

Research paper

Novel proteins from proteomic analysis of the trunk disease fungus *Lasiodiplodia theobromae* (Botryosphaeriaceae)

Carla C. Uranga^{a,*}, Majid Ghassemian^{b,1}, Rufina Hernández-Martínez^a

^a Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Carretera Ensenada-Tijuana 3918, Zona Playitas, 22860 Ensenada, B.C., Mexico

^b University of California, San Diego, Department of Chemistry and Biochemistry, 9500 Gilman Drive, La Jolla, CA 92093-0378, USA

Received 3 February 2017; accepted 2 March 2017

Available online 14 March 2017

Abstract

Many basic science questions remain regarding protein functions in the pathogen: host interaction, especially in the trunk disease fungi family, the Botryosphaeriaceae, which are a global problem for economically important plants, especially fruiting trees. Proteomics is a highly useful technology for studying protein expression and for discovering novel proteins in unsequenced and poorly annotated organisms. Current fungal proteomics approaches involve 2D SDS-PAGE and extensive, complex, protein extraction methodologies. In this work, a modified Folch extraction was applied to protein extraction to perform both *de novo* peptide sequencing and peptide fragmentation analysis/protein identification of the plant and human fungal pathogen *Lasiodiplodia theobromae*. Both bioinformatics approaches yielded novel peptide sequences from proteins produced by *L. theobromae* in the presence of exogenous triglycerides and glucose. These proteins and the functions they may possess could be targeted for further functional characterization and validation efforts, due to their potential uses in biotechnology and as new paradigms for understanding fungal biochemistry, such as the finding of allergenic enolases, as well as various novel proteases, including zinc metalloproteinases homologous to those found in snake venom. This work contributes to genomic annotation efforts, which, hand in hand with genomic sequencing, will help improve fungal bioinformatics databases for future studies of Botryosphaeriaceae. All data, including raw data, are available via the ProteomeXchange data repository with identifier PXD005283. This is the first study of its kind in Botryosphaeriaceae.

© 2017 The Authors. Published by Elsevier B.V. on behalf of Société Française de Biochimie et Biologie Moléculaire (SFBBM). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Trunk-disease fungi; Peptide fragmentation analysis; *de novo* peptide sequencing; Bioinformatics; Gene ontology

1. Introduction

Current work in fungal proteomics and bioinformatics of phytopathogenic filamentous fungi is in a very early stage. There are an estimated 1.5–5 million species of fungi thought to exist on earth [1], and most of them remain unsequenced and uncharacterized. Apart from the importance of efforts to understand fungal pathogenicity and pathogen: host-specific interactions, fungi are excellent model organisms for gaining

insight into the evolution of biochemical routes and their conservation across kingdoms and taxa [2]. Proteome analysis is a valuable technique for studying protein expression of organisms in different biological and experimental contexts [3], as well as for finding and annotating novel or unknown genome sequences [4]. Proteins are the final product of gene expression, and essentially the catalytic and metabolic force of an organism. Because of this and the many possible post-translational modifications of proteins [5,6], proteomics is an essential aspect of systematic gene expression studies for novel gene discovery.

Peptide sequencing and protein identification allows for the identification of biomarkers in pathological processes that may serve as indicators for disease and as targets for new

* Corresponding author.

E-mail addresses: curanga@cicese.edu.mx (C.C. Uranga), mghassemian@ucsd.edu (M. Ghassemian), ruherman@cicese.mx (R. Hernández-Martínez).

¹ Tel.: +1 534 822 0032.

treatments [7]. In the absence of genomic sequences for the organism of interest, proteomics is useful for designing biochemical characterization experiments via protein homology searches, which provide clues to enzymatic processes in uncharacterized proteins that may then be verified experimentally [3]. Protein characterization efforts are critical for improving bioinformatics databases available to the scientific community. Many programs exist to analyze peptides from a sample [8] as well as for reverse-translating protein identifications to then search genome sequences for homology, a valid approach for gene annotation and biochemical characterization efforts [9].

One example of an important family of mostly unsequenced fungi is the Botryosphaeriaceae. These are fungi that have been found to affect economically important woody plants around the world [10]. Several members of this family belong to “trunk disease fungi”, because they are able to invade and kill fruiting trees [11,12]. Emblematic symptoms from Botryosphaeriaceae consist of necrotic cankers in the trunks of the infected trees, reduced stature, and fruit rot [13]. A member of this family, *Lasiodiplodia theobromae* (teleomorph *Botryosphaeria rhodina*) has been found to be the most virulent species among those reported in grapevine [14]. *L. theobromae* is able to colonize a broad range of plant species, including important plants such as the rubber tree, (*Hevea brasiliensis*) [15], and the biofuel-producing plant *Jatropha curcas* [16]. Intriguingly, *L. theobromae* is also able to invade and colonize humans, and has been reported to cause corneal ulcers, keratitis, onychomycosis, pneumonia in a transplant patient [17], and skin lesions [18]. Only partial sequences are available in the NCBI database, mostly consisting of ITS regions used for species identification purposes.

Proteome studies from Botryosphaeriaceae (the proteome from *Diplodia seriata* is the only Botryosphaeriaceae studied to date reported in the literature) have involved the use of 1D and 2D SDS-PAGE [19] followed by mass spectrometry; laborious and technically difficult techniques that yield limited protein identification data [20]. Additionally, fungal proteins are notoriously difficult to extract, because fungal cell walls contain chitin, and, in the case of *L. theobromae*, extensive pigmentation (melanin) that interferes with the protein extraction [21]. Currently, protein extraction methods in fungi involve the use of painstaking, multi-step protein precipitation protocols requiring controlled precipitating reagents like trichloroacetic acid (TCA), or detergent extractions to precipitate proteins from aqueous suspensions, that result in protein and information loss [22]. In this work, lyophilized material was extracted directly using a modified Folch extraction, a method conventionally used to extract lipids from biological material [23]. Previous metabolomic studies of this fungus showed that when cultivated in exogenous triglycerides, a variety of fatty acid esters were produced [24], however, little is known about the proteins expressed by this fungus in this context. *L. theobromae* is a wound pathogen and is a problem in many vineyards especially because of grafting practices [25–27]. Lipids being one of the exposed substrates, the objective of this work was to evaluate the proteome of *L. theobromae* in the

presence of exogenous triglycerides, employing a multi-algorithm approach using database-dependent *in silico* fragmentation analysis in which a variety of databases were used. *De novo* peptide sequencing was also applied for comparison and for potential novel protein discovery.

2. Materials and methods

2.1. Fungal strains and incubation conditions

L. theobromae UCD256Ma (isolated in Madera County, California, USA) was provided by Dr. Douglas Gubler from the University of California at Davis [13]. The fungus isolate was incubated in triplicate in 50 mL Vogel's salts supplemented with 5% glucose and 5% grapeseed oil. All biological replicates were incubated for 20 days at 25 °C in the dark and then lyophilized. A modified Folch extraction consisted in the addition of 75 mL dichloromethane (DCM), 75 mL methanol and 0.01% of the antioxidant butylated hydroxytoluene (BHT) to the lyophilized material in each replicate and allowed to extract overnight at 4 °C. The solvents were removed from the solid material. For mass spectrometry-based peptide fragmentation analysis, a 0.5 g portion of the solid material from each replicate was used to create a pool.

2.2. Mass spectrometry methodology

The pooled samples were submitted to the University of California, San Diego proteomics mass spectrometry department to be processed according to standard procedure. Briefly, 0.5 g of the solids from the 50 mL fungal incubations (*L. theobromae* incubated in 5% glucose and 5% grapeseed oil and Vogel's salts for 20 days) remaining from the Folch extraction were dried under a stream of nitrogen and re-suspended in 50 mM Tris buffer, pH 8.00. Acetonitrile was added to the sample to a final concentration of 10%. The samples were then boiled for 5 min and cooled to room temperature. TCEP (Tris (2-carboxyethyl) phosphine) was added to 1 mM (final concentration) and the samples were incubated at 37 °C for 30 min. Subsequently, the samples were carboxymethylated with 0.5 mg/ml of iodoacetamide for 30 min at 37 °C in dark followed by neutralization with 2 mM TCEP (final concentration). Samples were boiled for 10 min followed by protease digestion with a 1:100 ratio of trypsin: protein (Pierce™ Trypsin Protease, MS Grade Catalog number: 90057 with K, R specificity). After an overnight digestion, samples were centrifuged on a desktop microfuge at max speed (15,000 rpm) for 10 min to remove the insoluble fraction. The soluble fraction was adjusted to 0.2% formic acid and 5% acetonitrile and its peptide content isolated using C-18 solid phase extraction (Thermo Scientific, PI-87782) as described by the manufacturer.

The nano-spray ionization experiments were performed using a TripleTOF 5600 hybrid mass spectrometer (ABSCIEX) interfaced with a nano-scale reversed-phase UPLC (Waters nano ACQUITY) using a 20 cm-75 μM ID glass capillary packed with 2.5-μM C18 (130) CSH™ beads

(Waters). Peptides were eluted from the C18 column into the mass spectrometer with a linear gradient (5–80%) of acetonitrile (ACN) at a flow rate of 250 $\mu\text{L}/\text{min}$ for 90 min. The buffers used to create the ACN gradient were Buffer A (98% H_2O , 2% ACN, 0.1% formic acid and 0.005% TFA) and Buffer B (100% ACN, 0.1% formic acid, and 0.005% TFA). MS/MS data were obtained in a data-dependent manner in which the MS1 data was acquired for 250 ms at m/z of 400–1250 Da and the MS/MS data was acquired from m/z of 50 to 2000 Da. An MS1-TOF acquisition time of 250 ms was set, followed by 50 MS2 events of 48 ms acquisition time for each event. The threshold to trigger the MS2 event was set to 150 counts, when the ion had the charge state +2, +3 and +4. The ion exclusion time was set to 4 s.

2.3. Protein identification

Peak lists obtained from MS/MS spectra were identified via fragmentation analysis (database dependent identification) using X! Tandem Vengeance (2015.12.15.2) [28], MS-GF+ version Beta (v10282) [29] and either OMSSA version 2.1.9 [30] or, in the case of the all-Uniprot database search only, Comet version 2016.01 rev. 2 [31]. The search was conducted using SearchGUI version 3.1.2 [32]. The data was searched against a whole Uniprot/Swissprot database search (manually annotated and reviewed) [33], as well as a non-redundant Botryosphaeriaceae-only database downloaded from NCBI [34]. An all-human database from Uniprot was also used for further assessing protein identifications. Because of the large amount of data collected, all identification data from each database may be found as a Data in Brief article as [Supplementary data S2, S3 and S4](#) [35].

The identification settings were as follows: Trypsin with a maximum of 2 missed cleavages; 60.0 ppm as MS1 and 0.8 Da as MS2 tolerances; fixed modifications: Carbamidomethylation of C (+57.021464 Da) and Oxidation of M (+15.994915 Da), variable modifications: Acetylation of protein N-term (+42.010565 Da), Pyrrolidone from E (+18.010565 Da), Pyrrolidone from Q (+17.026549 Da) and Pyrrolidone from carbamidomethylated C (+17.026549 Da). All algorithm-specific settings are listed in the Certificate of Analysis available in [Supplementary data S1](#) in the associated data in brief article [35].

Peptides and proteins were inferred from the spectrum identification results using PeptideShaker version 1.13.6 [36]. Peptide Spectrum Matches (PSMs), peptides and proteins were validated at a 1.0% False Discovery Rate (FDR) estimated using a decoy-hit distribution. Because of the large quantity of data, a Data in Brief article is cited when referring to the data [35]. All validation thresholds are listed in the Certificate of Analysis available in [Supplementary data S1A, S1B, and S1C](#) for all databases searched in the Data in Brief article [35]. Post-translational modification localizations were scored using the D-score [37] and the A-score [38] with a threshold of 95.0 as implemented in the compomics-utilities package [39].

The mass spectrometry raw data files along with the identification results have been deposited to the ProteomeXchange

Consortium [40] via the PRIDE partner repository [41] with the data set identifier PXD005283.

Gene ontology (GO) analysis of enriched proteins was done on all those hits obtained from the Uniprot database [33]. The software Cytoscape [42] with the BiNGO plugin [43] was used for GO and enrichment analysis using up-to-date databases, applying a hypergeometric test with a significance level (p -value) < 0.05 , as well as a Benjamini and Hochberg false discovery rate (FDR) correction. Interactive Cytoscape BiNGO networks were created with data from the all-Uniprot database search, and annotated with an all-Uniprot ontology database, with an interactive Cytoscape network available in [Fig. 1A](#) in the associated Data in Brief article [35]. Venn diagrams were created from the output of the hypergeometric test performed for enriched ontology categories with the Cytoscape BiNGO Plugin, using the “R”-based program VennDiagram [44]. An interactive cytoscape network is available in [Fig. 1B](#) in the associated Data in Brief article [35], as well as gene ontology annotations as [Supplementary data S6](#) in the same.

2.4. De novo peptide sequencing

De novo peptide sequencing was performed in order to compare results and explore peptides via sequence homology with sequenced proteins found in the entire Uniprot database using BLASTp. The program DeNovoGUI version 1.14.5 was used for this purpose [45], and both Novor [46] and PepNovo [47] were used for peptide sequencing. The mass allowance parameters were, for precursor mass tolerance: 10 ppm, and a fragment mass tolerance of 0.5 Da. Post-translational modification settings consisted in carbamidomethylation of cysteine (fixed) and oxidation of methionine (variable). All peptides were searched against the entire Uniprot database using a standalone version of NCBI-BLASTp [48], with one peptide match per spectrum (most significant) and one BLASTp match per peptide (most significant, lowest E -value). The BLASTp match data was also analyzed for gene ontology (molecular functions) as described above, and found in [Fig. 1B](#) and as [Supplementary material S6](#) in the Data in Brief article [35].

3. Results and discussion

This is the first LC-nanoESI-MS peptidome fragmentation analysis and *de novo* peptide sequencing of *L. theobromae* or any Botryosphaeriaceae. The Folch extraction was utilized because the literature reports protein loss when using other methods, which rely on precipitating proteins from aqueous solutions [20]. In this work, it is argued that using a combination of non-polar and semi-polar solvents (a 1:1 ratio of dichloromethane and methanol + 0.01% of the antioxidant BHT) in the Folch extraction of lyophilized (freeze-dried) material, removes interfering lipids and compounds without having to solubilize the proteins in aqueous solution, thus minimizing protein oxidation and loss. This is the first report of this in the literature, and has not been previously applied to proteomics of fungi. Although an unorthodox method, the

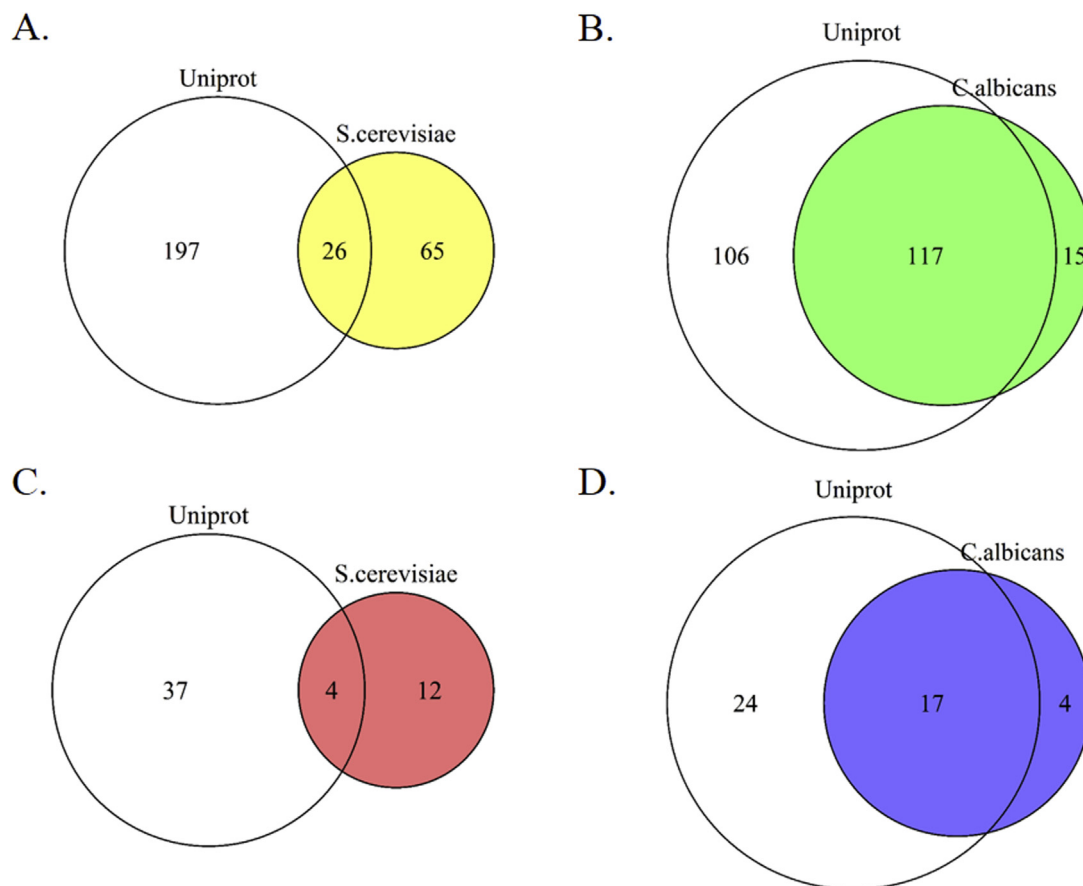


Fig. 1. Venn diagrams of gene ontology data from Cytoscape BiNGO showing enriched biological processes in the proteome from *Lasiodiplodia theobromae* using an all-Uniprot ontology database, compared to A; An all-*Saccharomyces cerevisiae* ontology database, and B; an all-*Candida albicans* ontology database. Also included are enriched molecular functions of the proteome from *L. theobromae* assessed with an all-Uniprot ontology database compared to molecular functions from C; an all-*S. cerevisiae* ontology database, as well as compared to molecular functions of D; an all-*C. albicans* ontology database. Venn diagrams were created with the “R”-based program VennDiagram.

quantity and quality of peptide information is unprecedented in proteomics studies of the fungal family Botryosphaeriaceae.

Using a conservative approach that adheres to established Paris Guidelines for proteomics [49], for the database-dependent fragmentation analysis, 224 peptide identification hits with 100% confidence were obtained from a Uniprot decoy database search with a 1% FDR (Supplementary data S2 in the Data in Brief article [35]). Of these, 76 protein hits were validated to 100% confidence with 2 unique peptides or high PSM number, and the remaining 148 identified with one unique peptide.

A special case is made of the identification of a protein homologous to human POTE ankyrin. Ankyrins are adaptor proteins that mediate the attachment of integral membrane proteins to the spectrin-actin based membrane cytoskeleton, and are poorly characterized in fungi [50–52]. One unique peptide was identified and validated to be homologous using an all-Uniprot database. To explore this further, the same data set was searched against a human-only Uniprot database, which yielded three different peptides that all matched POTE ankyrin (Uniprot accession number A5A3EO). In humans, the POTE ankyrins have been more thoroughly studied, are

considered primate-specific [53], and are expressed in testes, ovaries and prostate, as well as in embryonic stem cells, possessing both ankyrin and spectrin domains [52]. The presence of the ankyrin-binding protein spectrin has not been well-established in fungi, and little information exists on ankyrin-binding proteins in the fungal cytoskeleton [54,55], however, this is an example of proteomics serving to provide important clues for identifying novel proteins in poorly annotated organisms such as fungi that merit further research.

Seven hundred and forty-seven peptides yielded protein hits with 100% confidence using a Botryosphaeriaceae-only NCBI database for protein identification with the same data set. Of these, 361 proteins were validated to 100% confidence with at least two validated peptides (Supplementary data S3 in Data in Brief article [35]). Three hundred and eighty-six proteins were identified with 100% confidence with one validated peptide. Of those validated with two unique peptides, many proteins with important biotechnological applications were found, such as a variety of different fungal-specific alcohol dehydrogenases. For example, aryl alcohol dehydrogenase (NCBI accession 821064119) as well as saccharopine dehydrogenase (NCBI accession 821064554) were identified

homologous to *D. seriata*, the latter possessing a biochemical function specific to fungi (involved in the lysine synthesis pathway) and both potential targets for new antifungal compounds [56]. Aryl-alcohol dehydrogenase is involved in degrading lignin, and of biotechnological interest for the production of flavor compounds [57,58]. Along the lines of fungal-specific amino acid synthesis pathways that may serve as targets for new antifungals [59], a protein homologous to homocysteine synthase from the *de novo* methionine synthesis pathway was also identified (Uniprot ID# P50125).

Many proteins relevant to canonical metabolic pathways and fermentation (an important part of metabolism in microorganisms) were detected (Supplementary data S2–S5 in data article [35]). However, many enriched molecular functions never before reported in *L. theobromae* were found. Intriguingly, none of the enriched gene ontology categories yielded lipases as an enriched protein group using a 1% FDR search using fragmentation analysis. Not detecting more lipases was an unexpected result, since it is well known that lipases are induced by triglycerides in the medium in many fungi [60] and fatty acid ester analysis indicated a high production of a variety of these lipase/esterase-derived compounds by the fungus under the same cultivation conditions and carbon sources [24]. This is likely due to the fact that the genome from *L. theobromae* has not been sequenced, and evidently, the lipases detected in this work do not share homology with any lipases from sequenced fungal organisms.

Of the confidently identified proteins from the NCBI Botryosphaeriaceae-only database, thirty-one were identified as “hypothetical” proteins, i.e. proteins without known functions. The accession numbers from these were input into NCBI BLAST, and the hit with the highest max score was listed as a potential match (Table 1).

From the Uniprot database search, an unexpected annotation category found to be enriched is kininogen binding, with enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) homologous to *Candida albicans* identified and validated in this category to be produced by *L. theobromae*. Although binding targets in plants have not been identified, kininogens are known to possess physiological activity, and in humans are cleaved by proteases into their physiologically active form [61]. Besides having an important role in glycolysis, enolase from *Candida* spp. is present in biofilms and is known to bind to host cells and induce IgE-mediated allergy responses in humans [62]. In *Candida parapsilosis* and *Candida tropicalis*, it was shown that glycolytic enzymes, including enolase, are exposed at the surface of the fungus, and bind human host proteins such as laminin, vitronectin [63] and plasminogen [61]. *C. albicans* enolase is also involved in human respiratory fungal allergies and required for the colonization of the intestinal epithelium [62]. Although very little is known about enolases from *L. theobromae*, they are homologous to many allergenic enolases from other filamentous fungi were such as *Alternaria alternata* (Uniprot ID

Table 1
Proteins identified as “hypothetical” by the MSGF, X! Tandem and OMSSA algorithms, further searched with NCBI BLAST. The top-scoring homologous protein is reported for each search in the last column.

Main accession	Description	Coverage [%]	Spectrum counting	#Validated unique	Confidence [%]	Validation	Homologous protein sequence (BLAST)
407929156	Hypothetical protein MPH_00684 [<i>Macrophomina phaseolina</i> MS6]	6.0	30.1	2	100	Confident	<i>Neofusicoccum parvum</i> UCRNP2 putative protein mitochondrial targeting protein XM_007580308.1
407928922	Hypothetical protein MPH_00885 [<i>Macrophomina phaseolina</i> MS6]	17.1	228.8	3	100	Confident	<i>Neofusicoccum parvum</i> UCRNP2 putative phosphotransmitter protein ypd1 protein XM_007584095.1
407927556	Hypothetical protein MPH_02256 [<i>Macrophomina phaseolina</i> MS6]	9.1	124.1	4	100	Confident	<i>Neofusicoccum parvum</i> UCRNP2 putative nuclear and cytoplasmic polyadenylated rna-binding protein pub1 protein mRNA, XM_007585675.1
407923581	Hypothetical protein MPH_06104 [<i>Macrophomina phaseolina</i> MS6]	11.6	108.6	2	100	Confident	<i>Sphaerulina musiva</i> SO2202 GTP-binding protein mRNA XM_016910103.1
407922440	Hypothetical protein MPH_07279 [<i>Macrophomina phaseolina</i> MS6]	9.9	242.1	2	100	Confident	<i>Neofusicoccum parvum</i> UCRNP2 putative fk506-binding protein, XM_007581103.1
407921582	Hypothetical protein MPH_07998 [<i>Macrophomina phaseolina</i> MS6]	4.9	11.8	2	100	Confident	<i>Neofusicoccum parvum</i> UCRNP2 putative transcription factor (snd1 p100) protein, XM_007582992.1
407921297	Hypothetical protein MPH_08297 [<i>Macrophomina phaseolina</i> MS6]	12.3	74.8	3	100	Confident	<i>Neofusicoccum parvum</i> UCRNP2 putative g-protein complex beta subunit protein XM_007582859.1
407920925	Hypothetical protein MPH_08717 [<i>Macrophomina phaseolina</i> MS6]	23.4	128.1	7	100	Confident	<i>Neofusicoccum parvum</i> UCRNP2 putative glycolipid transfer protein het-c2 protein XM_007581974.1
407916430	Hypothetical protein MPH_13123 [<i>Macrophomina phaseolina</i> MS6]	21.2	581.6	4	100	Confident	<i>Neofusicoccum parvum</i> UCRNP2 putative surface protein 1 protein XM_007588394.1

Table 2
 Novel *de novo* peptide sequences from the Novor and PepNovo algorithms and BLASTp search results using an all-Uniprot database of the peptidome from *Lasiodiplodia theobromae* incubated in 5% grapeseed oil + 5% glucose for 20 days, showing the most probable homolog.

Query sequence	Peptide Score	<i>m/z</i>	Charge	E-value	Uniprot ID	Organism	Protein name
ELGFLASLGWNAPPAFPGPELTALDKLSAALSNR	28.7	1175.2	3+	2	Q8T9W1	<i>Dictyostelium discoideum</i> (Slime mold)	Serine protease/ABC transporter B family protein tagD (EC 3.4.21.-) (Serine protease/ABC transporter tagD)
DDGSMDTCEDDGAGR	43.1	808.8	2+	8.6	Q7RTY9	<i>Homo sapiens</i> (Human)	Serine protease 41 (EC 3.4.21.-) (Testis serine protease 1) (TESSP-1)
QETSVPGDSQATLDALNEWNR	66.5	739.7	3+	1.6	Q10749	<i>Naja mossambica</i> (Mozambique spitting cobra)	Snake venom metalloproteinase-disintegrin-like mocarhagin (MOC) (Mocarhagin-1) (SVMP) (EC 3.4.24.-) (Zinc metalloproteinase)
LLGTAMHVAGHLCAMMGWYRLTPSLHK	17.0	1033.8	3+	4.6	J3RY93	<i>Crotalus adamanteus</i>	Snake venom serine proteinase 12 (SVSP) (EC 3.4.21.-)
WLSAKGWAQGVAEQGDPR	35.9	991.5	2+	7.7	P0DL42	<i>Daboia siamensis</i> (Eastern Russel's viper)	Snake venom vascular endothelial growth factor toxin VR-1' (svVEGF) (VEGF-F)
DYPPVLEHLLAWEEEQR	19.0	708.7	3+	7	B2MVK7	<i>Rhynchium brunneum</i> (Potter wasp)	Venom allergen 5 (Antigen 5) (Cysteine-rich venom protein) (CRVP)
VLTVKGCSCNLTMLVSWTFSGNLTNGASTGSHK	37.4	1172.3	3+	5.7	C0ITL3	<i>Pachycondyla chinensis</i> (Asian needle ant)	Venom allergen 5 (Antigen 5) (Cysteine-rich venom protein) (CRVP) (allergen Pac c 3) (Fragment)
MLPAFLVDVHHSKKVGVPPNYYYQHSK	15.10	793.4	4+	7.1	P83370	<i>Hoplocephalus stephensii</i> (Stephens' banded snake)	Venom prothrombin activator hopsarin-D (vPA) (EC 3.4.21.6) (Venom coagulation factor Xa-like protease) [Cleaved into: Hopsarin-D light chain; Hopsarin-D heavy chain]
APTPALDGGKLMFVSAK	38.6	568.3	3+	0.63	Q5WFT5	<i>Bacillus clausii</i> (strain KSM-K16)	Zinc metalloprotease RasP (EC 3.4.24.-) (Regulating alternative sigma factor protease) (Regulating anti-sigma-W factor activity protease)
LASHKPNCFAPPLGTHDVVPK	34.5	800.1	3+	6	B8K1W0	<i>Daboia russelii</i> (Russel's viper) (<i>Vipera russelii</i>)	Zinc metalloproteinase-disintegrin-like daborhagin-K (EC 3.4.24.-) (Haemorrhagic metalloproteinase russelysin) (Snake venom metalloproteinase) (SVMP)
VSDSFTVGLAAQWAETDLYK	49.3	743.7	3+	7	C5H5D6	<i>Lachesis muta rhombeata</i> (Bushmaster)	Zinc metalloproteinase-disintegrin-like lachestatin-2 (EC 3.4.24.-) (Snake venom metalloprotease) (SVMP) (Vascular apoptosis-inducing protein-like) (VAP-like)
YGMKWSLLMLAAGGGR	54.8	863.9	2+	5.4	Q5XJ13	<i>Danio rerio</i> (Zebrafish) (Brachydanio rerio)	Ankyrin repeat and SAM domain-containing protein 6 (Ankyrin repeat domain-containing protein 14) (SamCystin)
LMAAWGQNPQAALYVEWNPEVMSAVWK	37.10	1054.8	3+	1.2	Q68DC2	<i>Homo sapiens</i> (Human)	Ankyrin repeat and SAM domain-containing protein 6 (Ankyrin repeat domain-containing protein 14) (SamCystin)
LGAKWCHHCFAVPMK	19.9	929.4	2+	0.95	Q6S8J3	<i>Homo sapiens</i> (Human)	POTE ankyrin domain family member E (ANKRD26-like family C member 1A) (Prostate, ovary, testis-expressed protein on chromosome 2) (POTE-2)
ALKGFYLMNPPAPVWWTR	49.6	716.4	3+	5.6	B8I4B9	<i>Clostridium cellulolyticum</i>	ATP-dependent zinc metalloprotease FtsH (EC 3.4.24.-)

(continued on next page)

Table 2 (continued)

Query sequence	Peptide Score	m/z	Charge	E-value	Uniprot ID	Organism	Protein name
LAAPNPYLWWHDLVPR	40.5	974.5	2+	6.7	Q2EEX7	<i>Helicosporidium</i> sp. subsp. <i>Simulium jonesii</i>	ATP-dependent zinc metalloprotease FtsH homolog
ELKAKKVGAMDVASSEFYK	52.7	701.04	3+	1.5	P30575	<i>Candida albicans</i>	Enolase 1 (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase)
TLVAGFFFLAASLTCLLPSE- APTFSETTR	45.0	812.4	4+	1.8	Q9HDT3	<i>Alternaria alternata</i>	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) (Allergen Alt a 11) (Allergen Alt a 5) (Allergen Alt a XI)
ELGAHTLGFLESNPKFDPGSY- EQLADLYK	53.2	1080.5	3+	1.7	Q96X30	<i>Neosartorya fumigata</i> (<i>Aspergillus fumigatus</i>)	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) (allergen Asp f 22)
HCLAAVPCMGQSPAAAALLNSSEVQNMAMGFK	40.4	1025.5	2+, 3+, 4+	1.6	Q59121	<i>Arcanobacterium haemolyticum</i>	Phospholipase D (EC 3.1.4.4) (Choline phosphatase) (PLD-A)
VFRLKDSGNSKPKDAAPPPGPLPR	29.2	848.8	3+	8.3	O14939	<i>Homo sapiens</i> (Human)	Phospholipase D2 (PLD 2) (hPLD2) (EC 3.1.4.4) (Choline phosphatase 2) (PLD1C) (Phosphatidylcholine-hydrolyzing phospholipase D2)
LHMGTDCAELDCVRSAPK	33.2	687.3	3+	4.5	Q9LI83	<i>Arabidopsis thaliana</i> (Mouse- ear cress)	Phospholipid-transporting ATPase 10 (AtALA10) (EC 3.6.3.1) (Aminophospholipid flippase 10)
DLADALKAAMGAPKHAVDHVAEPMQVLAQLAPGNK	36.9	867.7	4+	4.2	P06681	<i>Homo sapiens</i> (Human)	Complement C2 (EC 3.4.21.43) (C3/C5 convertase) [Cleaved into: Complement C2b fragment; Complement C2a fragment]
NTLHAPEGTVVHLNAHK	54.2	919.5	2+	8.7	P98094	<i>Eptatretus burgeri</i> (Inshore hagfish)	Complement C3 [Cleaved into: Complement C3 beta chain; Complement C3 alpha chain; C3a anaphylatoxin; Complement C3 gamma chain] (Fragment)
HNDRCPCMCNCQNDRCRCDTSNPAK	16.2	825.8	4+	6.3	Q00685	<i>Lethenteron camtschaticum</i> (Japanese lamprey) (<i>Lampetra japonica</i>)	Complement C3 [Cleaved into: Complement C3 beta chain; Complement C3 alpha chain; C3a anaphylatoxin; Complement C3 gamma chain] (Fragment)
SWFVNVPCDGCALAGMTASFVR	30.9	800.4	3+	9.7	P01031	<i>Homo sapiens</i> (Human)	Complement C5 (C3 and PZP-like alpha-2-macroglobulin domain-containing protein 4) [Cleaved into: Complement C5 beta chain; Complement C5 alpha chain; C5a anaphylatoxin; Complement C5 alpha chain]
KDLKYLDVYYWRYFWKWRQVDYLVAPAK	22.6	928.9	4+	4.1	P15304	<i>Rattus norvegicus</i> (Rat)	Hormone-sensitive lipase (HSL) (EC 3.1.1.79)
DPGFLGTLATASGKFGEGDPHAKTSGLNRK	40.0	1010.5	3+	0.21	A7L035	<i>Chironex fleckeri</i> (Box jellyfish)	Toxin CfTX-1 (Toxin 1)
TLMGYLLAVDGCALGCFVMPGLGLVAALSK	28.3	815.9	4+	3.2	P60262	<i>Tityus discrepans</i> (Venezuelan scorpion)	Toxin TdII-3 (Fragment)
SAPDPGTGDEVVMGQMDPR	37.1	658.9	3+	4.8	P42211	<i>Oryza sativa</i> subsp. japonica (Rice)	Aspartic proteinase (EC 3.4.23.-)
PAYHDLGVVPMADTR	41.7	821.4	2+	9.2	Q8VYV9	<i>Arabidopsis thaliana</i> (Mouse- ear cress)	Aspartyl protease family protein 1 (EC 3.4.23.-)

Q9HDT3), a known human allergen of clinical significance [64,65].

Although *L. theobromae* shares many biological processes with *Saccharomyces cerevisiae*, as well as molecular functions and cellular components, the proteins distinguished to be significantly enriched and involved in fermentation, such as alcohol dehydrogenases and pyruvate decarboxylase, were unique to fungal pathogens and clearly distinct to *S. cerevisiae* fermentation genes. For example, the enriched alcohol fermentation-specific proteins pyruvate decarboxylase and alcohol dehydrogenase from *L. theobromae* were found to be homologous to *Aspergillus* spp. and *Botryosphaeriaceae* spp., respectively.

Membrane-related processes such as binding and cell signaling were attributed to the 14-3-3 protein, which is known to act in a variety of important lipid/membrane signaling processes [66], and identified in *L. theobromae* to be expressed under the described cultivation conditions (See [Supplementary data S3](#) in Data in Brief article [35]). The glycolipid transfer protein Het-C2 was identified (only after manually searching the confidently identified hypothetical proteins) (Table 1). Het-C2 is also important in heterokaryon incompatibility signaling and programmed cell death in fungi [67]. The anti-viral lectin-type protein cyanovirin-N was detected as well, which is a carbohydrate-binding protein important in nutrient sensing [68], and of great biotechnological interest for use in anti-viral therapies.

As aforementioned, very few of the enriched categories were shared between *L. theobromae* and *S. cerevisiae*. Instead, the enriched categories from *L. theobromae* were more similar to the pathogenic yeast *C. albicans* (see full *S. cerevisiae*-specific ontology in [Supplementary data 5](#) in associated Data in Brief article [35] and Fig. 1).

The data provides insight into enzymatic differences between *L. theobromae* and the non-pathogen *S. cerevisiae*, as well as metabolic similarities between pathogens such as *L. theobromae* and *C. albicans*. All ferment glucose, but *L. theobromae*

evidently has fermentation enzymes that have evolved differently than either of these yeasts, and possesses a generally wider metabolic repertoire in comparison, as shown in the identified ontology categories analyzed via Venn Diagrams (Fig. 1).

De novo peptide sequencing and a BLASTp search against the entire Uniprot database yielded many novel proteins for *L. theobromae*. Out of the 5983 peptide hits, many peptides were homologous to lipases and toxins (Table 2 and [Supplementary information S6](#) in Data in Brief article [35]). The most probable homolog is reported, with relatively high E-values that are reflective of potentially uncharacterized, novel sequences in this fungus, and a lack of genomic sequences for Botryosphaeriaceae and filamentous fungi in general. *De novo* sequencing supported some findings from the database-dependent fragmentation analysis-based protein identifications, especially the finding of novel ankyrins, including the POTE ankyrin. Many peptides homologous to an assortment of proteases were detected, including some found in venom from a variety of insects and snakes, representing an area of new research possibilities in the host: pathogen interaction. The finding of a zinc metalloproteinase-disintegrin-like daborhagin-K (EC 3.4.24.-) is highly significant and novel because of the function of this type of protease in causing dermal hemorrhage in mice [69]. The role of this protein in *L. theobromae* remains to be studied.

An allergenic enolase was found with homology to the same in *A. alternata*, further evidence of immune system modulation by *L. theobromae* in an IgE-dependent manner (Table 2). More tellingly, a significant *de novo* peptide was found homologous to enolase from *C. albicans*, supporting and confirming the results from the database-dependent fragmentation analysis, with many potential novel functions in plants and humans, and ruling out database bias in the identification of this enolase via a Uniprot database-dependent fragmentation analysis.

The fact that the proteome and related molecular functions from *L. theobromae* more closely resembled *C. albicans*

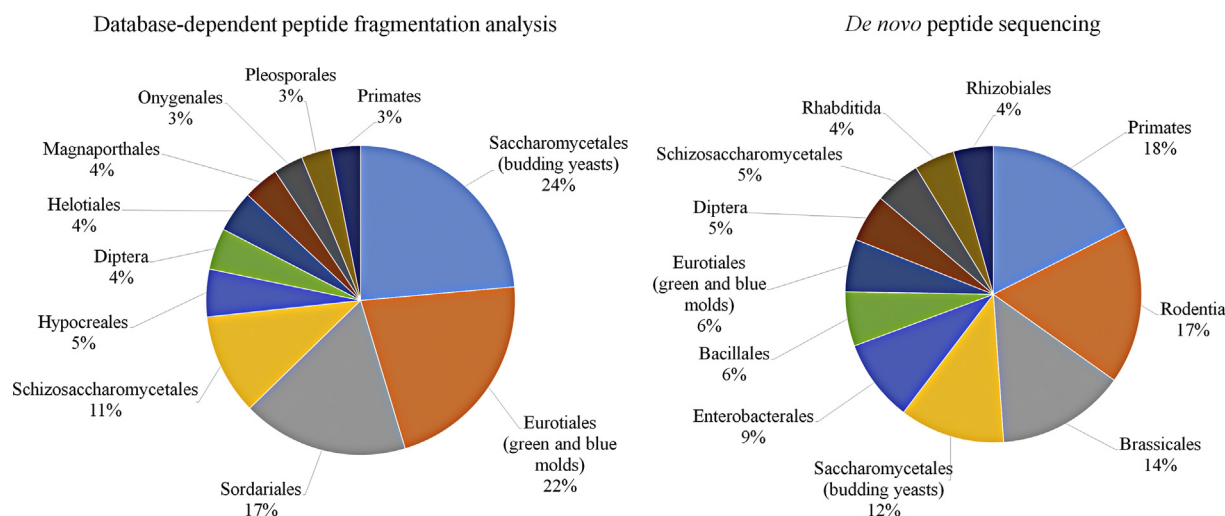


Fig. 2. Taxonomic order distribution of peptide hits using the entire Uniprot database from both peptidome analysis approaches, showing the top ten most represented orders in each peptide identification approach with most homology to the peptidome from *Lasiodiplodia theobromae*.

instead of *Saccharomyces cerevisiae* is of interest because *L. theobromae*, along with *Macrophomina phaseolina*, also a member of the Botryosphaeriaceae, is reported to be able to infect both plants and humans [18], making this an avenue for further research for understanding host-specific pathogenicity. It is proposed that finding common proteins or physiological functions between known human pathogens such as *Candida* spp. and plant pathogens that are also able to infect humans, may be the key to understand cross-domain colonization. For example, the enolase in *L. theobromae* may be a part of *L. theobromae*'s pathogenicity in both plants and humans in a manner similar to *C. albicans*, potentially causing IgE-based allergic reactions in humans, or binding compounds in plants critical to normal physiological functioning.

Many other peptides sequenced in this work were found to be involved in immune system regulation (such as peptides homologous to complement C3). This indicates an ancient conserved proteolytic system and the ability of this fungus to manipulate the host's immune system as an exciting avenue of research in cross-domain colonization by pathogenic fungi. The finding of this type of protein is of interest in this case especially because in humans, it is known to contribute to inflammation and lipid metabolism [70]. The complement C3 version detected here includes the type cleaved into the anaphylatoxin version of this protein, which is also involved in causing histamine release from mast cells and enhanced vascular permeability and various other processes in the innate immune response [71]. In humans, C3 production is also activated by chylomicrons [72], also involved in lipid metabolism. The function of this protein in *L. theobromae* remains to be elucidated, although its role in lipid metabolism may be especially relevant in exogenous lipid digestion by *L. theobromae*, an important substrate for fungi and one of the substrates used for this experiment.

Gene ontology of the *de novo* sequencing results revealed binding and catalytic activities to be enriched, as found in the database-dependent protein identification results (Fig. 1A and B and Supplementary information S6 in Data in Brief article [35]). Both peptide identification methods demonstrated peptides that were homologous to a variety of taxonomic orders (Fig. 2).

Uniprot database-dependent fragmentation analysis showed homology mostly to Saccharomycetales (budding yeasts), followed by homology to Eurotiales (green and blue molds), and Sordariales, whereas *de novo* peptide sequencing generated peptides with homology to mostly Primates, followed by Rodentia, Brassicales (flowering plants) and Saccharomycetales (budding yeasts). These differences are thought to arise from a lack of sequence information in the Uniprot database for filamentous fungi (Ascomycetes) from different taxonomic orders. Additionally, the differences are likely due to uncharacterized and unannotated sequences in *L. theobromae*. These dissimilarities are expected to be resolved as the genome sequence from *L. theobromae* and other Botryosphaeriaceae are sequenced and annotated.

It brings to attention the fact that genomic sequencing and detailed, functional characterization efforts are crucial for

effectively strengthening bioinformatics databases and for implementing powerful technological platforms such as proteomics, especially because protein expression studies are critical for pursuing promising new avenues of research in fungal biochemistry. These efforts will lead to the characterization and annotation of new genes, and ontological categories attributed to each gene, which will allow for more thorough protein expression studies in any experimental context. Apart from genomic sequencing, characterization efforts in the form of cloning and purification, or GFP-tagging and cellular localization efforts would be the next step forward from this work. A comprehensive bioinformatics study such as this (the first of its kind for Botryosphaeriaceae), is highly useful, and in this case has allowed the targeting of many new proteins for further research, such as zinc metalloproteinases and lipases. The data is freely available to the scientific community interested in this area of fungal biochemistry and cellular biology, and will be re-visited once the genome from *L. theobromae* becomes available.

4. Conclusions

This proteomic exploration has led to the identification of host *binding* as a critically important aspect in fungal pathogenicity. Protein targets for developing strategies to control *L. theobromae* have also been identified. Novel proteins have been detected that merit further research as unexplored paradigms in fungal biochemistry, such as peptides with homology to ankyrins, and particularly the POTE ankyrin from humans. Novel zinc metalloproteinases homologous to those from snake venom were found, among many other interesting proteins such as lipases, peptides homologous to complement from the human immune system, and allergenic enolases that await further functional description.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors state that there are no conflicts of interest in this work.

Acknowledgments

This work was made possible by a doctoral stipend from CONACyT and a UCMEXUS fellowship for Carla Uranga. Thanks to Dr. Elizabeth Komives from the UCSD proteomics mass spectrometry department for UCMEXUS support and sponsorship. Special thanks to Dr. Kristina Herbert and Dr. Genaro Pimienta for “omica” server establishment and access at CICESE. We are grateful to Dr. Marc Vaudel and Dr. Harald Barsnes for help with the SearchGUI and DeNovoGUI programs. Thanks to Claudio Espinosa de los Monteros for help with Adobe Photoshop.

Appendix A. Supplementary material

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.biopen.2017.03.001>.

References

- [1] M. Blackwell, The fungi: 1, 2, 3... 5.1 million species? *Am. J. Bot.* 98 (3) (2011) 426–438.
- [2] R. Brambl, Fungal physiology and the origins of molecular biology, *Microbiology* 155 (12) (2009) 3799–3809.
- [3] C.M. Carnielli, F.V. Winck, A.F. Paes Leme, Functional annotation and biological interpretation of proteomics data, *Biochim. Biophys. Acta – Proteins Proteomics* 1854 (1) (2015) 46–54.
- [4] G.D. Findlay, M.J. Maccoss, W.J. Swanson, Proteomic discovery of previously unannotated, rapidly evolving seminal fluid genes in *Drosophila*, *Genome Res.* 19 (2009) 886–896.
- [5] A. Ambrogelly, S. Palioura, D. Söll, Natural expansion of the genetic code, *Nat. Chem. Biol.* 3 (1) (2007) 29–35.
- [6] K. Grützmann, K. Szafranski, M. Pohl, K. Voigt, A. Petzold, S. Schuster, Fungal alternative splicing is associated with multicellular complexity and virulence: a genome-wide multi-species study, *DNA Res.* 21 (1) (2014) 27–39.
- [7] P. Álvarez-Chaver, O. Otero-Estévez, M.P. de la Cadena, F.J. Rodríguez-Berrocal, V.S. Martínez-Zorzano, Proteomics for discovery of candidate colorectal cancer biomarkers, *World J. Gastroenterol.* 20 (14) (2014) 3804–3824.
- [8] S. Pevtsov, I. Fedulova, H. Mirzaei, C. Buck, X. Zhang, Performance evaluation of existing de novo sequencing algorithms, *J. Proteome Res.* 5 (11) (2006) 3018–3028.
- [9] M. Gırdea, L. Noé, G. Kucherov, Back-translation for discovering distant protein homologies, in: S.L. Salzberg, T. Warnow (Eds.), *Algorithms in Bioinformatics: 9th International Workshop, WABI 2009, Philadelphia, PA, USA, September 12–13, 2009 Proceedings*, Springer, Berlin, Heidelberg, 2009, pp. 108–120.
- [10] J.R. Urbez-Torres, G.M. Leavitt, J.C. Guerrero, J. Guevara, W.D. Gubler, Identification and pathogenicity of *Lasiodiplodia theobromae* and *Diplodia seriata*, the causal agents of bot canker disease of grapevines in Mexico, *Plant Dis.* 92 (4) (2008) 519–529.
- [11] M.A. Khazada, A.M. Lodhi, S. Shahzad, Chemical control of *Lasiodiplodia theobromae*, the causal agent of mango decline in Sindh, *Pak. J. Bot.* 37 (4) (2005) 1023–1030.
- [12] M. Mbenoun, E.H. Momo Zeutsa, G. Samuels, F. Nsouga Amougou, S. Nyasse, Dieback due to *Lasiodiplodia theobromae*, a new constraint to cocoa production in Cameroon, *Plant Pathol.* 57 (2) (2008) 381.
- [13] J.R. Urbez-Torres, The status of *Botryosphaeriaceae* species infecting grapevines, *Phytopathol. Mediterr.* 50 (Suppl.) (2011) 5–45.
- [14] J.R. Urbez-Torres, W.D. Gubler, Pathogenicity of *Botryosphaeriaceae* species isolated from grapevine cankers in California, *Plant Dis.* 93 (6) (2009) 584–592.
- [15] N.A. Nghia, V. Thi, Q. Chi, N.X. Dong, Molecular analysis of *Botryodiplodia theobromae* isolates from rubber in Vietnam using rDNA its sequencing and ISSR markers, *Int. Rubber Conf.* (2012) 28–31.
- [16] A.R. Machado, O.L. Pereira, Major diseases of the biofuel plant, physic nut (*Jatropha curcas*), *Biodiesel Feed Prod. Appl.* (2012) 18.
- [17] P.C.Y. Woo, S.K.P. Lau, A.H.Y. Ngan, H. Tse, E.T.K. Tung, K.Y. Yuen, *Lasiodiplodia theobromae* pneumonia in a liver transplant recipient, *J. Clin. Microbiol.* 46 (1) (2008) 380–384.
- [18] L.J. Papacostas, A. Henderson, K. Choong, D. Sowden, An unusual skin lesion caused by *Lasiodiplodia theobromae*, *Med. Mycol. Case Rep.* 8 (2015) 44–46.
- [19] R. Cobos, C. Barreiro, R.M. Mateos, J.-J.R. Coque, Cytoplasmic- and extracellular-proteome analysis of *Diplodia seriata*: a phytopathogenic fungus involved in grapevine decline, *Proteome Sci.* 8 (2010) 46.
- [20] L. Bianco, G. Perrotta, Methodologies and perspectives of proteomics applied to filamentous fungi: from sample preparation to secretome analysis, *Int. J. Mol. Sci.* 16 (3) (2015) 5803–5829.
- [21] V. Bhadauria, Y.-L. Peng, Optimization of a protein extraction technique for fungal proteomics, *Indian J. Microbiol.* 50 (Suppl. 1) (2010) 127–131.
- [22] S. Kumar Bhuyan, P. Bandyopadhyay, P. Kumar, D. Kumar Mishra, R. Prasad, A. Kumari, et al., Interaction of *Piriformospora indica* with *Azotobacter chroococcum*, *Sci Rep.* 5 (2015) 13911.
- [23] J. Folch, M. Lees, G.H. Sloan Stanley, A simple method for the isolation and purification of total lipides from animal tissues, *J. Biol. Chem.* 226 (1957) 497–509.
- [24] C.C. Uranga, J. Beld, A. Mrse, I. Córdova-Guerrero, M.D. Burkart, R. Hernández-Martínez, Fatty acid esters produced by *Lasiodiplodia theobromae* function as growth regulators in tobacco seedlings, *Biochem. Biophys. Res. Commun.* 472 (2) (2016).
- [25] W. Bester, P.W. Crous, P.H. Fourie, Evaluation of fungicides as potential grapevine pruning wound protectants against *Botryosphaeria* species, *Australas. Plant Pathol.* 36 (1) (2007) 73–77.
- [26] J. Moral, C. Muñoz-Díez, N. González, A. Trapero, T.J. Michailides, Characterization and pathogenicity of *Botryosphaeriaceae* species collected from olive and other hosts in Spain and California, *Phytopathology* 100 (9) (2010) 1340–1351.
- [27] N. Wunderlich, G.J. Ash, C.C. Steel, H. Raman, S. Savocchia, Association of *Botryosphaeriaceae* grapevine trunk disease fungi with the reproductive structures of *Vitis vinifera*, *Vitis – J. Grapevine Res.* 50 (2) (2011) 89–96.
- [28] R. Craig, R.C. Beavis, TANDEM: matching proteins with tandem mass spectra, *Bioinformatics* 20 (9) (2004) 1466–1467.
- [29] S. Kim, P.A. Pevzner, MS-GF+ makes progress towards a universal database search tool for proteomics, *Nat. Commun.* 5 (2014) 5277.
- [30] L.Y. Geer, S.P. Markey, J.A. Kowalak, L. Wagner, M. Xu, D.M. Maynard, et al., Open mass spectrometry search algorithm research articles, *J. Proteome Res.* (2004) 1–25.
- [31] J.K. Eng, T.A. Jahan, M.R. Hoopmann, Comet: an open-source MS/MS sequence database search tool, *Proteomics* 13 (1) (2013) 22–24.
- [32] M. Vaudel, H. Barsnes, F.S. Berven, A. Sickmann, L. Martens, Search-GUI: an open-source graphical user interface for simultaneous OMSSA and X!Tandem searches, *Proteomics* 11 (5) (2011) 996–999.
- [33] A. Bateman, M.J. Martin, C. O'Donovan, M. Magrane, R. Apweiler, E. Alpi, et al., UniProt: a hub for protein information, *Nucleic Acids Res.* 43 (D1) (2015) D204–D212.
- [34] L.Y. Geer, A. Marchler-Bauer, R.C. Geer, L. Han, J. He, S. He, et al., The NCBI BioSystems database, *Nucleic Acids Res.* 38 (Suppl. 1) (2009) 492–496.
- [35] C.C. Uranga, M. Ghassemian, R. Hernandez-Martinez, Data from proteome analysis of *Lasiodiplodia theobromae* (*Botryosphaeriaceae*), *Data Br.* (2017).
- [36] M. Vaudel, J.M. Burkhardt, R.P. Zahedi, E. Oveland, F.S. Berven, A. Sickmann, et al., PeptideShaker enables reanalysis of MS-derived proteomics data sets, *Nat. Biotechnol.* 33 (1) (2015) 22–24.
- [37] M. Vaudel, D. Breiter, F. Beck, J. Rahnenführer, L. Martens, R.P. Zahedi, D-score: a search engine independent MD-score, *Proteomics* 13 (6) (2013) 1036–1041.
- [38] S.A. Beausoleil, J. Villén, S.A. Gerber, J. Rush, S.P. Gygi, A probability-based approach for high-throughput protein phosphorylation analysis and site localization, *Nat. Biotechnol.* 24 (10) (2006) 1285–1292.
- [39] H. Barsnes, M. Vaudel, N. Colaert, K. Helsens, A. Sickmann, F.S. Berven, et al., compomics-utilities: an open-source Java library for computational proteomics, *BMC Bioinformatics* 12 (March) (2011) 70.
- [40] J. Vizcaíno, E.E.W. Deutsch, R. Wang, J.A. Vizcaino, E.E.W. Deutsch, R. Wang, et al., ProteomeXchange provides globally coordinated proteomics data submission and dissemination, *Nat. Biotechnol.* 32 (3) (2014) 223–226.
- [41] L. Martens, H. Hermjakob, P. Jones, M. Adamsk, C. Taylor, D. States, et al., PRIDE: the proteomics identifications database, *Proteomics* 5 (13) (2005) 3537–3545.

- [42] R. Christmas, I. Avila-Campillo, H. Bolouri, B. Schwikowski, M. Anderson, R. Kelley, et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Am. Assoc. Cancer Res. Educ. B* (2005) 12–16 (Karp. 2001).
- [43] S. Maere, K. Heymans, M. Kuiper, BiNGO: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks, *Bioinformatics* 21 (16) (2005) 3448–3449.
- [44] H. Chen, P.C. Boutros, VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R, *BMC Bioinformatics* 12 (1) (2011) 35.
- [45] T. Muth, L. Weilnböck, E. Rapp, C.G. Huber, L. Martens, M. Vaudel, et al., DeNovoGUI: an open source graphical user interface for de novo sequencing of tandem mass spectra, *J. Proteome Res.* 13 (2) (2014) 1143–1146.
- [46] B. Ma, Novor: real-time peptide de novo sequencing software, *J. Am. Soc. Mass Spectrom.* 26 (11) (2015) 1885–1894.
- [47] A. Frank, P. Pevzner, PepNovo: de novo peptide sequencing via probabilistic network modeling, *Anal. Chem.* 77 (4) (2005) 964–973.
- [48] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, et al., BLAST+: architecture and applications, *BMC Bioinformatics* 10 (2009) 421.
- [49] R.A. Bradshaw, A.L. Burlingame, S. Carr, R. Aebersold, Reporting protein identification data: the next generation of guidelines, *Mol. Cell. Proteomics* 5 (5) (2006) 787–788.
- [50] V. Bennett, J. Healy, Membrane domains based on ankyrin and spectrin associated with cell–cell interactions, *Cold Spring Harb. Perspect. Biol.* 1 (6) (2009) 1–17.
- [51] C.N. Leetola, M. Knight, D. Cascio, S. Hoffman, J.U. Bowie, P. Wilson, et al., Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3, *BMC Struct. Biol.* 14 (1) (2014) 17.
- [52] T.K. Bera, N. Huynh, H. Maeda, B.K. Sathyanarayana, B. Lee, I. Pastan, Five POTE paralogs and their splice variants are expressed in human prostate and encode proteins of different lengths, *Gene* 337 (Suppl.) (2004) 45–53.
- [53] T.K. Bera, A. Saint Fleur, D. Ha, M. Yamada, Y. Lee, B. Lee, et al., Selective POTE paralogs on chromosome 2 are expressed in human embryonic stem cells, *Stem Cells Dev.* 17 (2) (2008) 325–332.
- [54] M. Cotado-Sampayo, M. Ojha, R. Ortega-Pérez, M.L. Chappuis, F. Barja, Proteolytic cleavage of a spectrin-related protein by calcium-dependent protease in *Neurospora crassa*, *Curr. Microbiol.* 53 (4) (2006) 311–316.
- [55] M. Cotado-Sampayo, P.O. Ramos, R.O. Perez, M. Ojha, F. Barja, Specificity of commercial anti-spectrin antibody in the study of fungal and Oomycete spectrin: cross-reaction with proteins other than spectrin, *Fungal Genet. Biol.* 45 (6) (2008) 1008–1015.
- [56] K. Jastrzębowska, I. Gabriel, Inhibitors of amino acids biosynthesis as antifungal agents, *Amino Acids* 47 (2) (2015) 227–249.
- [57] A. Hernández-Ortega, P. Ferreira, A.T. Martínez, Fungal aryl-alcohol oxidase: a peroxide-producing flavoenzyme involved in lignin degradation, *Appl. Microbiol. Biotechnol.* 93 (4) (2012) 1395–1410.
- [58] T.D.H. Bugg, M. Ahmad, E.M. Hardiman, R. Rahmanpour, Pathways for degradation of lignin in bacteria and fungi, *Nat. Prod. Rep.* 28 (12) (2011) 1883–1896.
- [59] D. Ubhi, G. Kago, A.F. Monzingo, J.D. Robertus, Structural analysis of a fungal methionine synthase with substrates and inhibitors, *J. Mol. Biol.* 426 (8) (2014) 1839–1847.
- [60] V.L. Colin, M.D. Baigori, L.M. Pera, Effect of environmental conditions on extracellular lipases production and fungal morphology from *Aspergillus niger* MYA 135, *J. Basic Microbiol.* 50 (1) (2010) 52–58.
- [61] J.D. Crowe, I.K. Sievwright, G.C. Auld, N.R. Moore, N.A.R. Gow, N.A. Booth, *Candida albicans* binds human plasminogen: identification of eight plasminogen-binding proteins, *Mol. Microbiol.* 47 (6) (2003) 1637–1651.
- [62] R.C. Silva, A.C.B. Padovan, D.C. Pimenta, R.C. Ferreira, C.V. da Silva, M.R.S. Briones, Extracellular enolase of *Candida albicans* is involved in colonization of mammalian intestinal epithelium, *Front. Cell. Infect. Microbiol.* 4 (June) (2014) 66.
- [63] A. Kozik, M. Gogol, O. Bochenska, J. Karkowska-Kuleta, N. Wolak, W. Kamysz, et al., Kinin release from human kininogen by 10 aspartic proteases produced by pathogenic yeast *Candida albicans*, *BMC Microbiol.* 15 (1) (2015) 60.
- [64] I. Postigo, A. Gutiérrez-Rodríguez, J. Fernández, J.A. Guisantes, E. Suñín, J. Martínez, Diagnostic value of Alt a 1, fungal enolase and manganese-dependent superoxide dismutase in the component-resolved diagnosis of allergy to pleosporaceae, *Clin. Exp. Allergy* 41 (3) (2011) 443–451.
- [65] I. Kustrzeba-Wójcicka, E. Siwak, G. Terlecki, A. Wolańczyk-Mędrala, W. Mędrala, *Alternaria alternata* and its allergens: a comprehensive review, *Clin. Rev. Allergy Immunol.* 47 (3) (2014) 354–365.
- [66] U.F. Benning, B. Tamot, B.S. Guelette, S. Hoffmann-Benning, New aspects of Phloem-mediated long-distance lipid signaling in plants, *Front. Plant Sci.* 3 (March) (2012) 53.
- [67] C. Hall, J. Welch, D.J. Kowbel, N.L. Glass, Evolution and diversity of a fungal self/nonself recognition locus, *PLoS ONE* 5 (11) (2010).
- [68] L.M.I. Koharudin, A.R. Viscomi, J. Jee, S. Ottonello, A.M. Gronenborn, The evolutionarily conserved family of cyanovirin-N homologs: structures and carbohydrate specificity, *Structure* (2008) 570–584.
- [69] Chen H. Sen, H.Y. Tsai, Y.M. Wang, I.H. Tsai, P-III hemorrhagic metalloproteinases from Russell's viper venom: cloning, characterization, phylogenetic and functional site analyses, *Biochimie* 90 (10) (2008) 1486–1498.
- [70] A. Barbu, O.A. Hamad, L. Lind, K.N. Ekdahl, B. Nilsson, The role of complement factor C3 in lipid metabolism, *Mol. Immunol.* 67 (1) (2015) 101–107.
- [71] J.R. Dunkelberger, W. Song, Complement and its role in innate and adaptive immune responses, *Cell Res.* 20 (1) (2010) 34–50.
- [72] T. Fujita, T. Fujioka, T. Murakami, A. Satomura, Y. Fuke, K. Matsumoto, Chylomicron accelerates C3 tick-over by regulating the role of factor H, leading to overproduction of acylation stimulating protein, *J. Clin. Lab. Anal.* 21 (2) (2007) 71–76.