



HAL
open science

Comparative analysis of five *Mucor* species transcriptomes

Annie Lebreton, Laurence Meslet-Cladière, Stéphanie Morin-Sardin, Emmanuel Coton, Jean-Luc Jany, Georges Barbier, Erwan Corre

► **To cite this version:**

Annie Lebreton, Laurence Meslet-Cladière, Stéphanie Morin-Sardin, Emmanuel Coton, Jean-Luc Jany, et al.. Comparative analysis of five *Mucor* species transcriptomes. *Genomics*, 2019, 111, pp.1306 - 1314. 10.1016/j.ygeno.2018.09.003 . hal-03488854

HAL Id: hal-03488854

<https://hal.science/hal-03488854>

Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Comparative analysis of five *Mucor* species transcriptomes.

Lebreton Annie¹, Meslet-Cladière Laurence¹, Morin-Sardin Stéphanie¹, Coton Emmanuel¹,
Jany Jean-Luc¹, Barbier Georges¹, Corre Erwan²

1: Université de Brest, EA 3882 Laboratoire Universitaire de Biodiversité et Ecologie
Microbienne, ESIAB, Technopôle Brest-Iroise, 29280 Plouzané, France.

2: Station biologique de Roscoff, plateforme ABiMS, CNRS: FR2424, Sorbonne Université
(UPMC), 29682 Roscoff, France

†Corresponding author: Corre Erwan

Abstract

Mucor species belong to the Mucorales order within the Mucoromycota phylum, an early diverging fungal lineage. Although *Mucor* species are often ubiquitous some species have been reported to specifically occur in certain ecological niches. In this study, similarities and differences of a representative set of *Mucor* species with contrasted lifestyles were investigated at the transcriptome level. Five strains pertaining to five different species were studied, namely *M. fuscus* and *M. lanceolatus*, two species used in cheese production (during ripening), *M. racemosus*, a recurrent cheese spoiler sometimes described as an opportunistic pathogen, *M. circinelloides*, often described as an opportunistic pathogen and *M. endophyticus*, a plant endophyte. A core transcriptome was delimited and a phylogenetic analysis led to an altered phylogenetic placement of *M. endophyticus* compared to previously published topologies. Interestingly, the core transcriptome comprising 5566

26 orthogroups included genes potentially involved in secondary metabolism. As expected,
27 given the wide taxonomic range investigated, the five transcriptomes also displayed
28 specificities that can be, for some of them, linked to the different lifestyles such as
29 differences in the composition of transcripts identified as virulence factors or carbohydrate
30 transporters.

31

32 Introduction

33 Within the filamentous fungi, the *Mucor* genus belongs to the Mucorales order within the
34 Mucoromycota phylum, an early diverging fungal lineage [1]. *Mucor* species are common
35 and often ubiquitous [2, 3], their fast growing and high sporulating mycelium, consisting of
36 coenocytic hyphae, are encountered in a large variety of environments, with the exception of
37 low water activities (a_w) substrates. Growth of several *Mucor* species has been documented
38 to be limited to relatively high a_w (> 0.90) [4]. The *Mucor* genus mainly comprises mesophilic
39 species but also some thermotolerant and thermophilic species [2], some of them being
40 animal and human opportunistic pathogens responsible for mucormycoses [5, 6] which are
41 increasingly frequent, especially in immunocompromised patients [7]. *Mucor* spp. are mostly
42 saprobes, with some species being described as plant endophytes [8].

43 Interestingly, several *Mucor* species have an obvious biotechnological interest, for
44 metabolite production (e.g., biofuels) and biotransformations (e.g., terpenoid
45 biotransformations) but also in food production, especially in fermented Asian and African
46 food but also in cheese ripening (e.g. Tommes or Saint-Nectaire in France) (for a review,
47 see [9]). Since *Mucor* strains used for cheese ripening can be considered as technological
48 and have been only described so far in cheese, the question of their potential adaptation to
49 this matrix has been raised [9]. An adaptation hypothesis in cheese technological strains
50 was supported by the results of a recent study that showed that, contrary to other *Mucor*

51 strains tested (*M. racemosus*, *M. circinelloides*, *M. brunneogriseus*, *M. spinosus* and *M.*
52 *endophyticus*), *M. lanceolatus* and *M. fuscus* (technological strains) showed higher optimal
53 growth rates (μ_{opt}) on cheese matrices than on Potato Dextrose Agar (PDA) medium [4].
54 Moreover, lag times of the *M. endophyticus* endophyte strain were strongly extended on
55 cheese related matrices. The apparent adaptation to the cheese environment of *M.*
56 *lanceolatus* was also confirmed by morphological observations as well as by a higher ratio of
57 over accumulated proteins on Cheese agar versus PDA [2].

58 A recent large effort to generate genome data concerning the early diverging fungi has
59 helped refine their taxonomy [10] and has shed new light on *Mucor* genome evolution and
60 functions such as sensory perception [11], lipid metabolism [12] or pathogenesis [13]. The
61 lifestyle diversity within the genus *Mucor* offers interesting perspectives to better understand
62 evolutive adaptation to different life modes, e.g. saprobic, pathogenic and even adaptation to
63 anthropogenic conditions. The present study aimed to provide an overview of the common or
64 specific patterns of gene expression of five *Mucor* species with contrasting lifestyles, grown
65 in standard fungal culture.

66 Materials and methods

67 *Mucor* strains, culture conditions

68 Five *Mucor* strains belonging to different species reported to have contrasting lifestyles and
69 habitats, namely *M. fuscus* (MF) and *M. lanceolatus* (ML) (two species used in cheese
70 ripening), *M. racemosus* (MR) (a recurrent cheese spoiler), *M. circinelloides* (MC) (often
71 classified as an opportunistic pathogen) and *M. endophyticus* (ME) (a plant endophyte
72 species) were used in the present study (Table 1). They were obtained from the Université
73 de Bretagne Occidentale Culture Collection, (UBOCC, France; [http://www.univ-](http://www.univ-brest.fr/ubocc)
74 [brest.fr/ubocc](http://www.univ-brest.fr/ubocc)) or ordered from the Westerdijk Fungal Biodiversity Institute (The Netherlands,
75 <http://www.westerdijkinstitute.nl/>). Strains were maintained in the dark at 25 °C on PDA

76 medium (Difco Laboratories, Detroit, Michigan). Spore suspensions of each strain were
77 produced as previously described by Morin-Sardin *et al.* (2016) [4]. Concentrations were
78 adjusted to 10^7 to 10^8 spores·mL⁻¹ prior to storage at -80 °C until use. For RNA extraction,
79 each fungal strain was grown on PDA solid medium at 25°C for 7 days. The whole organism
80 was collected for sequencing.

81 **RNA extraction and sequencing**

82 Total RNA extraction from each strain thallus was performed using the RNeasy plant mini kit
83 (Qiagen, Courtaboeuf, France) following the manufacturer's instructions. The RNA-seq
84 libraries were prepared from total RNA using the Illumina Ribo-Zero rRNA Removal Kit
85 (Epicenter, Madison, WI) and converted to cDNA with the TruSeq RNA Sample Preparation
86 Kit (Illumina, San Diego, CA) following the manufacturer's instructions. The libraries
87 containing the cDNA from each sample were sequenced on an Illumina HiSeq 2000 system
88 (Illumina) with a sequencing configuration for 100 bp paired-end reads.

89 RNAseq quality check and assembly

90 Data quality of the sequencing files was checked using the FastQC software [14]. For each
91 strain, *de novo* transcriptome was reconstructed using Trinity (release 2014-07-17; [15]) with
92 the '-trimmomatic -normalize_reads -normalize_by_read_set' options. Weakly expressed
93 transcripts (isoform percentage < 1 and fragment per kilobase of transcript per million
94 mapped reads (FPKM) < 1) were removed from the dataset with the RSEM tool [16] included
95 in the Trinity package. For each Trinity-predicted "gene" only the most abundant isoform was
96 conserved in the final dataset.

97 Transcript annotation

98 Eukaryotic and prokaryotic ribosomal RNA were screened with Rnammer (v.1.2; [17]).
99 Transcript open reading frames (ORFs) were predicted using Transdecoder (r2014-07-04,
100 available at <http://transdecoder.github.io>), protein domains were predicted with HMMER
101 (v.3.1b1, [18]) based on the PFAM-A database [19], transmembrane helices were identified

102 with tmhmm (v.2.0, [20]), signal peptide cleavage sites were predicted with SignalP (v.4.1,
103 [21]). All predicted transcripts and predicted proteins were identified using respectively
104 BLASTx and BLASTp [22] against Swissprot-Uniprot (r2017-01-18) and Uniref90 (r201611-
105 02) databases and against the filtered predicted proteomes of *M. circinelloides* CBS 277.49
106 (v2, [11]), *Rhizopus oryzae* 99-880 [23] and *Phycomyces blackesleeenanus* NRRL1555 (v2,
107 [11]) available on the Joint Genome Institute (JGI) database (genome.jgi.doe.gov/). An E-
108 value cut-off of 1e-4 was used for all BLASTp and BLASTx analysis, only the best match per
109 database and protein was conserved. Putative functions were assigned to predicted proteins
110 using the Gene Ontology (GO) database [24] and the Enzyme Commission number (EC)
111 classification [25]. Predicted proteins were assigned to GO categories by transferring the
112 annotation of BLAST matches and protein domains. EC numbers were assigned by profile
113 search with PRIAM (r2015-03-04, [26]) and transferred from the predicted proteomes BLAST
114 matches.

115 Transcriptome quality check

116 In order to assess the quality of the transcriptomes, statistics such as percentage of
117 realigned reads and N50 were calculated. Transcriptome completeness in terms of gene
118 content was assessed by searching with BUSCO (Benchmarking Universal Single-Copy
119 Orthologs, v2, <http://busco.ezlab.org/>, [27]) for the presence/absence of the conserved
120 eukaryotic orthologous genes available in the BUSCO software.

121 Comparative transcriptomics

122 Significant differences in GO category occurrence among species were identified with an
123 internal script using the Fisher's exact test (pvalue < 0.05, available on github). All EC
124 numbers for each species were mapped onto metabolic networks with iPath2 [28]. When a
125 species specific EC number was found, the transcript annotation was checked manually. All
126 predicted proteomes were compared against each other based on sequence similarity to
127 identify orthologous proteins using the software Orthofinder v.1.1.2 [29] (E-value 1e-2,

128 inflation 1.5). Orthologous proteins were clustered according to the reported lifestyle of the
129 producing fungal species. Clusters were then compared: cheese/non-cheese, pathogen/non-
130 pathogen and core transcriptome/non-endophyte (orthogroups composed of proteins of all
131 species except ME). GO categories were assigned to orthogroup by transfer of protein
132 annotation. A single copy of each annotation was kept to avoid annotation redundancy.

133 Phylogenomic analysis

134 Single copy orthologs shared by the five studied *Mucor* spp. as well as *Rhizopus oryzae*
135 strain 99-880 and *Phycomyces blackesleeianus* strain NRRL1555 (both later species being
136 considered as outgroups) were kept for phylogenetic reconstruction.
137 For each of the 1289 obtained clusters, a multiple alignment was inferred using PRANK
138 v.170427 [30], run with default settings. Spuriously aligned regions were excluded with trimAl
139 v1.4.r15 [31] with a 0.2 gap threshold. Subsequent alignments were concatenated in a
140 supermatrix of 727479 sites. This matrix was used to reconstruct species tree by maximum
141 likelihood inference and by Bayesian Monte Carlo Markov Chain (MCMC) samples. RAxML
142 PTHREADS v. 8.2.9 [32], a program for Maximum Likelihood based inference, was used
143 with a partitioned WAG+G model, where each data partition represented a single input gene
144 family. A bootstrap analysis with 100 replicates under the same model was performed in
145 RAxML in order to assess branch support of the tree. Alternatively, the PhyloBayes v3.3
146 MCMC samplers [33] was used with a CAT+GTR model and 3 chains.

147 Investigation of specific gene families

148 Transcripts involved in secondary metabolism (*PKS* and *NRPS*), sugar transport, aminoacyl
149 degradation (decarboxylases, transaminases and deaminase), fatty acid degradation
150 (alkaline and acidic lipases) and virulence (*M. circinelloides* CBS 277-49 ferroxidases *fet3*
151 and *ID112092*, *Rhizopus delmar fob*, *FTR1* and *coth*) were identified in each transcriptome.
152 *PKs* and *NRPs* transcripts were identified using the genes annotated in the *Mucor*

153 *circinelloides* CBS277.49 genome available at the JGI. Sugar transporters, *cotH* transcripts
154 and *FTR1* transcripts were identified using the corresponding domain profiles: respectively
155 Pfam ID PF00083, PF08757 and PF03239. Alkaline and acidic lipases and sequences of
156 genes involved in virulence were collected from the NCBI database and used to search the
157 *Mucor* predicted proteomes using diamond BLASTp [34]. Matching predicted proteins were
158 then used as queries to search *Mucor* predicted proteomes to identify predicted proteins that
159 might have significant sequence variation compared to the reference sequence. Annotation
160 of each matching sequence was then checked using the conserved domain search tool
161 available on NCBI [35].

162 A gene tree of putative sugar transporters was reconstructed using similar methodology as
163 presented above. *Aspergillus niger* putative sugar transporters and 61 experimentally
164 characterized fungal sugar transporters presented in Peng et al [36] were appended to the
165 analysis. Gene tree was displayed using iTOL v3 [37].

166 Availability of supporting data

167 Sequence data are available in the ArrayExpress database at EMBL-EBI
168 (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-6453. Transcriptomes are
169 available upon request. Internal script allowing the analysis of GO terms can be found on
170 github (https://github.com/anlebreton/GO_terms_FisherTest.git).

171 Results

172 RNA sequencing, transcriptome assembly and annotation

173 Sequencing of the five *Mucor* transcriptomes resulted in 25 to 35 million pairs of 101 base
174 paired-end reads (Table 2). Reads were assembled into 16,950 to 21,556 filtered transcripts

175 grouped into 13,655 to 15,554 Trinity “genes”. Except for ML, assemblies correctly
176 represented the read sets. Indeed, the percentage of realigned reads was above 93% for all
177 species except ML which was 70% (73% for the transcriptome including all isoforms). Except
178 for ME, the average transcript length was close to 1,200 bp for all species. In addition, more
179 than 97% of single-copy orthologs were complete reflecting the high-quality of the
180 assemblies. The assembly of ME was of lower quality as shown by the N50 of 1,292bp, a
181 higher number of predicted genes and the identification of single-copy orthologs against the
182 BUSCO eukaryotic database of which 67% were complete, 20% fragmented and 13%
183 missing. Despite this difference, functional annotation was similar among species: the
184 number of transcripts with protein prediction varied from 9,383 to 10,808 and among these
185 predicted proteins 61% to 72% had GO term annotation and 19% to 24% had an EC number
186 assignment (Table 2).

187 Functional comparisons

188 Overall, 1212 different EC numbers were assigned, among them 956 (79%) were shared
189 among all strains (Figure 1A). We did not identify any strains-specific pathways (data not
190 shown). Even though a significant proportion of EC were lacking for ME (7%) compared to
191 the other strains, all pathways were complete and present in this strain.

192 Nineteen GO categories were differently represented among strains (level 2) (Fisher exact,
193 p -value < 0.05). These categories corresponded to primary catabolism and anabolism (e.g.
194 “nitrogen compound metabolic process”), transport (e.g. “carbohydrate transporter activity”)
195 as well as secondary metabolism (e.g. “pigmentation”) (Figure 1B).

196 Ortholog groups

197 The predicted proteins were clustered in ortholog groups (orthogroups).
198 The core transcriptome of the five *Mucor* species grown on PDA medium comprised 5566
199 orthogroups, whereas 5017 predicted proteins could not be grouped (singletons) (Figure
200 2A).

201 Based on the obtained single copy orthologs, a phylogenomic tree was reconstructed. This
202 tree was concordant with the previously published results [38] except for the placement of
203 ME which was found basal compared to the other species in the present study (Figure 2B).

204 Some orthogroups were composed of multiple predicted proteins associated to the same
205 species. The 21 orthogroups consisting of more than 10 predicted proteins associated with
206 one species were investigated. Seven orthogroups corresponded to unknown proteins while
207 14 could be assigned to a putative function (Figure 2C). Interestingly, two orthogroups were
208 composed of predicted proteins corresponding to transposons (OG03 and OG12). In the first
209 orthogroup MF and ML had a higher number of predicted proteins by fourfold compared to
210 other species whereas in the second group MF had a higher number of predicted proteins by
211 at least twofold. In addition, the GO category transposase activity was overrepresented in
212 MR and MC compared to the other studied strains. Notably, the number of predicted proteins
213 was found higher by at least twofold in MR and MC than in the other strains in an orthogroup
214 associated with a multidrug resistance gene family. Among the singletons, predicted proteins
215 with a predicted signal peptide were specifically investigated in order to identify putative
216 secreted proteins involved in strain-specific substrate exploitation. However, no predicted
217 proteins with signal peptide were identified in MF while only a few sets could be identified in
218 the species representative strains (e.g. carbohydrate esterase and glycoside hydrolase in
219 MC) (Figure 2D). At least 48% of the singletons had EC numbers assignment and/or GO
220 term annotation. Although most of the EC numbers were strains-specific, and mapping onto
221 metabolic networks did not reveal any strain-specific pathways.

222 Investigation of specific gene families

223 The expression on PDA medium of (i) different gene families that could be involved in
224 exploitation of the cheese substrate, (ii) virulence factors that could play an important role for
225 *Mucor* opportunistic pathogens (iii) sugar transporters that may vary depending on the

226 fungus lifestyle, and (iv) PKS and NRPS that are important for secondary metabolite
227 production were analyzed in more detail.

228 MF and ML, two technological strains used for cheese ripening, did not harbor, cultivated on
229 PDA medium, a richer repertoire of transcripts corresponding to genes potentially involved in
230 the exploitation of the cheese substrate such as lactate permease and dehydrogenase and
231 in aroma production through lipolysis and aminoacyl degradation such as lipases,
232 decarboxylases and aminases. Indeed, unexpectedly more of these genes were transcribed
233 in MC and MR.

234 Among the virulence factors identified in this study, three categories harbour different
235 number of transcripts among strains: spore coat protein homologs (*coth*), high affinity iron
236 permease *FTR1* and iron transport multicopper ferroxidase *fet3* homologs.

237 The number of *coth* transcripts detected was two time higher in MC and MR than in ML and
238 MF transcriptomes (Table 3). A previously described motif involved in the invasin function of
239 *coth* was identified in three *Rhizopus delmar* *coth* genes: *coth2*, *coth3* and the *coth*
240 *RO3G_15938*. One homolog to *coth2/coth3* was found in each *Mucor* transcriptome and
241 one homolog of *RO3G_15938* was found in MC, MR and ME.

242 MC and MR clearly showed a higher number of transcripts identified as sugar transporters
243 (STs) than the other strains. However, only half of these transcripts were clustered with
244 experimentally characterized fungal transporters (Figure 3). Three clusters of transcripts
245 were expanded, all three in MR and MC compared to the other *Mucor* strains: Unknown STs
246 group 3, D_galacturonic, quinic acid STs and Glucose, pentose STs.

247 Among secondary metabolism transcripts, two fragments of non-ribosomal peptides (*NRPs*)
248 were identified in ME and two fragments of *PKs/FAS* were found in MC and MR. In both
249 cases when fragments were concatenated, they formed a complete sequence. If we

250 considered them as one gene, one NRPS and one *PKS/FAS* was found in all
251 transcriptomes.

252 Discussion

253 The transcriptomes of five *Mucor* strains pertaining to five species within a broad
254 phylogenetic range and reported to have different lifestyles were investigated in order to
255 shed new light on *Mucor* gene expression on a standard medium and get information on
256 specificities that could be linked to specific lifestyles within the *Mucor* genus. Indeed, *Mucor*
257 *lanceolatus* (ML) and *Mucor fuscus* (MF) that form a clade within the tree reconstructed
258 using 1289 orthogroups composed of single copy orthologs, were encountered in cheeses
259 [39] and were reported to have an optimal growth on cheese [4] whereas *Mucor*
260 *endophyticus* (ME), basal to the other species was only described as a wheat leaf endophyte
261 [8]. *Mucor circinelloides* (MC) and *Mucor racemosus* (MR) forming a sister clade to the
262 ML/MF clade are possibly more ubiquitous, MR being known as a recurrent cheese spoiler
263 and MC as an opportunistic pathogen [9]. It is however worth to note that the two latter
264 species include different forms which might display different ecological behaviors.

265

266 Some common traits were observed among the five transcriptomes. Global comparison of
267 transcriptome functional annotations in these five species did not reveal noticeable
268 differences among them in terms of enzyme composition nor putative secreted proteins in
269 the medium suggesting that the sets of enzymes mobilized by *Mucor* species to grow on
270 PDA medium were similar among the five species with no lost/alterd nor gained pathways
271 despite the different lifestyles and habitats reported for these species. Besides functions
272 related to basal metabolism we noticed interesting traits among the core transcriptome. It is
273 noteworthy that repeated elements seemed still active in *Mucor* genomes as shown by the
274 orthogroups OG03 and OG12 that were identified as transposons and expressed in the five

275 *Mucor* species, as well as GO term transposase activity that was differentially represented
276 among species with more expressed genes related to this GO term found in MC which is a
277 thermophilic opportunistic pathogen [38]. It is also worth to note that GO categories related
278 to known factors associated to interactions with other organisms were found: "movement in
279 the environment of other organism involved in symbiotic interaction", "multi-organism
280 metabolic process", "host cell part and other organism membrane". This raises the possibility
281 that bacterial endosymbionts occur in these studied *Mucor* strains especially in the MF and
282 ML strains where the above categories were overrepresented compared to the other strains
283 studied. This has previously been shown in *Rhizopus microsporus*, another Mucorales,
284 which harbors endosymbiotic bacteria [40, 41].

285

286 Since secondary metabolites can provide a significant advantage for survival in a given
287 ecological niche [42] and might vary depending on the fungus lifestyle, our study
288 investigated transcripts corresponding to *PKS* and *NRPS*. Indeed, *PKs* and non-ribosomal
289 peptides (*NRPs*) are involved in the majority of secondary metabolite biosynthesis (usually
290 40% and 15% respectively). Despite the different lifestyles reported for the five species
291 studied here, no differences were found among the transcriptomes obtained from cultures on
292 PDA medium. One transcript of *NRPS* was systematically found as well as one transcript
293 corresponding to a gene annotated as a *PKS I* in the *M. circinelloides* genome [11] but
294 which, as noticed by Voigt *et al.* [3], has the typical structure and domain order of a Fatty
295 Acid Synthase (*FAS*) alpha subunit. *FAS* and *PKS I* share a common evolutionary history
296 [43] and many homologies/similarities, such as the chemistry of catalyzed reactions (a
297 sequence of simple unit condensation resulting in the synthesis of a molecule of higher
298 molecular weight), or their enzymatic activity characteristics (condensation and modification
299 of fatty acids or polyketides). Further studies are needed to determine to which production
300 the *NRPS* could be associated. *Mucor* species are known to produce secondary metabolites
301 such as pigments and terpenoids [3] and has been sometimes suspected to produce harmful

302 toxins [44] but little information has already been obtained concerning production of other
303 secondary metabolites in the *Mucor* genus.

304

305 This study did not demonstrate specificities linked to species encountered in cheese and that
306 can be considered as technological species (ML and MF).

307 The more ubiquitous species MC and MR, which have been isolated from clinical
308 environments [13] especially in the case of MC which is a thermotolerant species [9]
309 involved in mucormycoses [9, 45] had transcriptomes with an over-representation of
310 "intrinsic component of membrane", "carbohydrate transporter activity", "xenobiotic
311 transporter activity", "oxidoreductase activity" and "transcription factor activity" on the
312 standard PDA medium. Among carbohydrate transporters, three sugar transporter families
313 were expanded in MR and MC transcriptomes, corresponding to D-galacturonic and quinic
314 acid STs, glucose and pentose STs and an unknown STs group 3. D-Galacturonic acid is
315 the main component of pectin which is an important plant cell-wall polysaccharide. Since
316 pectin is most abundant in the primary cell walls of soft and growing tissues, fruits and
317 vegetables are particularly pectin-rich [46]. Quinic acid can also be extracted from plants
318 sources. Possessing more diverse sugar transporters of this type may be an asset to a plant
319 pathogenic lifestyle. The unknown STs group 3 was rare in ME (the endophyte species)
320 transcriptome while it was expanded in MR and MC. Determining the substrate(s) of this
321 transporter family might contribute to our understanding of how MR and MC are more
322 ubiquitous. The GO category xenobiotic transporter activity was over-represented in MR and
323 MC. Furthermore, the orthogroup OG27 associated to a multidrug resistance gene family
324 was also mainly composed of MC and MR transcripts. These genes might contribute to a
325 better resistance to xenobiotics, including drugs, which could facilitate opportunistic
326 infections of these two species which are known to be animal/human pathogens [5, 38] and

327 might contribute to the known problem with drug treatments against mucormycosis since
328 pathogenic *Mucor* species are resistant to many classical antifungal products [47-50].

329 Among the strains studied in this work, the two strains reported as potential opportunistic
330 human pathogens presented expression of all virulence factors checked whereas at least
331 one of them was not detected in the technological and the endophytic transcriptomes.

332 Indeed, the MC ferroxidases *fet3b* and *fet3c* were absent from ME transcriptome. These
333 genes along with *fet3a*, are overexpressed during infection in a mouse model for
334 mucormycosis [51]. *fet3a* is specifically expressed during yeast growth under anaerobic
335 conditions, whereas *fet3b* and *fet3c* are specifically expressed in mycelium during aerobic
336 growth, *fet3c* being required for virulence during *in vivo* infections. *FTR1*, another gene
337 involved in iron uptake and linked to virulence showed a different distribution among *Mucor*
338 species. Two *FTR1* transcripts were detected in ME where only one was found in other
339 species, ME display a higher pathogenic susceptibility concerning this aspect. In *R. oryzae*
340 the reduction of *FTR1* copy number by gene disruption reduces the virulence of the fungus
341 in animal models of mucormycosis [52, 53]. The most important differences regarding the
342 virulence factor transcripts were observed for spore coat protein homologs (*cotH*). During
343 Human cell invasion by Mucorales, *cotH* genes allow the fungi to bind to glucose-regulated
344 protein 78 (*GRP78*) which act as endothelial cell receptor [54], the *cotH* gene copy number
345 of the species being correlated with its clinical prevalence [13]. Our results concur with this
346 hypothesis as the number of *cotH* transcripts detected was two time higher in the
347 transcriptomes of potentially pathogenic strains than in cheese technological strains. It is
348 worth noting that *cotH* genes does not have the same impact on virulence. Indeed, *Rhizopus*
349 *delmar cotH3* has higher affinity to *GRP78* than *cotH1* leading to a reduce impact of *cotH1* in
350 virulence [55]. A motif corresponding to a surface-exposed region against which a
351 therapeutic antibody has been raised was previously proposed [13]. Searching for *cotH*
352 transcripts containing this motif allowed the identification of *cotH2* and *cotH3* known to be
353 important for virulence [55] and the *cotH RO3G_15938* in *Rhizopus delmar* genes.

354 According to this study the duplication of the ancestral gene leading to *cotH2* and *cotH3*
355 happened after the separation of *Rhizopus* and *Mucor* clades. Each of the *Mucor* species
356 used in this study expressed one ortholog of *cotH2/cotH3* gene. However, the ortholog of
357 *cotH RO3G_15938*, that might be an asset for a pathogenic lifestyle, lacked in
358 transcriptomes of MF/ML, the species used for cheese ripening.

359 Conclusion

360 The transcripts obtained from five different *Mucor* spp. cultivated on PDA allowed us to
361 describe the predicted core proteome of a representative set of *Mucor* species with
362 contrasting lifestyles. This analysis provided insight into *Mucor* characteristics by highlighting
363 the presence of *NRPS* which imply a potential of *Mucor* for the production of secondary
364 metabolites including pigments, siderophores, toxins [56]. It also provided hints as to how
365 *Mucor* may adapt to different lifestyles, for example through expression of a larger set of
366 sugar transporters and a comprehensive array of virulence factors in species that inhabit in
367 multiple environments. On the other hand, species that are associated with cheese did not
368 appear to have over-representation of set of transcripts involved in cheese media usage.
369 Further studies using media mimicking cheese and animal and plant media might highlight
370 more differences in transcripts expression associated to *Mucor* adaptation to a lifestyle.

371

372 ACKNOWLEDGMENTS

373 The authors are thankful to Dr Antoine Hermet for his involvement in the early steps of this
374 project. We also thank Nantes BiogenOuest “Nantes Genomics Core Facility” for
375 technological core facilities. This research was funded by the the Région Bretagne (ARED
376 program) and EQUASA, a technological platform of the Université de Bretagne Occidentale,

377 in the framework of the MUCORSCOPE project. The authors thank the Roscoff Bioinformatic
378 platform: ABiMS (<http://abims.sb-roscoff.fr>) for providing computational resources. The
379 authors are also grateful to the reviewers for their helpful comments.

380

381 LITERATURE CITED

- 382 [1] J.W. Spatafora, Y. Chang, G.L. Benny, K. Lazarus, M.E. Smith, M.L. Berbee, G. Bonito,
383 N. Corradi, I. Grigoriev, A. Gryganskyi, T.Y. James, K. O'Donnell, R.W. Roberson, T.N.
384 Taylor, J. Uehling, R. Vilgalys, M.M. White, J.E. Stajich, A phylum-level phylogenetic
385 classification of zygomycete fungi based on genome-scale data, *Mycologia*, 108 (2016)
386 1028-1046.
- 387 [2] S. Morin-Sardin, J.L. Jany, S. Artigaud, V. Pichereau, B. Bernay, E. Coton, S. Madec,
388 Proteomic analysis of the adaptative response of *Mucor* spp. to cheese environment, *J*
389 *Proteomics*, 154 (2017) 30-39.
- 390 [3] K. Voigt, T. Wolf, K. Ochsenreiter, G. Nagy, K. Kaerger, E. Shelest, T. Papp, 15 Genetic
391 and Metabolic Aspects of Primary and Secondary Metabolism of the Zygomycetes, in: D.
392 Hoffmeister (Ed.), Springer International Publishing, Cham, 2016, pp. 361-385.
- 393 [4] S. Morin-Sardin, K. Rigalma, L. Coroller, J.L. Jany, E. Coton, Effect of temperature, pH,
394 and water activity on *Mucor* spp. growth on synthetic medium, cheese analog and cheese,
395 *Food Microbiol*, 56 (2016) 69-79.
- 396 [5] K. Hoffmann, J. Pawłowska, G. Walther, M. Wrzosek, G.S. de Hoog, G.L. Benny, P.M.
397 Kirk, K. Voigt, The family structure of the <l>Mucorales</l>: a synoptic revision based on
398 comprehensive multigene-genealogies, *Persoonia - Molecular Phylogeny and Evolution of*
399 *Fungi*, 30 (2013) 57-76.
- 400 [6] G. Walther, J. Pawłowska, A. Alastruey-Izquierdo, M. Wrzosek, J.L. Rodriguez-Tudela, S.
401 Dolatabadi, A. Chakrabarti, G.S. de Hoog, DNA barcoding in Mucorales: an inventory of
402 biodiversity, *Persoonia*, 30 (2013) 11-47.
- 403 [7] L. Mendoza, R. Vilela, K. Voelz, A.S. Ibrahim, K. Voigt, S.C. Lee, Human Fungal
404 Pathogens of Mucorales and Entomophthorales, in, 2015.
- 405 [8] R. Zheng, H. Jiang, *Rhizomucor endophyticus* sp.nov., an endophytic zygomycetes from
406 higher plants, *Mycotaxon*, 56 (1995) 455-466.
- 407 [9] S. Morin-Sardin, P. Nodet, E. Coton, J.-L. Jany, *Mucor*: A Janus-faced fungal genus with
408 human health impact and industrial applications, *Fungal Biology Reviews*, 31 (2017) 12-32.
- 409 [10] J.W. Spatafora, M.C. Aime, I.V. Grigoriev, F. Martin, J.E. Stajich, M. Blackwell, The
410 Fungal Tree of Life: from Molecular Systematics to Genome-Scale Phylogenies,
411 *Microbiology spectrum*, 5 (2017).

- 421 [11] L.M. Corrochano, A. Kuo, M. Marcet-Houben, S. Polaino, A. Salamov, J.M. Villalobos-
422 Escobedo, J. Grimwood, M.I. Álvarez, J. Avalos, D. Bauer, E.P. Benito, I. Benoit, G. Burger,
423 L.P. Camino, D. Cánovas, E. Cerdá-Olmedo, J.-F. Cheng, A. Domínguez, M. Eliáš, A.P.
424 Eslava, F. Glaser, G. Gutiérrez, J. Heitman, B. Henrissat, E.A. Iturriaga, B.F. Lang, J.L.
425 Lavín, S.C. Lee, W. Li, E. Lindquist, S. López-García, E.M. Luque, A.T. Marcos, J. Martin, K.
426 McCluskey, H.R. Medina, A. Miralles-Durán, A. Miyazaki, E. Muñoz-Torres, J.A. Oguiza,
427 R.A. Ohm, M. Orejas, L. Ortiz-Castellanos, A.G. Pisabarro, J. Rodríguez-Romero, J. Ruiz-
428 Herrera, R. Ruiz-Vázquez, C. Sanz, W. Schackwitz, M. Shahriari, E. Shelest, F. Silva-
429 Franco, D. Soanes, K. Syed, V.G. Tagua, N.J. Talbot, M.R. Thon, H. Tice, R.P. de Vries, A.
430 Wiebenga, J.S. Yadav, E.L. Braun, S.E. Baker, V. Garre, J. Schmutz, B.A. Horwitz, S.
431 Torres-Martínez, A. Idnurm, A. Herrera-Estrella, T. Gabaldón, I.V. Grigoriev, Expansion of
432 signal transduction pathways in fungi by extensive genome duplication, *Current biology* : CB,
433 26 (2016) 1577-1584.
434
- 435 [12] X. Tang, H. Chen, Y.Q. Chen, W. Chen, V. Garre, Y. Song, C. Ratledge, Comparison of
436 Biochemical Activities between High and Low Lipid-Producing Strains of *Mucor*
437 *circinelloides*: An Explanation for the High Oleaginity of Strain WJ11, *PloS one*, 10 (2015)
438 e0128396-e0128396.
439
- 440 [13] M.C. Chibucos, S. Soliman, T. Gebremariam, H. Lee, S. Daugherty, J. Orvis, A.C.
441 Shetty, J. Crabtree, T.H. Hazen, K.A. Etienne, P. Kumari, T.D. O'Connor, D.A. Rasko, S.G.
442 Filler, C.M. Fraser, S.R. Lockhart, C.D. Skory, A.S. Ibrahim, V.M. Bruno, An integrated
443 genomic and transcriptomic survey of mucormycosis-causing fungi, *Nature Communications*,
444 7 (2016) 1-11.
445
- 446 [14] S. Andrews, FASTQC: A quality control tool for high throughput sequence data, in,
447 2010.
448
- 449 [15] M.G. Grabherr, B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis,
450 L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind,
451 F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Trinity:
452 reconstructing a full-length transcriptome without a genome from RNA-Seq data, *Nature*
453 *biotechnology*, 29 (2011) 644-652.
454
- 455 [16] B. Li, C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or
456 without a reference genome, *BMC Bioinformatics*, 12 (2011) 323.
457
- 458 [17] K. Lagesen, P. Hallin, E.A. Rodland, H.H. Staerfeldt, T. Rognes, D.W. Ussery,
459 RNAmmer: consistent and rapid annotation of ribosomal RNA genes, *Nucleic Acids Res*, 35
460 (2007) 3100-3108.
461
- 462 [18] R.D. Finn, J. Clements, S.R. Eddy, HMMER web server: interactive sequence similarity
463 searching, *Nucleic Acids Res*, 39 (2011) W29-37.
464
- 465 [19] R.D. Finn, J. Tate, J. Mistry, P.C. Coggill, S.J. Sammut, H.R. Hotz, G. Ceric, K.
466 Forslund, S.R. Eddy, E.L. Sonnhammer, A. Bateman, The Pfam protein families database,
467 *Nucleic Acids Res*, 36 (2008) D281-288.
468
- 469 [20] a. Krogh, B. Larsson, G. von Heijne, E. Sonnhammer, Predicting transmembrane
470 protein topology with a hidden Markov model: application to complete genomes, *Journal of*
471 *molecular biology*, 305 (2001) 567-580.
472
- 473 [21] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal
474 peptides from transmembrane regions, *Nature Methods*, 8 (2011) 785.

- 475 [22] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search
476 tool, *J Mol Biol*, 215 (1990) 403-410.
477
- 478 [23] L.J. Ma, A.S. Ibrahim, C. Skory, M.G. Grabherr, G. Burger, M. Butler, M. Elias, A.
479 Idnurm, B.F. Lang, T. Sone, A. Abe, S.E. Calvo, L.M. Corrochano, R. Engels, J. Fu, W.
480 Hansberg, J.M. Kim, C.D. Kodira, M.J. Koehrsen, B. Liu, D. Miranda-Saavedra, S. O'Leary,
481 L. Ortiz-Castellanos, R. Poulter, J. Rodriguez-Romero, J. Ruiz-Herrera, Y.Q. Shen, Q. Zeng,
482 J. Galagan, B.W. Birren, C.A. Cuomo, B.L. Wickes, Genomic analysis of the basal lineage
483 fungus *Rhizopus oryzae* reveals a whole-genome duplication, *PLoS Genet*, 5 (2009)
484 e1000549.
485
- 486 [24] M.A. Harris, Developing an ontology, *Methods Mol Biol*, 452 (2008) 111-124.
487
- 488 [25] A.G. McDonald, K.F. Tipton, Fifty-five years of enzyme classification: advances and
489 difficulties, *FEBS Journal*, 281 (2014) 583-592.
490
- 491 [26] C. Claudel-Renard, C. Chevalet, T. Faraut, D. Kahn, Enzyme-specific profiles for
492 genome annotation: PRIAM, *Nucleic Acids Res*, 31 (2003) 6633-6639.
493
- 494 [27] F.A. Simao, R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO:
495 assessing genome assembly and annotation completeness with single-copy orthologs,
496 *Bioinformatics*, 31 (2015) 3210-3212.
497
- 498 [28] T. Yamada, I. Letunic, S. Okuda, M. Kanehisa, P. Bork, iPath2.0: interactive pathway
499 explorer, *Nucleic Acids Res*, 39 (2011) W412-415.
500
- 501 [29] D.M. Emms, S. Kelly, OrthoFinder: solving fundamental biases in whole genome
502 comparisons dramatically improves orthogroup inference accuracy, *Genome Biology*, 16
503 (2015) 157-157.
504
- 505 [30] A. Loytynoja, Phylogeny-aware alignment with PRANK, *Methods Mol Biol*, 1079 (2014)
506 155-170.
507
- 508 [31] S. Capella-Gutiérrez, J.M. Silla-Martínez, T. Gabaldón, trimAl: a tool for automated
509 alignment trimming in large-scale phylogenetic analyses, *Bioinformatics*, 25 (2009) 1972-
510 1973.
511
- 512 [32] A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of
513 large phylogenies, *Bioinformatics*, 30 (2014) 1312-1313.
514
- 515 [33] N. Lartillot, T. Lepage, S. Blanquart, PhyloBayes 3: a Bayesian software package for
516 phylogenetic reconstruction and molecular dating, *Bioinformatics*, 25 (2009) 2286-2288.
517
- 518 [34] B. Buchfink, C. Xie, D.H. Huson, Fast and sensitive protein alignment using DIAMOND,
519 *Nature Methods*, 12 (2014) 59.
520
- 521 [35] A. Marchler-Bauer, Y. Bo, L. Han, J. He, C.J. Lanczycki, S. Lu, F. Chitsaz, M.K.
522 Derbyshire, R.C. Geer, N.R. Gonzales, M. Gwadz, D.I. Hurwitz, F. Lu, G.H. Marchler, J.S.
523 Song, N. Thanki, Z. Wang, R.A. Yamashita, D. Zhang, C. Zheng, L.Y. Geer, S.H. Bryant,
524 CDD/SPARCLE: functional classification of proteins via subfamily domain architectures,
525 *Nucleic Acids Research*, 45 (2017) D200-D203.
526
- 527 [36] M. Peng, M.V. Aguilar-Pontes, R.P. de Vries, M.R. Mäkelä, In Silico Analysis of Putative
528 Sugar Transporter Genes in *Aspergillus niger* Using Phylogeny and Comparative
529 Transcriptomics, *Frontiers in Microbiology*, 9 (2018).

- 530 [37] I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and
531 annotation of phylogenetic and other trees, *Nucleic Acids Research*, 44 (2016) W242-W245.
532
- 533 [38] G. Walther, J. Pawłowska, A. Alastruey-Izquierdo, M. Wrzosek, J.L. Rodriguez-Tudela,
534 S. Dolatabadi, A. Chakrabarti, G.S. de Hoog, DNA barcoding in *Mucorales*: an
535 inventory of biodiversity, *Persoonia - Molecular Phylogeny and Evolution of Fungi*, 30 (2013)
536 11-47.
537
- 538 [39] A. Hermet, D. Meheust, J. Mounier, G. Barbier, J.L. Jany, Molecular systematics in the
539 genus *Mucor* with special regards to species encountered in cheese, *Fungal Biol*, 116 (2012)
540 692-705.
541
- 542 [40] S.J. Mondo, O.A. Lastovetsky, M.L. Gaspar, N.H. Schwardt, C.C. Barber, R. Riley, H.
543 Sun, I.V. Grigoriev, T.E. Pawłowska, Bacterial endosymbionts influence host sexuality and
544 reveal reproductive genes of early divergent fungi, *Nature communications*, 8 (2017) 1843-
545 1843.
546
- 547 [41] L.P. Partida-Martinez, I. Groth, I. Schmitt, W. Richter, M. Roth, C. Hertweck,
548 *Burkholderia rhizoxinica* sp. nov. and *Burkholderia endofungorum* sp. nov., bacterial
549 endosymbionts of the plant-pathogenic fungus *Rhizopus microsporus*, *International journal*
550 *of systematic and evolutionary microbiology*, 57 (2007) 2583-2590.
551
- 552 [42] E.M. Fox, B.J. Howlett, Secondary metabolism: regulation and role in fungal biology,
553 *Current opinion in microbiology*, 11 (2008) 481-487.
554
- 555 [43] H. Jenke-Kodama, A. Sandmann, R. Muller, E. Dittmann, Evolutionary implications of
556 bacterial polyketide synthases, *Molecular biology and evolution*, 22 (2005) 2027-2039.
557
- 558 [44] S.C. Lee, R.B. Billmyre, A. Li, S. Carson, S.M. Sykes, E.Y. Huh, P. Mieczkowski, D.C.
559 Ko, C.A. Cuomo, J. Heitman, Analysis of a Food-Borne Fungal Pathogen Outbreak:
560 Virulence and Genome of a *Mucor circinelloides* Isolate from Yogurt, *mBio*, 5 (2014) e01390-
561 01314.
562
- 563 [45] S.C. Lee, A. Li, S. Calo, J. Heitman, Calcineurin Plays Key Roles in the Dimorphic
564 Transition and Virulence of the Human Pathogenic Zygomycete *Mucor circinelloides*, *PLOS*
565 *Pathogens*, 9 (2013) e1003625.
566
- 567 [46] J.P. Benz, R.J. Protzko, J.M. Andrich, S. Bauer, J.E. Dueber, C.R. Somerville,
568 Identification and characterization of a galacturonic acid transporter from *Neurospora crassa*
569 and its application for *Saccharomyces cerevisiae* fermentation processes, *Biotechnology for*
570 *Biofuels*, 7 (2014) 20.
571
- 572 [47] R.E. Lewis, O. Lortholary, B. Spellberg, E. Roilides, D.P. Kontoyiannis, T.J. Walsh, How
573 Does Antifungal Pharmacology Differ for Mucormycosis Versus Aspergillosis?, *Clinical*
574 *Infectious Diseases*, 54 (2012) S67-S72.
575
- 576 [48] A. Muszewska, J. Pawłowska, P. Krzyściak, Biology, systematics, and clinical
577 manifestations of Zygomycota infections, *European Journal of Clinical Microbiology &*
578 *Infectious Diseases*, 33 (2014) 1273-1287.
579
- 580 [49] T.T. Riley, C.A. Muzny, E. Swiatlo, D.P. Legendre, Breaking the Mold:A Review of
581 Mucormycosis and Current Pharmacological Treatment Options, *Annals of*
582 *Pharmacotherapy*, 50 (2016) 747-757.
583

- 584 [50] A. Skiada, L. Pagano, A. Groll, S. Zimmerli, B. Dupont, K. Lagrou, C. Lass-Flörl, E.
585 Bouza, N. Klimko, P. Gaustad, M. Richardson, P. Hamal, M. Akova, J.F. Meis, J.L.
586 Rodriguez-Tudela, E. Roilides, A. Mitrousia-Ziouva, G. Petrikkos, Zygomycosis in Europe:
587 analysis of 230 cases accrued by the registry of the European Confederation of Medical
588 Mycology (ECMM) Working Group on Zygomycosis between 2005 and 2007, *Clinical*
589 *Microbiology and Infection*, 17 (2011) 1859-1867.
- 590
- 591 [51] M.I. Navarro-Mendoza, C. Pérez-Arques, L. Murcia, P. Martínez-García, C. Lax, M.
592 Sanchis, J. Capilla, F.E. Nicolás, V. Garre, Components of a new gene family of ferroxidases
593 involved in virulence are functionally specialized in fungal dimorphism, *Scientific Reports*, 8
594 (2018) 7660.
- 595
- 596 [52] A.S. Ibrahim, B. Spellberg, T.J. Walsh, D.P. Kontoyiannis, Pathogenesis of
597 Mucormycosis, *Clinical Infectious Diseases*, 54 (2012) S16-S22.
- 598
- 599 [53] M. Liu, L. Lin, T. Gebremariam, G. Luo, C.D. Skory, S.W. French, T.-F. Chou, J.E.
600 Edwards, Jr., A.S. Ibrahim, Fob1 and Fob2 Proteins Are Virulence Determinants of *Rhizopus*
601 *oryzae* via Facilitating Iron Uptake from Ferrioxamine, *PLOS Pathogens*, 11 (2015)
602 e1004842.
- 603
- 604 [54] M. Liu, B. Spellberg, Q.T. Phan, Y. Fu, Y. Fu, A.S. Lee, J.E. Edwards, Jr., S.G. Filler,
605 A.S. Ibrahim, The endothelial cell receptor GRP78 is required for mucormycosis
606 pathogenesis in diabetic mice, *The Journal of Clinical Investigation*, 120 (2010) 1914-1924.
- 607
- 608 [55] T. Gebremariam, M. Liu, G. Luo, V. Bruno, Q.T. Phan, A.J. Waring, J.E. Edwards, Jr.,
609 S.G. Filler, M.R. Yeaman, A.S. Ibrahim, CotH3 mediates fungal invasion of host cells during
610 mucormycosis, *The Journal of Clinical Investigation*, 124 (2014) 237-250.
- 611
- 612 [56] P.J. Bhetariya, M. Prajapati, A. Bhaduri, R.S. Mandal, A. Varma, T. Madan, Y. Singh,
613 P.U. Sarma, Phylogenetic and Structural Analysis of Polyketide Synthases in *Aspergilli*,
614 *Evolutionary Bioinformatics Online*, 12 (2016) 109-119.
- 615
- 616 [57] J. CERNING, J. GRIPON, C., G. LAMBERET, J. LENOIR, Les activités biochimiques
617 des *Penicillium* utilisés en fromagerie, *Lait*, 67 (1987) 3-39.
- 618
- 619 [58] P.L.H. McSweeney, Chapter 14 - Biochemistry of Cheese Ripening: Introduction and
620 Overview, in: P.L.H. McSweeney, P.F. Fox, P.D. Cotter, D.W. Everett (Eds.) *Cheese* (Fourth
621 Edition), Academic Press, San Diego, 2017, pp. 379-387.
- 622
- 623 [59] P.L.H. McSweeney, P.F. Fox, F. Ciocia, Chapter 16 - Metabolism of Residual Lactose
624 and of Lactate and Citrate, in: P.L.H. McSweeney, P.F. Fox, P.D. Cotter, D.W. Everett (Eds.)
625 *Cheese* (Fourth Edition), Academic Press, San Diego, 2017, pp. 411-421.
- 626
- 627 [60] L. López-Fernández, M. Sanchis, P. Navarro-Rodríguez, F.E. Nicolás, F. Silva-Franco,
628 J. Guarro, V. Garre, M.I. Navarro-Mendoza, C. Pérez-Arques, J. Capilla, Understanding
629 *Mucor circinelloides* pathogenesis by comparative genomics and phenotypical studies,
630 *Virulence*, 9 (2018) 707-720.

631

632

633 **Table 1:** List of *Mucor* strains used in the present study, their origin and reported habitats for
 634 the corresponding species according to Walther *et al.* (2013) [38]; Hermet *et al.* (2012) [39];
 635 Zheng and Jiang (1995) [8].

Species	Strain	Strain origin	Reported habitat	Reported role
<i>M. racemosus</i>	UBOCC-A-109155	Cheese	Cheese, yogurt, walnuts, sausages, grassland soil, decaying vegetables, human	Food contaminant, technological in cheese production, pathogen
<i>M. fuscus</i>	UBOCC - A-109160	Cheese	Cheese, dung, sediment	Technological in cheese production
<i>M. lanceolatus</i>	UBOCC-A-109153	Cheese	Cheese	Technological in cheese production
<i>M. endophyticus</i>	CBS 385-95	<i>Triticum aestivum</i> ; leaves;	<i>Triticum aestivum</i> endophyte	None
<i>M. circinelloides</i>	CBS 277-49	Unknown	Sufu, corn grain, fungi (basidiomycota), human, forest soil, decaying vegetables	Food contaminant, technological in sufu production, pathogen

636

637

638 **Table 2:** Overview statistics of *Mucor racemosus* (MR), *Mucor fuscus* (MF), *Mucor*
639 *lanceolatus* (ML), *Mucor endophyticus* (ME) and *Mucor circinelloides* (MC) including
640 transcriptome size, assembly quality and annotation. Gene* refers to cluster of transcripts
641 predicted to be transcribed from the same gene by the assembler Trinity.

species	MR	MF	ML	ME	MC
Filtered reads (2×101 bp)	30,203,513	25,511,572	24,835,844	34,852,641	24,642,543
No. genes*	14,035	14,299	14,041	15,554	13,655
No. transcripts	17,368	20,898	21,556	16,950	19,891
Average transcript length	1,205	1,231	1,209	836	1,295
Remapped reads (%)	93.08	95.33	70.93	96.94	93.60
Complete BUSCO genes (%)	99,34	97,36	97,36	67,00	97,69
no. of predicted proteins coding genes*	10808	9963	9383	10434	10194
No. predicted proteins with EC numbers assignment	3134	2886	2801	2727	3085
No. predicted proteins with GO term annotation	10202	9021	8729	11280	9506

642

643

644 **Table 3:** Number of predicted proteins, potentially involved in adaptation to different
645 lifestyles, across the translated products of the five *Mucor* transcriptomes. *Mucor racemosus*
646 (MR), *Mucor fuscus* (MF), *Mucor lanceolatus* (ML), *Mucor endophyticus* (ME) and *Mucor*
647 *circinelloides* (MC).

Target	Gene name	References	MC	MR	MF	ML	ME
Cheese exploitation	lactate permease		2	3	2	2	2
	lactate dehydrogenase		2	2	1	1	1
	alkaline lipase		3	1	4	2	0
	acidic lipase		3	2	1	1	1
	decarboxylase	[57-59]	21	25	19	22	22
	transaminase		26	27	25	23	26
	deaminase		21	20	16	16	21
	<i>cotH2, cotH3</i>		1	1	1	1	1
	<i>RO3G_15938</i>	[13, 55]	1	1	0	0	1
	all <i>cotH</i>		16	17	7	8	11
Virulence factors	<i>ID112092 Mucor circinelloides</i>	[60]	1	1	1	1	1
	<i>fob1 (RO3G_11000), fob2 (RO3G_11000)</i>	[52, 53]	1	1	1	1	1
	<i>FTR1 (RO3G_03470)</i>		1	1	1	1	2
	<i>fet3a</i>		0	0	0	0	1
	<i>fet3b</i>	[51]	1	1	1	1	0
	<i>fet3c</i>		1	1	1	1	0
Substrat exploitation	sugar transporter		47	39	29	25	25
secondary metabolism	<i>NRPs</i>		1	1	1	1	2 fragments
	<i>PKs/FAS</i>	[3]	2 fragments	2 fragments	1	1	1

648

649

650 Legends of Figures

651 **Figure 1** : Global comparisons of the five transcriptomes functional annotation. (A) The
652 Venn diagram shows the distribution of Enzyme Commission (EC) numbers assigned across
653 the translated products of the five *Mucor* transcriptomes. Shaded sections indicate
654 categories with manual curation of the annotation. (B) Comparison of GO sub-categories
655 (level 2) annotations across the translated products of the five *Mucor* transcriptomes. (Fisher
656 exact, * p value < 0.05, ** p value < 0.005, *** p value < 0.0005). bp: biological process, cc:
657 cellular component, mf: molecular function, MR: *Mucor racemosus*, MF: *M. fuscus*, ML: *M.*
658 *lanceolatus*, ME: *M. endophyticus*, MC: *M. circinelloides*.

659

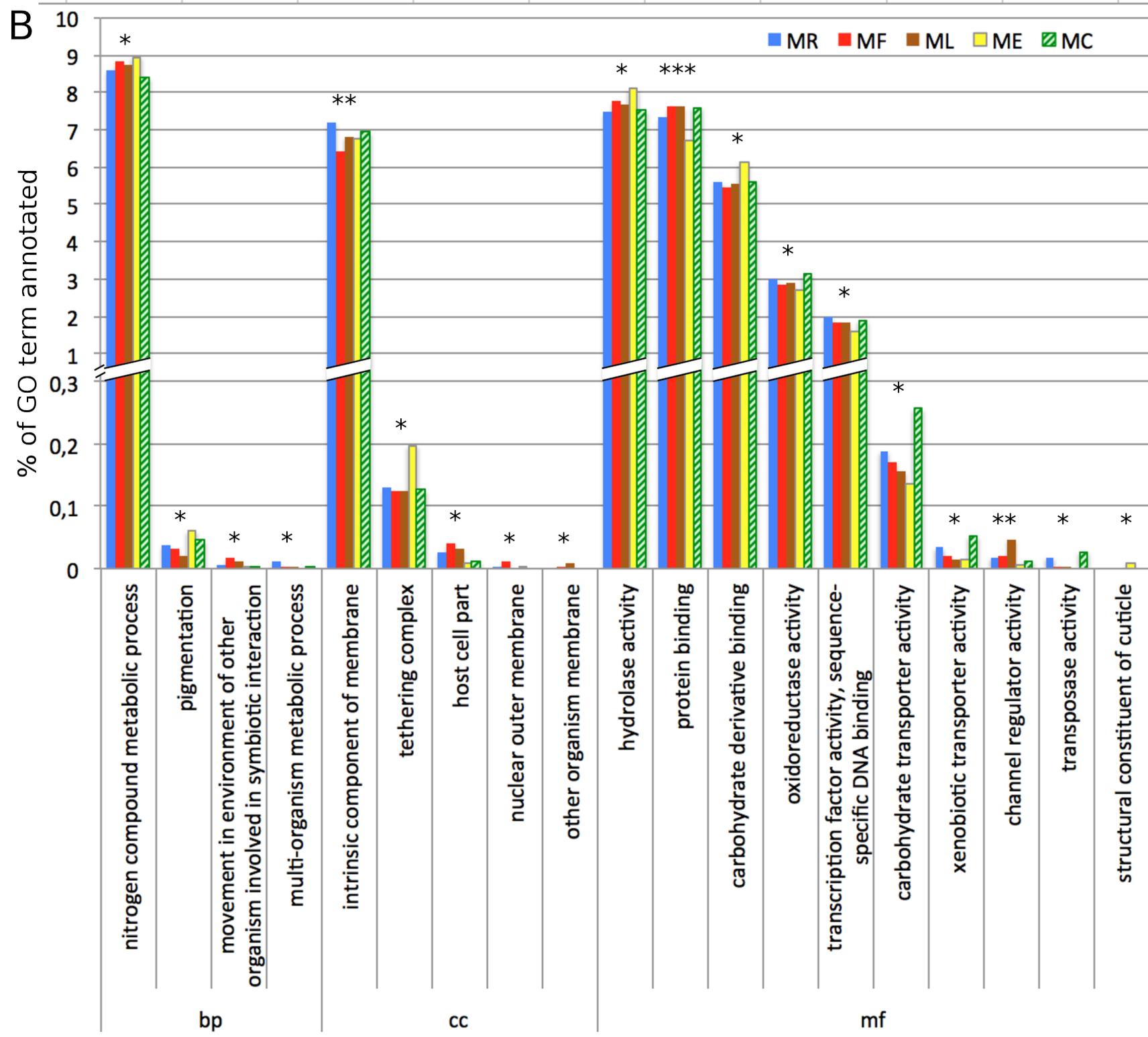
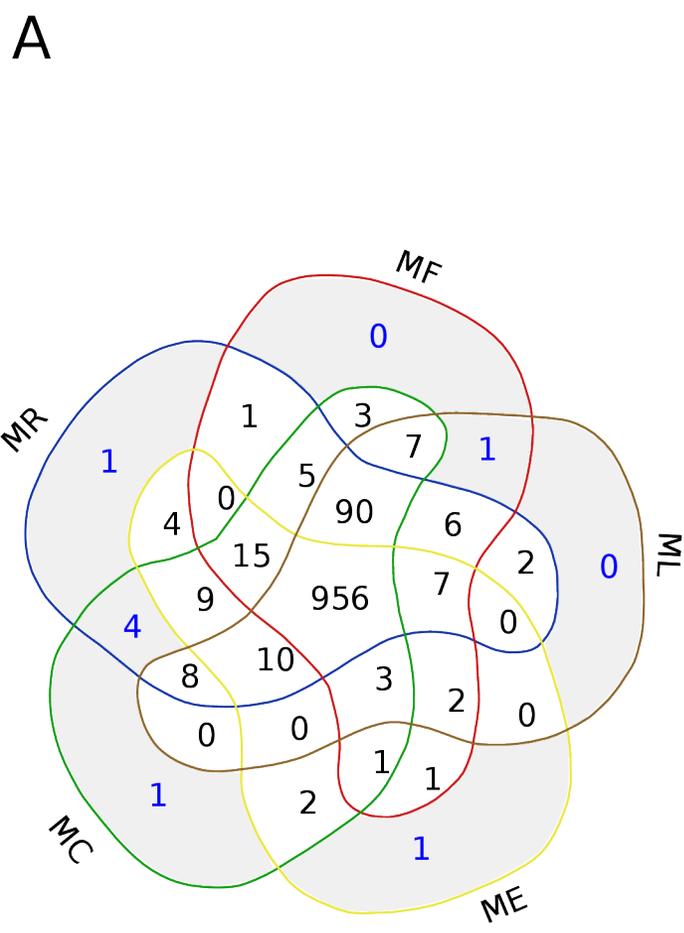
660 **Figure 2:** (A) Distribution of orthogroups among strains. Hereafter, predicted proteins that
661 could not be grouped are referred as singletons. (B) Phylogenetic species tree and reported
662 lifestyle of the five studied *Mucor* species: *M. racemosus* (MR), *M. fuscus* (MF), *M.*
663 *lanceolatus* (ML), *M. endophyticus* (ME) and □*M. circinelloides* (MC). Branch length
664 represents the substitution per site, numbers on node represent the bootstrap support. (C)
665 The heat map represents orthogroups composed of more than 10 proteins for at least one
666 strain. For each orthogroup, the number of predicted proteins of each strain is represented
667 by a color gradient going from blue (no predicted protein) to red (more than 20 predicted
668 proteins). Only orthogroups with putative function assignment are displayed. (D) The table
669 lists the function of singletons with signal peptides found in the five transcriptomes. Stars (*)
670 indicate partial predicted proteins. Predicted proteins of unknown function are not shown.

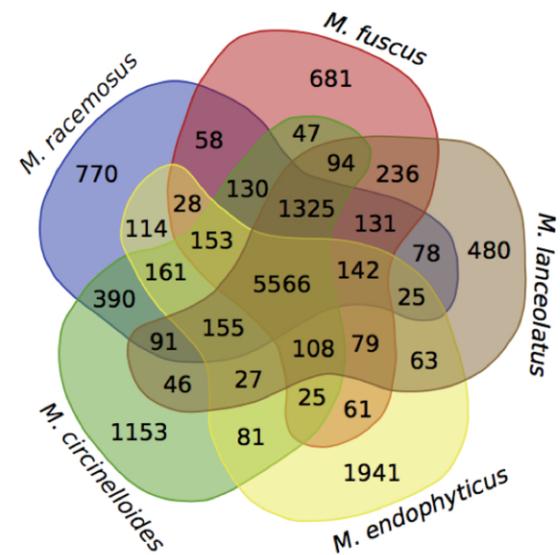
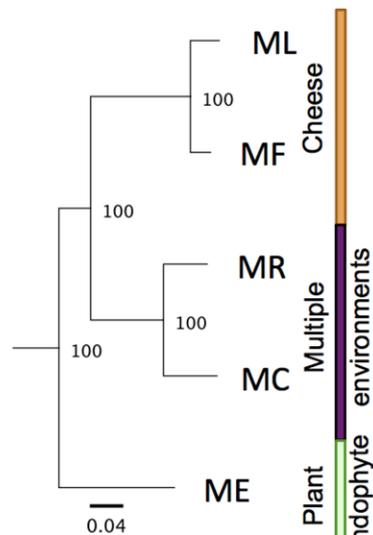
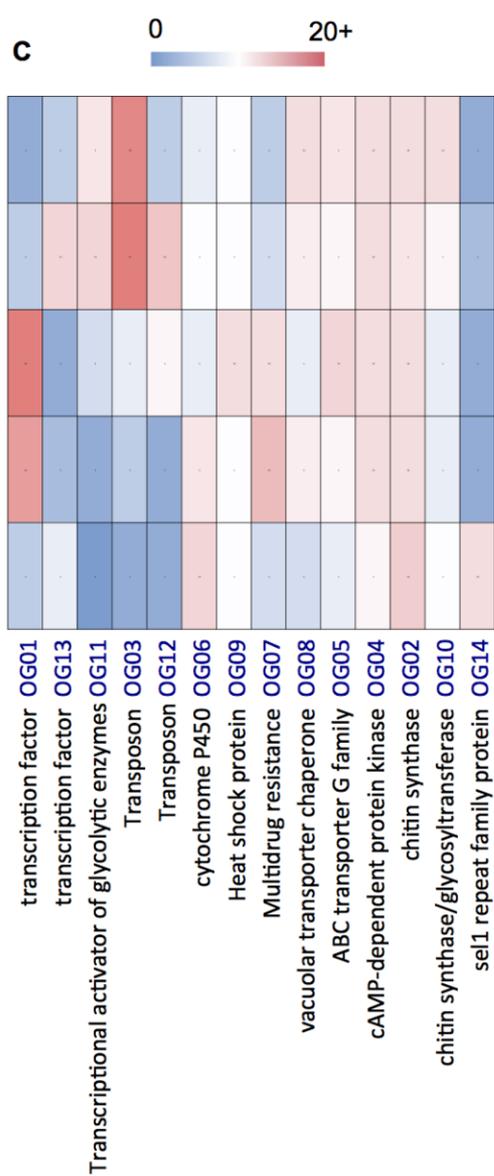
671

672 **Figure 3:** Phylogenetic classification of putative sugar transporters (STs) in five *Mucor*
673 transcriptomes. The tree includes 165 *Mucor* sugar transporters containing Pfam domain
674 PF00083, 86 *Aspergillus niger* putative sugar transporters containing Pfam domain PF00083
675 and 61 experimentally characterized fungal transporters. Predicted protein of *Mucor* species
676 found in multiple environments are highlighted with violet font, those found in the endophytic
677 species in green and technological species in orange. Four clades were collapsed as they
678 contain only references genes. Bars around the tree indicates the substrate of
679 experimentally characterized fungal transporters within the cluster. Branches with bootstrap

680 values above 50% are indicated by circles with larger circles representing higher bootstraps.
681 *Mucor* genes begin with the species name tag: MC = *Mucor circinelloides*, MR = *Mucor*
682 *racemosus*, MF = *Mucor fuscus*, ML = *Mucor lanceolatus* and ME= *Mucor endophyticus*. An
683 represent *Aspergillus niger* putative sugar transporter genes. The abbreviation of reference
684 fungal species name is attached to each known transporter gene (*Anid* = *Aspergillus*
685 *nidulans*, *Anig* = *Aspergillus niger*, *Aory* = *Aspergillus oryzae*, *Amon* = *Ambrosiozyma*
686 *monospora*, *Bcin* = *Botrytis cinerea*, *Kmar* = *Kluyveromyces marxianus*, *Klac* =
687 *Kluyveromyces lactis*, *Ncra* = *Neurospora crassa*, *Psti* = *Pichia stipitis*, *Scer* =
688 *Saccharomyces cerevisiae*, *Spas* = *Saccharomyces pastorianus*, *Stip* = *Scheffersomyces*
689 *stipitis*, *Tree* = *Trichoderma reesei*, *Thar* = *Trichoderma harzianum*, *umay* = *Ustilago*
690 *maydis*).

691



A**B****C****D**

Serine protease *
Putative endo-rhamnogalacturonase*
-
Mitochondrial NADH dehydrogenase
UDP-glucosyl transferase family protein*
Glycoside hydrolase family 5
Glycoside hydrolase family 9*
carbohydrate esterase family
Arginase
Alkaline phosphatase *

