford, IL) was added to a final DTT concentration of 5 mM. Incubation for 30 min at 56 °C followed to reduce the protein disulfide bonds. The samples were then mixed with 200 µl of 8 M urea in 0.1 M Tris/HCl (pH 8.5) on Vivacon 500 30 kDa MW cutoff filter units (Vivaconproducts, Littleton, MA) and centrifuged at constant 21 °C to prevent carbamylation at $14,000 \times g$ in a refrigerated bench-top centrifuge (Eppendorf 5810R) (Eppendorf, Hauppauge, NY) equipped with a fixed angle rotor (F45-30-11) (Eppendorf, Hauppauge, NY) for 15 min. This step was repeated one more time and the flow through solvent was discarded before 100 µl of 0.05 M iodoacetamide was added. The samples were incubated for 30 min in the dark at room temperature for alkylation of the free thiols and centrifuged afterwards for 10 min. Three 100 µl 8M urea buffer washes and three 100 µl 0.05 M ammonium bicarbonate in H₂O washes were performed. The flow-through collection vials were switched out with clean ones and 80 µl ammonium bicarbonate buffer and 2.5 µl of 0.2 µg/µl stock trypsin solution were added at a 1:100 enzyme:protein ratio. Digestion was performed for 16 h at 37 °C. The peptides were eluted with 40 µl ammonium bicarbonate buffer. Nanodrop 2000c (Thermo Scientific, Wilmington, DE) readings at 280 nm with 2 µl of each sample were taken and indicated ~98% peptide recovery.

2.2 StageTip Anion Fractionation

A Kel-F Hub, point style 3, gauge 16 (1.19 mm) needle with a 100 µl plunger assembly was used to extract 6 anion disks from an Empore anion extraction disk (3M, St. Paul, MN) and stacked into a 200 µl pipette tip (Eppendorf, Hauppauge, NY). The same needle and plunger were used to extract 18 C18-disks from an Empore C18 extraction disk (3M, St. Paul, MN) and 6 desalting StageTips were prepared by stacking 3 C18 disks into six 200 µl pipette tips. Seven StageTips (1 anion, 6 C18) were prepared for every sample¹⁷. A swing bucket rotor (A-2-DWP) (Eppendorf, Hauppauge, NY) was placed into the refrigerated centrifuge with 2 robust pipette tip boxes in order to perform the StageTip conditionings. The anion StageTips were conditioned by adding 100 µl methanol (Burdick and Jackson, Muskegon, MI) from

the top of the pipette tip, placing into the pipette tip box and centrifuging at $1,800 \times g$ for 2 min. Subsequently a 100 μ l 1M NaOH and two 100 μ l of Britton Robinson buffer pH 11 (Ricca Chemicals, Arlington, TX) washes were performed. C18 StageTips were conditioned by sequential washes with 50 μ l of methanol, 50 μ l 80% (v/v) acetonitrile (Burdick and Jackson, Muskegon, MI) containing 3% (v/v) trifluoroacetic acid and 50 μ l H₂O (Burdick and Jackson, Muskegon, MI).

The peptide samples were diluted to ~200 μ l by adding 100 μ l of Britton Robinson buffer pH 11 and then loaded on top of the anion StageTips. Peptides were eluted onto a C18 StageTip placed underneath the anion StageTip by centrifugation. Additional peptide elutions from the anion StageTip were performed with Britton Robinson buffers pH 8, 6, 5, 4, and 3 into individual C18 StageTips. Sodium chloride (0.25 M) was added to buffer pH 3 in order to ensure elution of all peptides from the anion StageTips. The C18 StageTips containing the fractionated samples were then washed with 50 μ l 0.1% (v/v) trifluoroacetic acid in H₂O and samples eluted with 100 μ l 60% (v/v) acetonitrile using a syringe. Fractions were dried to completion and stored at -80°C until analysis.

2.3 NanoLC-MS/MS

The nanoLC-MS/MS was performed on a cHIPLC-Nanoflex system (Eksigent, Dublin, CA) at 26°C . A trap column $200 \mu\text{m} \times 0.5 \text{ mm}$ packed with Chrom XP C18-CL $3 \mu\text{m}$ $120 \text{ \AA}$ was in-line with a $75 \mu\text{m} \times 15 \text{ cm}$ analytical column packed with the same packing material. The fractions were diluted in 20 μ l mobile phase A (98% H₂O, 2% acetonitrile, 0.1% formic acid), each run 4 μ l of the sample was injected into the sample loop and loaded onto the trap column. A 350 nl/min flow rate was used during the run. Sixty min 5-50% B gradients were run for the preliminary analysis. A 210 min 5-50% B gradient and a shallow 5-35% B gradient for 210 minutes were also explored. The reproducibility experiments were all run with the shallow 210 min gradient. Three technical replicates of each fraction were recorded.

MS analysis was performed on a hybrid LTQ-Orbitrap XL MS (Thermo Fischer Scientific, Bremen, Germany). Optimized instrument parameters for achieving maximum proteome coverage recently published by our group were used *verbatim*²² with one minor change. The MS/MS data was collected in centroid mode and not in profile mode, due to very large file sizes (above 2 GB) that could not be handled by Xcalibur version 2.0.7 SP1. A summary of the instrument parameters are also provided in the Supplementary Material Section (Table S1).

2.4 Data analysis

Raw files were analyzed in the RawMeat program version 2.0, downloaded from <http://vastscientific.com/rawmeat>. Raw files were also searched against a concatenated target-reverse *M. oryzae* database MG8 (*Magnaporthe* comparative Sequencing Project, Broad Institute of Harvard and MIT, <http://www.broadinstitute.org>) and converted into .dat files through MASCOT Distiller version 2.3.2 (Matrix Science Inc., Boston, MA). In MASCOT the distiller option Orbitrap_low_res_MS2_2.opt was used as peak picking algorithm and search parameters were the same as in Gokce *et al.*¹¹. ProteoIQ version 2.3.02 (BioInquire, Athens, GA) was used to determine the protein identifications (ID's) at 1% protein FDR²³. Tandem mass spectra for protein identifications based on single peptide assignments are provided in the Supplementary Material Section (MSIonsData.rar). The WoLF PSORT program²⁴ was utilized to assign subcellular localization to all predicted proteins in version 8 of the *M. oryzae* genome.

3. Results & Discussion

Gradient Study

The FASP-digestion and StageTip fractionation workflow is shown in Figure 1. The procedure is rapid and amenable to processing multiple samples simultaneously. *M. oryzae* conidial proteins extracted by bead beating were reduced with DTT and loaded onto molecular weight cutoff filters. Utilizing a 30 kDa molecular weight cutoff filter, 150 µg of total protein per sample was processed in one filter within 3 hours. In our experience, up to 30 samples can be processed simultaneously. SDS, which is advantageous during protein extraction but is incompatible with electrospray ionization mass spectrometry (ESI-MS), was thoroughly depleted from samples during processing. Highly concentrated Urea (8M) was utilized to weaken the hydrophobic SDS-protein interactions and as an aid in removing the SDS by interacting electrostatically with the peptides. After SDS depletion, alkylation with iodoacetamide and buffer exchange with ammonium bicarbonate, digestion was performed with trypsin. Anion exchange fractionation using a pH gradient (pH 11, 8, 6, 5, 4, 3) and C18 clean-up was performed using StageTips resulting in extremely clean samples. The fractions were subjected to triplicate nanoLC-MS analysis and no trap or column clogging was experienced using FASP in combination with the StageTip clean-up.

Running 60 min gradients (5-50% B) for the six fractions we identified 1511 proteins, comparable to our GeLC study on similar samples analyzing 10 gel-fractions as previously published¹¹. By extending our gradients to 210 min we increased protein ID's by 60% (60 min: 687 → 210 min: 1091) for a single fraction (fraction pH 8) (Figure 2A). Almost all identified proteins in the 60 min gradient run were within the identifications in the 210 min run. Furthermore, running a shallower gradient 5-35% B increased our number of protein ID's by an additional 5% (210 min: 1091 → 210 min shallow: 1145). Although the intensity of the total ion chromatograms (TIC's) dropped with both increased gradient times and shallower gradients, the chromatograms still appeared to be very complex attributable to the fact that more material can be loaded per run. Figure 2B shows the number of consecutive MS² scans, allowed was a maximum of 8 MS/MS per precursor ion scan. In the 60 min gradient run those 8 MS/MS events were hit most of the time (green colored bars), indicating under-sampling of the proteome. By performing longer separations (210 min gradients) we saw an overall increase in the total number of MS/MS events and the shallow gradient decreased the frequency of 8 consecutive MS/MS events (rose colored bars) suggesting a more efficient sampling of the proteome. Table 1 shows additional information about the three different gradient runs. An approximate three-fold increase was observed in the number of precursor spectra (60 min: 3387 → 210 min: 9427 → 210 min shallow: 9753) as well as fragment ion spectra (60 min: 5672 → 210 min: 14341 → 210 min shallow: 17185) by switching to the longer gradients. The number of identified unique peptides was doubled (60 min: 2017 → 210 min: 3977 → 210 min shallow: 4290). The peptide yield was calculated by dividing the number of identified peptides by the number of MS/MS spectra, and although the peptide yield decreased from 36% to 29% to 25%, the number of protein ID's still increased due to the identification of new unique peptides derived from previously undetected proteins.

Reproducibility Experiment

Applying 210 min LC gradients on all fractions in the reproducibility experiment resulted in the identification of 2912 proteins, corresponding to 22.4% of the predicted proteome of *M. oryzae*. The 2912 identified proteins are provided in Table S2 ranked by abundance derived from Normalized Spectral Abundance Factor (NSAF)²⁵ values with spectral count data included for each protein. The FASP and StageTip approach with 6 fractions run on 210 min gradients required 1 day of instrument time for each replicate per sample. It is possible that

more extensive processing and fractionation may provide additional protein identifications. For example, utilizing milligram quantities of protein, protein level fractionation by gel filtration, multi-enzyme FASP digestion and dedicating 12 days of instrument time to a single sample, Mann and coworkers were able to identify 10255 proteins from HeLa cells²⁶. Limitations based upon the availability and efficient use of instrument time as well as the ability to generate milligram quantities of protein sample preclude identification of this many proteins in most analyses.

In order to determine cutoffs for the detection of real biological change in future experiments via label free quantification for the FASP StageTip sample preparation and fractionation method, a false positive rate (FPR) table (Table 2) was created using data from replicate analyses of the a single conidial proteome sample where no difference in protein abundance is expected. Different fold-change and p-value cutoffs were applied across spectral count bin width (s stands for spectral count per replicate) and very similar FPR and protein distributions across the bin widths (N) were seen when compared to previous GeLC data¹¹. Low FPR's are colored in blue while high FPR's are red. The FPRs indicate that lower fold change-cutoffs can be applied on higher spectral count bin widths while maintaining FPRs consistent with low spectral count bin widths. This is especially beneficial for very abundant proteins where commonly used 2 fold-changes are rarely observed. Defining 10% FPR to be acceptable for each spectral count bin width, a p-value of 0.05 and the appropriate fold-change cutoff will be applied on the spectral bin widths in further large scale experiments where detection of differential protein expression by label free quantification is desired.

Subcellular Localization Analysis

An analysis of the subcellular localization of the conidial proteome is presented in Figure 3. Predictions of subcellular localization for the entire *M. oryzae* predicted proteome were performed using the WoLF PSORT program²⁴. A qualitative analysis of the identified proteins indicates that the distribution of proteins in the predicted proteome (Figure 3A) is highly similar to that of the conidial proteome (Figure 3B). A higher percentage of cytoplasmic proteins were observed in the conidial proteome relative to the whole proteome (14.8% of the whole proteome, 26.7% of the proteins identified), likely due to the high solubility of this group of proteins. Conversely, nuclear proteins were not as well represented in the conidial proteome (27.5% of the whole proteome, 21.3% of the proteins identified). Interestingly, 249 proteins predicted to be extracellular were identified from conidia without specifically targeting extracellular proteins during sample preparation. However, when compared to the whole proteome, extracellular proteins represent a smaller percentage of the identified proteins (8.6% compared to 15.8% for the entire proteome).

To obtain a more quantitative view of the distribution of proteins identified from conidia, NSAF values were calculated for each protein (Table S2) and a sum of NSAF values was generated for all proteins within each subcellular localization category (Figure 3C). The results of this analysis indicate that cytoplasmic and mitochondrial proteins represent not only the largest groups of identified proteins but also include proteins of high abundance. In contrast, the nuclear and extracellular protein categories that are underrepresented qualitatively also generally contain proteins of lower abundance or in the case of extracellular proteins were not specifically targeted in this study. Finally, although membrane localized proteins were identified at a rate comparable to that observed for the whole proteome, these proteins appear much lower in abundance. The FASP methodology was previously reported to identify membrane proteins in an unbiased manner¹⁴. However, it is unclear if the apparent low abundance of membrane proteins in conidia is due to inefficiencies in their isolation and detection using the current methodologies or simply reflects their true biological status.

Gene Ontology Analysis

Further classification of the conidial proteome was performed using gene ontology (GO) analysis. GO terms were assigned to the *M. oryzae* proteome using the BLAST2GO program²⁷ and tested for enrichment relative to the entire proteome. Fisher's Exact test was used within BLAST2GO at an FDR of 5% and reduced for most specific terms (Table S3). A total of 12 GO terms were underrepresented and are primarily related to DNA metabolism, transcription and extracellular proteins in agreement with subcellular localization analysis and further supporting the observation that nuclear and extracellular proteins are under sampled. A total of 35 GO terms were overrepresented and primarily reflect classes of highly abundant proteins. The number of proteins identified for selected GO Process categories is presented in Figure 4 with a calculated sum of NSAF values for all proteins associated with a specific GO term presented above. Comparison of the sum of NSAF values to the total number of identified proteins for a given GO term indicates that five GO terms disproportionately contain highly abundant proteins and include: translation, carbohydrate metabolism, generation of precursor metabolites and energy, response to stress and secondary metabolic process.

Biological Insights into Identified Proteins

Examination of the 10% of proteins (291 proteins) with the highest NSAF values provides clues regarding the metabolic processes required for conidial germination. The translation GO term (GO:0006412) contains primarily ribosomal proteins, translation factors and tRNA synthetases and 80 of these proteins fall within the highest 10% of NSAF values. A total of 11 proteins corresponding to components of the ATP synthase complex were identified in the conidial proteome and belong to the generation of precursor metabolites and energy GO term (GO:0006091). Eight of the ATP synthase complex proteins are amongst the top 10% of NSAF values and four are among the top 20 most abundant proteins as determined by NSAF values. Glycolysis (8 proteins), the pyruvate dehydrogenase complex (4 proteins), and the citrate cycle (11 proteins) account for the highly abundant proteins in the carbohydrate metabolism GO term (GO:0005975). The response to stress GO term (GO:0006950) contains 15 chaperone proteins that account for 42% of the sum of NSAF values for that GO term.

Proteins involved in synthesis of the polyketide melanin, ALB1 (MGG_07219), RSY1 (MGG_05059), BUF1 (MGG_02252) in addition to a BUF1 homolog, 4HNR, (MGG_07216) (Table S4), account for 33% of the sum of NSAF values for the Secondary metabolism GO Term (GO:0019748). Three of these proteins, RSY1, BUF1, and 4HNR are within the top 10% of most abundant proteins. Previous studies have established a melanin requirement for pigmentation of *M. oryzae* cultures and also formation of functional appressoria prior to plant infection²⁸⁻³¹. *M. oryzae* strains carrying mutations in the hydroxynaphthalene reductase *BUF1* (MGG_02252), the polyketide synthase *ALB1* (MGG_07219), and the scytalone dehydratase *RSY1* all display abnormal culture pigmentation and are nonpathogenic due to formation of defective appressoria. BUF1 and 4HNR were amongst the most abundant proteins in our data set. Interestingly 4HNR was shown previously to catalyze the reduction of tetrahydroxynaphthalene to scytalone, a function often associated with BUF1³². *M. oryzae* mutant strains lacking a functional 4HNR retained the ability to produce melanin and form functional appressoria presumably due to the ability BUF1 to catalyze the reduction of both tri- and tetrahydroxynaphthalene. The high abundance of 4HNR in conidia supports the hypothesis that, although not specifically required, it contributes to the normal pathway for melanin biosynthesis *M. oryzae*.

M. oryzae conidia utilize storage reserves including glycogen, mannitol and trehalose for germination, hyphal development and appressorium formation³³. Glycogen is synthesized

from UDP-Glucose in part via the activity of glycogen synthase. UDP-glucose is derived from Glucose-6P, an intermediate of glycolysis, via the activities of phosphoglucomutase and UTP-glucose-1-phosphate uridylyltransferase. We identified four highly abundant conidial proteins predicted to be involved in synthesis of glycogen including a putative glycogen synthase (MGG_07289), a putative phosphoglucomutase (MGG_04495), and a putative UTP-glucose-1-phosphate uridylyltransferase (MGG_01631) (Table S4). A fourth protein, 1,4- α -glucan-branching enzyme (MGG_03186), is a homolog of yeast GLC3³⁴ (63% amino acid identity), a protein involved in the generation of 1,6- α -glucosidic linkages within the glycogen molecule. The utilization of conidial glycogen reserves during germination and appressorium formation was documented previously³⁵. Glycogen is broken down to glucose-1P via the activity of glycogen phosphorylase³⁶ (MGG_01819), also detected at high levels in conidia. The importance of glycogen homeostasis in conidiation and conidial germination of *M. oryzae* is reflected in the abundance of conidial proteins involved in glycogen metabolism.

M. oryzae conidia also contain lipid droplets that are mobilized during germination and appressorium formation³⁷. One of the most abundant proteins in the dataset was MGG_11916, a homolog of the *Colletotrichum gloeosporioides* CAP20 protein³⁸ and the *Metarhizium anisopliae* MPL1 protein³⁹. The *CAP20* gene of *C. gloeosporioides* is expressed in conidia and during appressorium formation and is necessary for pathogenicity. The *MPL1* gene of *M. anisopliae*, a homolog of mammalian perilipins, was highly expressed in conidia and localized to lipid droplets. In addition, Δ *MPL1* mutant conidia contain fewer lipid droplets. MPL1 is also required for generation of turgor pressure in developing appressoria and is required for full virulence in *M. anisopliae*. The high conidial abundance of MGG_11916 and its homology to MPL1 (47% amino acid identity) suggests a role in maintenance and hydrolysis of lipid droplets of *M. oryzae* conidia.

Utilization of lipid reserves requires beta oxidation of fatty acids liberated via lipase activity. A prior investigation of the *M. oryzae* genome indicated the presence of 16 proteins involved in β -oxidation of fatty acids⁴⁰. A total of 15 of these proteins were identified in the conidial proteome (Table S4), including a peroxisomal multifunctional β -oxidation protein, MFP1, previously demonstrated to be required for appressorium-mediated plant infection. Furthermore, two putative Acyl-CoA dehydrogenases (MGG_08661 and MGG_05949) predicted to catalyze the first step in β -oxidation of fatty acids were also identified at high abundance.

β -oxidation of fatty acids results in the production of acetyl-CoA which is further metabolized via the glyoxylate cycle. Isocitrate lyase is required for efficient germination of conidia as well as pathogenicity in *M. oryzae*⁴¹. Both enzymes unique to the glyoxylate cycle, isocitrate lyase (MGG_04895) and a putative malate synthase (MGG_02813) were identified (Table S4) and their relative abundances are similar to other enzymes of the citrate cycle indicating that *M. oryzae* conidia contain a functional glyoxylate bypass and are capable of utilizing Acetyl-CoA derived from fatty acid β -oxidation.

Several proteins involved in the detoxification of reactive oxygen species (ROS) were also detected at high levels in conidia (Table S4). The most abundant of these is the peroxiredoxin, HYR1 (MGG_07460). HYR1 is a homolog of the yeast glutathione-dependent peroxidase HYR1p and *M. oryzae* Δ *HYR1* mutants display growth impairments in the presence of H₂O₂⁴². HYR1 was demonstrated to be localized to vacuoles during the earliest stages of conidial germination on the surfaces of barley leaves and is required for efficient response to plant-derived ROS during infection. Additional abundant ROS detoxification proteins identified include a putative superoxide dismutase (MGG_02625), a putative monofunctional catalase CATA (MGG_10061), and the bifunctional catalase/

peroxidases CPXA/KATG1⁴³ (MGG_04337) and CPXB^{44,45} (MGG_09834). Catalase activity was previously reported in supernatants of both conidial suspensions and hyphal cultures. CPXB was described as an extracellular catalase/peroxidase and Δ CPXB mutants have reduced extracellular catalase activity and increased sensitivity to H₂O₂ in culture. The detection of several highly abundant proteins involved in ROS detoxification suggests that *M. oryzae* conidia are subject to oxidative stress and maintain high levels of proteins capable of mitigating the negative effects of ROS exposure.

Overall, the analysis of identified proteins suggests that conidia are primed for germination and rich in proteins required for basic housekeeping functions. Proteins required for utilization of storage reserves during germination are also highly abundant in conidia. Furthermore, proteins with known roles in the biology of the conidium were often identified in high abundance.

4. Conclusions

We successfully identified 2912 proteins of the *M. oryzae* proteome by applying FASP and StageTip fractionation combined with long and shallow nanoLC gradients to conidial protein samples. In comparison to GeLC we achieved a doubling in the number of protein identifications with this new methodology and observed marginally lower FPR values for the reproducibility experiment. Subcellular localization analysis of the proteome revealed in general a comparable distribution of our data to the predicted proteome. Gene ontology analysis indicates that housekeeping proteins and components of central metabolic processes are well represented in the highly abundant proteins of the *M. oryzae* conidial proteome. In addition, analysis of the conidial proteome indicates that proteins important to the physiology of conidia such as those involved in glycogen metabolism and β -oxidation of fatty acids are present and often highly abundant. The depth of proteome coverage for *M. oryzae* conidia presented here provides an excellent foundation for future investigations of pathogenic development of *M. oryzae* including germination of conidia and development of appressoria.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Mr. Ian Jones for thoughtful comments, the National Science Foundation (Grant Number MCB-0918611), the W. M. Keck Foundation, and North Carolina State University for financial support. W.L.F. is supported by the NIH Molecular Mycology and Pathogenesis Training Program.

References

- (1). Ou SH. Pathogen Variability and Host-Resistance in Rice Blast Disease. Annu. Rev. Phytopathol. 1980; 18:167–187.
- (2). Greer CA, Scardaci SC, Webster RK. First Report of Rice Blast Caused by *Pyricularia grisea* in California. Plant Dis. 1997; 81:1094–1094.
- (3). Scardaci SCW,RK, Greer CA, Hill JE, Williams JF, Mutters RG, Brandon DM, McKenzie KS, Oster JJ. Rice Blast: A New Disease in California. Agronomy Fact Sheet Series. 1997; 2:1–2.
- (4). Howard RJ, Valent B. Breaking and Entering: Host Penetration by the Fungal Rice Blast Pathogen *Magnaporthe grisea*. Annu Rev Microbiol. 1996; 50:491–512. [PubMed: 8905089]
- (5). Dean RA. Signal Pathways and Appressorium Morphogenesis. Annu Rev Phytopathol. 1997; 35:211–234. [PubMed: 15012522]

- (6). deJong JC, McCormack BJ, Smirnov N, Talbot NJ. Glycerol Generates Turgor in Rice Blast. *Nature*. 1997; 389:244–245.
- (7). Talbot NJ. On the Trail of a Cereal Killer: Exploring the Biology of *Magnaporthe grisea*. *Annual Review of Microbiology*. 2003; 57:177–202.
- (8). Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan HQ, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeier C, Li WX, Harding M, Kim S, Lebrun MH, Bohnert H, Coughlan S, Butler J, Calvo S, Ma LJ, Nicol R, Purcell S, Nusbaum C, Galagan JE, Birren BW. The Genome Sequence of the Rice Blast Fungus *Magnaporthe grisea*. *Nature*. 2005; 434:980–986. [PubMed: 15846337]
- (9). Schwanhauser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Global Quantification of Mammalian Gene Expression Control. *Nature*. 2011; 473:337–342. [PubMed: 21593866]
- (10). Shevchenko A, Wilm M, Vorm O, Mann M. Mass Spectrometric Sequencing of Proteins from Silver Stained Polyacrylamide Gels. *Anal. Chem*. 1996; 68:850–858. [PubMed: 8779443]
- (11). Gokce E, Shuford CM, Franck WL, Dean RA, Muddiman DC. Evaluation of Normalization Methods on GeLC-MS/MS Label-Free Spectral Counting Data to Correct for Variation during Proteomic Workflows. *J Am Soc Mass Spectr*. 2011; 22:2199–2208.
- (12). Manza LL, Stamer SL, Ham AJ, Codreanu SG, Liebler DC. Sample Preparation and Digestion for Proteomic Analyses using Spin Filters. *Proteomics*. 2005; 5:1742–1745. [PubMed: 15761957]
- (13). Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal Sample Preparation Method for Proteome Analysis. *Nat. Methods*. 2009; 6:359–363. [PubMed: 19377485]
- (14). Wisniewski JR, Zougman A, Mann M. Combination of FASP and StageTip-Based Fractionation Allows In-Depth Analysis of the Hippocampal Membrane Proteome. *J. Proteome Res*. 2009; 8:5674–5678. [PubMed: 19848406]
- (15). Ishihama Y, Rappsilber J, Mann M. Modular Stop and Go Extraction Tips with Stacked Disks for Parallel and Multidimensional Peptide Fractionation in Proteomics. *J. Proteome Res*. 2006; 5:988–994. [PubMed: 16602707]
- (16). Rappsilber J, Ishihama Y, Mann M. Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics. *Anal. Chem*. 2003; 75:663–670. [PubMed: 12585499]
- (17). Rappsilber J, Mann M, Ishihama Y. Protocol for Micro-Purification, Enrichment, Pre-Fractionation and Storage of Peptides for Proteomics Using StageTips. *Nat. Protoc*. 2007; 2:1896–1906. [PubMed: 17703201]
- (18). Kim ST, Yu S, Kim SG, Kim HJ, Kang SY, Hwang DH, Jang YS, Kang KY. Proteome Analysis of Rice Blast Fungus (*Magnaporthe grisea*) Proteome During Appressorium Formation. *Proteomics*. 2004; 4:3579–3587. [PubMed: 15378734]
- (19). Wang Y, Wu J, Park ZY, Kim SG, Rakwal R, Agrawal GK, Kim ST, Kang KY. Comparative Secretome Investigation of *Magnaporthe oryzae* Proteins Responsive to Nitrogen Starvation. *J Proteome Res*. 2011; 10:3136–3148. [PubMed: 21563842]
- (20). Jung YH, Jeong SH, Kim SH, Singh R, Lee JE, Cho YS, Agrawal GK, Rakwal R, Jwa NS. Secretome Analysis of *Magnaporthe oryzae* Using In Vitro Systems. *Proteomics*. 2012; 12:878–900. [PubMed: 22539438]
- (21). Bhaduria V, Wang LX, Peng YL. Proteomic Changes Associated with Deletion of the *Magnaporthe oryzae* Conidial Morphology-Regulating Gene COM1. *Biol Direct*. 2010; 5:1–19. [PubMed: 20056006]
- (22). Andrews GL, Dean RA, Hawkridge AM, Muddiman DC. Improving Proteome Coverage on a LTQ-Orbitrap Using Design of Experiments. *J. Am. Soc. Mass Spec*. 2011; 22:773–783.
- (23). Weatherly DB, Atwood JA, Minning TA, Cavola C, Tarleton RL, Orlando R. A Heuristic Method for Assigning a False-Discovery Rate for Protein Identifications from Mascot Database Search Results. *Mol. Cell. Proteomics*. 2005; 4:762–772. [PubMed: 15703444]
- (24). Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: Protein Localization Predictor. *Nucleic Acids Res*. 2007; 35:1–3. [PubMed: 16920744]

- (25). Zybaylov B, Mosley AL, Sardu ME, Coleman MK, Florens L, Washburn MP. Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces cerevisiae*. J. Proteome Res. 2006; 5:2339–2347. [PubMed: 16944946]
- (26). Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Paabo S, Mann M. Deep Proteome and Transcriptome Mapping of a Human Cancer Cell Line. Mol Syst Biol. 2011; 7:1–8.
- (27). Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: A Universal Tool for Annotation, Visualization and Analysis in Functional Genomics Research. Bioinformatics. 2005; 21:3674–3676. [PubMed: 16081474]
- (28). Chumley FG, Valent B. Genetic Analysis of Melanin-Deficient, Nonpathogenic Mutants of *Magnaporthe grisea*. Mol Plant Microbe Interact. 1990; 3:135–143.
- (29). Howard RJ, Ferrari MA. Role of Melanin in Appressorium Function. Experimental Mycology. 1989; 13:403–418.
- (30). Howard RJ, Ferrari MA, Roach DH, Money NP. Penetration of Hard Substrates by a Fungus Employing Enormous Turgor Pressures. Proc Natl Acad Sci U S A. 1991; 88:11281–11284. [PubMed: 1837147]
- (31). Vidalcros A, Viviani F, Labesse G, Boccara M, Gaudry M. Polyhydroxynaphthalene Reductase Involved in Melanin Biosynthesis in *Magnaporthe-grisea* - Purification, cDNA Cloning and Sequencing. European Journal of Biochemistry. 1994; 219:985–992. [PubMed: 8112349]
- (32). Thompson JE, Fahnestock S, Farrall L, Liao DI, Valent B, Jordan DB. The Second Naphthol Reductase of Fungal Melanin Biosynthesis in *Magnaporthe grisea*: Tetrahydroxynaphthalene Reductase. J Biol Chem. 2000; 275:34867–34872. [PubMed: 10956664]
- (33). Dixon KP, Xu JR, Smirnoff N, Talbot NJ. Independent Signaling Pathways Regulate Cellular Turgor During Hyperosmotic Stress and Appressorium-Mediated Plant Infection by *Magnaporthe grisea*. Plant Cell. 1999; 11:2045–2058. [PubMed: 10521531]
- (34). Thon VJ, Vigneron-Lesens C, Marianne-Pepin T, Montreuil J, Decq A, Rachez C, Ball SG, Cannon JF. Coordinate Regulation of Glycogen Metabolism in the Yeast *Saccharomyces cerevisiae*. Induction of Glycogen Branching Enzyme. J Biol Chem. 1992; 267:15224–15228. [PubMed: 1634552]
- (35). Thines E, Weber RW, Talbot NJ. MAP Kinase and Protein Kinase A-Dependent Mobilization of Triacylglycerol and Glycogen During Appressorium Turgor Generation by *Magnaporthe grisea*. Plant Cell. 2000; 12:1703–1718. [PubMed: 11006342]
- (36). Deng YZ, Ramos-Pamplona M, Naqvi NI. Autophagy-Assisted Glycogen Catabolism Regulates Asexual Differentiation in *Magnaporthe oryzae*. Autophagy. 2009; 5:33–43. [PubMed: 19115483]
- (37). Weber RW, Wakley GE, Thines E, Talbot NJ. The Vacuole as Central Element of the Lytic System and Sink for Lipid Droplets in Maturing Appressoria of *Magnaporthe grisea*. Protoplasma. 2001; 216:101–112. [PubMed: 11732192]
- (38). Hwang CS, Flaishman MA, Kolattukudy PE. Cloning of a Gene Expressed During Appressorium Formation by *Colletotrichum Gloeosporioides* and a Marked Decrease in Virulence by Disruption of this Gene. Plant Cell. 1995; 7:183–193. [PubMed: 7756829]
- (39). Wang C, St Leger RJ. The Metarhizium Anisopliae Perilipin Homolog MPL1 Regulates Lipid Metabolism, Appressorial Turgor Pressure, and Virulence. J Biol Chem. 2007; 282:21110–21115. [PubMed: 17526497]
- (40). Wang ZY, Soanes DM, Kershaw MJ, Talbot NJ. Functional Analysis of Lipid Metabolism in *Magnaporthe grisea* Reveals a Requirement for Peroxisomal Fatty Acid Beta-Oxidation During Appressorium-Mediated Plant Infection. Mol Plant Microbe In. 2007; 20:475–491.
- (41). Wang ZY, Thornton CR, Kershaw MJ, Li DB, Talbot NJ. The Glyoxylate Cycle is Required for Temporal Regulation of Virulence by the Plant Pathogenic Fungus *Magnaporthe grisea*. Mol Microbiol. 2003; 47:1601–1612. [PubMed: 12622815]
- (42). Huang K, Czynn timer KJ, Caplan JL, Sweigard JA, Donofrio NM. HYR1-Mediated Detoxification of Reactive Oxygen Species is Required for Full Virulence in the Rice Blast Fungus. PLoS Pathog. 2011; 7:e1001335, 1–18. [PubMed: 21533213]

- (43). Zamocky M, Furtmuller PG, Bellei M, Battistuzzi G, Stadlmann J, Vlasits J, Obinger C. Intracellular Catalase/Peroxidase from the Phytopathogenic Rice Blast Fungus *Magnaporthe grisea*: Expression Analysis and Biochemical Characterization of the Recombinant Protein. *Biochem J*. 2009; 418:443–451. [PubMed: 19000033]
- (44). Tanabe S, Hayashi N, Nishizawa Y, Yamane H, Shibuya N, Minami E. Elicitor and Catalase Activity of Conidia Suspensions of Various Strains of *Magnaporthe grisea* in Suspension-Cultured Cells of Rice. *Biosci Biotechnol Biochem*. 2008; 72:889–892. [PubMed: 18323633]
- (45). Tanabe S, Ishii-Minami N, Saitoh K, Otake Y, Kaku H, Shibuya N, Nishizawa Y, Minami E. The Role of Catalase-Peroxidase Secreted by *Magnaporthe oryzae* During Early Infection of Rice Cells. *Mol Plant Microbe Interact*. 2011; 24:163–171. [PubMed: 21043575]

Figure 1.

Schematic of FASP and StageTip workflow. After protein extraction and reduction the sample is loaded onto the 30 kDa molecular weight cutoff filter. Urea was used to help deplete the SDS from the sample. Reduction, alkylation, buffer exchange to ammonium bicarbonate and trypsin digestion were performed on the filter. StageTip anion exchange fractionation and subsequent C18 clean-up was performed at the peptide level.

Figure 2.

Total ion chromatograms of a 60 min gradient (5-50% B) in comparison to 210 min (5-50% B), and 210 min with a shallow gradient (5-35% B) for the pH 8 fraction are shown (A). The frequency of consecutive MS/MS events for each gradient performed is diagrammed in B. While the maximum of 8 consecutive MS/MS events was hit most of the time in the 60 min gradient, a more uniform distribution between consecutive scans can be seen for the 210 min gradient. A drop in 8 MS/MS events was only seen for the 210 min shallow gradient, suggesting an optimized window for peptide elution and detection.

Figure 3.

Subcellular localization charts for the predicted whole proteome (A), sum of the number of proteins identified in each category (B), and the sum of NSAF values for proteins identified in each category (C). NSAFs indicate that while plasma membrane proteins are represented in our study (9.2% of total protein identifications) they are of lower abundance (3.6%). The same is true for extracellular proteins. More highly abundant proteins were mitochondrial and cytoplasmic proteins.

Figure 4.

Gene ontology analysis of biological process terms of observed proteins and their relative abundance levels determined by summing NSAF values of the proteins in each category. Beige bars highlight categories with disproportionately highly abundant proteins.

Table 1

Number of full scans, MS/MS scans, identified peptides, calculated peptide yield and identified proteins for the different gradients used for the single fraction at pH8.

	Full Scans	MS/MS Scans	Peptides	Peptide Yield	Proteins
60 min gradient	3387	5672	2017	36%	687
210 min gradient	9427	14341	3977	28%	1091
210 min shallow gradient	9753	17185	4290	25%	1145

Table 2

False Positive Rate table (s: spectral counts per replicate, N: number of proteins)

p&value	fold&change	ALL	1.67	s	1.66	1.66<s	2	2<s	5	5<s	10	10<s	15	15<s	20	s>20
0.1	1.25	17.69	17.28	19.4	19.1	19.1	19.6	18.6	21.6	12.0	11.2					
0.1	1.50	11.61	9.76	19.4	19.1	15.7	8.2	10.0	0.0	0.0	0.8					
0.1	2.00	7.04	4.37	18.3	16.7	7.3	1.5	1.2	0.0	0.0	0.0					
0.1	2.50	5.19	2.38	17.1	9.8	3.9	0.8	0.4	0.0	0.0	0.0					
0.1	3.00	3.91	1.06	16.0	4.2	1.7	0.6	0.0	0.0	0.0	0.0					
0.05	1.25	11.16	10.44	14.2	11.6	11.1	12.0	14.5	6.3	6.7						
0.05	1.50	7.35	5.73	14.2	11.6	8.6	5.5	6.2	0.0	0.4						
0.05	2.00	4.81	2.84	13.1	10.2	4.8	1.1	1.2	0.0	0.0						
0.05	2.50	3.81	1.91	11.9	8.8	2.7	0.8	0.4	0.0	0.0						
0.05	3.00	2.88	0.89	11.3	3.3	1.4	0.6	0.0	0.0	0.0						
0.01	1.25	4.57	2.97	11.3	2.8	3.4	3.2	4.6	1.9	1.8						
0.01	1.50	3.33	1.44	11.3	2.8	1.7	1.7	2.9	0.0	0.0						
0.01	2.00	2.54	0.72	10.3	1.9	0.9	0.8	0.8	0.0	0.0						
0.01	2.50	2.23	0.64	9.0	1.9	0.9	0.6	0.4	0.0	0.0						
0.01	3.00	1.89	0.34	8.5	0.5	0.5	0.6	0.0	0.0	0.0						
N		2912.00	2356.00	556.00	215.00	776.00	474.00	241.00	158.00	492.00						