Domestication of the emblematic white cheese-making fungus *Penicillium camemberti* and its diversification into two varieties

Jeanne Ropars[1], Estelle Didiot[1], Ricardo C. Rodríguez de la Vega[1], Bastien Bennetot[1], Monika Coton[2], Elisabeth Poirier[2], Emmanuel Coton[2], Alodie Snirc[1], Stéphanie Le Prieur[1] and Tatiana Giraud[1]

[1]Ecologie Systematique et Evolution, Universite Paris-Saclay, CNRS, AgroParisTech, 91400 Orsay, France

[2]Univ Brest, Laboratoire Universitaire de Biodiversité et Écologie Microbienne, F-29280 Plouzané, France

Corresponding author and lead contact: Jeanne Ropars, jeanne.ropars@universite-paris-saclay.fr

Keywords: domestication, fungi, cheese, *Penicillium,* mycotoxin, horizontal gene transfer

Summary

Domestication involves recent adaptation under strong human selection and rapid diversification, and therefore constitutes a good model for studies of these processes. We studied the domestication of the emblematic white mold *Penicillium camemberti*, used for the maturation of soft cheeses, such as Camembert and Brie, about which surprisingly little was known, despite its economic and cultural importance. Whole genome-based analyses of genetic relationships and diversity revealed that an ancient domestication event led to the emergence of the gray-green *P. biforme* mold used in cheese-making, by divergence from the blue-green wild *P. fuscoglaucum* fungus. Another much more recent domestication event led to the generation of the *P. camemberti* clonal lineage as a sister group to *P. biforme*. Both these domestication events were associated with strong bottlenecks. *Penicillium biforme* displayed signs of phenotypic adaptation to cheese-making relative to *P. fuscoglaucum*, in terms of whiter color, faster growth on cheese medium under cave conditions, lower levels of toxin production and greater ability to prevent the growth of other fungi. The *P. camemberti* lineage displayed even stronger signs of domestication for all these phenotypic features. We also identified two differentiated *P. camemberti* varieties, apparently associated with different kinds of cheeses, and with contrasted phenotypic features in terms of color, growth, toxin production and competitive ability. We have, thus, identified footprints of domestication in these fungi, with genetic differentiation between cheese and wild populations, and specific phenotypic traits beneficial for cheese-making. This study has not only fundamental implications for our understanding of domestication but can also have important impacts on cheese-making.

Introduction

Understanding how organisms adapt to their environment is a key issue in evolutionary biology, requiring investigations of population subdivision, and levels of genetic and phenotypic diversity or adaptive divergence. Domestication is a good model for studies of adaptive divergence, as it involves recent adaptation events affecting known traits under strong human selection and rapid diversification. Several studies on domesticated animals (e.g. horse, dog, pig [1–3]) and plants (e.g. maize, apricot [4,5]) have improved our understanding of adaptive divergence by revealing how adaptation and diversification into varieties occurred. Maize, for example, has undergone major changes in phenotype compared to its wild relative (teosinte), including a decrease in tillering, and the development of larger, non-dehiscent grains [6]. Several different maize varieties have been selected for different usages, with sugar-rich varieties grown for human consumption as kernels and field corn varieties grown for animal feed. Similarly, a number of different *Brassica* varieties have been selected for their leaves (cabbage and kale), stems (kohlrabi), flower shoots (broccoli and cauliflower) or buds (Brussels sprouts). Dog breeds have also diversified greatly under human selection, with different breeds displaying very different behaviors and phenotypes. Such notable adaptation is often associated with a decrease in fitness in natural environments, with, for example, a decrease in or loss of sexual reproduction ability in bulldogs [7] and bananas [8], smaller brains and less acute sense organs in most domesticated animals [9].

Fungi are excellent models for studying evolution and adaptation in eukaryotes, given their many experimental assets [10], including their small genomes and tractability for laboratory experiments. They can therefore be used to address questions in evolutionary biology through complementary approaches, including genomic and experimental methods. Humans have domesticated several fungi for the fermentation of foods (e.g. for beer, bread, wine, dried sausage and cheese), the production of secondary metabolites used in pharmaceutics (e.g. penicillin), and for their nutritional and gustatory qualities (e.g. button and shiitake mushrooms) [11]. Despite their economic and industrial importance, and their utility as biological models for studying adaptive divergence, fungi used by humans have been little studied, with the exception of the budding yeast *Saccharomyces cerevisiae* used for beer, wine and bread production [12–22], the filamentous fungus *Aspergillus oryzae* used to ferment soy and rice products in Asia [23–25] and the blue-cheese mold *Penicillium roqueforti* [26–28]. Whole-genome analyses have revealed that *P. roqueforti* has been domesticated twice, in two independent events, resulting in one population specific to the Roquefort protected designation of origin (PDO), the other population being used worldwide for all types of blue cheeses [26–29]. The Roquefort population displays some genetic

diversity and had beneficial traits for cheese production before the industrial era, growing more slowly in cheese and displaying greater spore production on bread, the traditional multiplication medium. By contrast, the non-Roquefort cheese population is a clonal lineage with traits beneficial for industrial cheese production, such as high levels of lipolytic activity, efficient cheese cavity colonization and high salt tolerance. Some of these beneficial traits have been conferred by two large horizontally transferred genomic regions, *Wallaby* and *CheesyTer* [27,28], in the clonal non-Roquefort cheese population. The non-Roquefort cheese *P. roqueforti* population grows more slowly in harsh conditions (minimal medium) than the Roquefort cheese population [28]. Whole-genome analyses of *S. cerevisiae* have also revealed the existence of a cheese-specific population of this yeast differentiated from the populations used for alcohol or bread production. The cheese *S. cerevisiae* population assimilates galactose more rapidly than *S. cerevisiae* populations thriving in other food environments (e.g. beer, bread, wine) or in natural environments (oak) [30].

The white mold *Penicillium camemberti* is used for the maturation of soft cheeses, such as Camembert, Brie and Neufchatel, and also to a lesser extent for dry sausage maturation. It is thought to be a white mutant selected from the gray-green species *P. commune* for its color at the start of the 20th century [31], and cultured clonally ever since. The first records of Brie cheeses date from middle ages [32], being of blue aspect until the 20th century [33,34]. Very little is known about the *P. camemberti* taxonomic status, origin and diversity, despite its great economic and cultural importance. In particular, its relationships to the closely related species *P. biforme, P. caseifulvum, P. commune* and *P. fuscoglaucum,* and even possible overlaps with these species, remain unclear. *Penicillium camemberti* has only ever been found in the food environment. Strains identified as *P. commune* are used for the maturation of other types of cheese (e.g. hard and blue cheeses) and in the production of dried sausages, and can be commonly found as spoilers of dairy products [35] and also in non-food environments. Genetic analyses however suggested that *P. commune* was not monophyletic (Figure 1A), which supported the reinstatement of two ancient species names, *P. biforme* (named by Thom in 1910) and *P. fuscoglaucum* (named by Biourge in 1923) [36]. A few genetic sequences indeed separated *P. commune* strains isolated from cheese (including the *P. biforme* type) from *P. commune* strains isolated from other environments (including the *P. fuscoglaucum* type) [36]. However, the most recent taxonomic reference study of the *Penicillium* genus did not recognize *P. fuscoglaucum* as a species and maintained the *P. commune* name [37]. *Penicillium caseifulvum* has also sometimes been advocated to constitute a separate species in this clade [37], including cheese strains isolated from Danish blue and other German and French cheeses, on the basis of colony morphology and their lack of production of cyclopiazonic acid (CPA), a mycotoxin known to be produced by *P. camemberti.* However, a study based on a small

106 number of genetic markers was unable to differentiate these putative *P. caseifulvum* strains from *P.*
107 *camemberti* [36]. Parts of the horizontally transferred *CheesyTer* and *Wallaby* regions have also been
108 found in the reference *P. camemberti* and *P. biforme* genomes [27,28].

109 The taxonomy of this clade thus remains unclear, and population structure has not been studied
110 with powerful genetic markers, despite the paramount importance of genetic relationships to our
111 understanding of domestication. Differences in phenotypic traits have not been extensively
112 studied either, but are also essential for determining whether or not domestication has occurred.
113 Humans may, indeed, have selected strains for specific traits beneficial for cheese production,
114 such as faster growth in cheese, attractive color and a greater ability to compete against food-
115 spoiling microorganisms, but without the production of detrimental extrolites, called mycotoxins.
116 Mycotoxins are natural secondary metabolites produced by fungi under certain conditions that
117 can inhibit competitor growth and reproduction and which can have a toxicological impact on
118 humans depending on the exposure conditions [38]. Cyclopiazonic acid is one of these toxic
119 secondary metabolites produced by *P. camemberti* under certain conditions [39,40].

120 Our goal here was to investigate the footprints of domestication by assessing genetic and
121 phenotypic differentiation, by sequencing the genomes of strains isolated from the rinds of
122 different cheeses from around the world (e.g. Camembert, tommes, blue cheeses), from other
123 food products (e.g. dry sausages, bread) and from other environments (e.g. leaf litter, ice and
124 wood). We performed a whole genome-based analysis of population structure, which revealed
125 that *P. biforme, P. camemberti* and *P. fuscoglaucum* formed separate and specific genetic clusters, *P.*
126 *biforme* and *P. camemberti* being sister clades (Figure 1A). The species name *P. commune* should
127 therefore be avoided, being paraphyletic. We found that the sister clades *P. camemberti* and *P.*
128 *biforme* were both specific to the cheese environment, whereas the more distantly related *P.*
129 *fuscoglaucum* clade was mostly found in natural environments. These relationships suggest an
130 ancient domestication event separating *P. biforme* from *P. fuscoglaucum* and a much more recent
131 domestication event separating the *P. camemberti* clonal lineage from *P. biforme.* Consistent with
132 this scenario, we found evidence of phenotypic adaptation to cheese-making in *P. biforme* relative
133 to *P. fuscoglaucum*, with a whiter color, faster growth on cheese medium under cave conditions and
134 lower levels of toxin production. We also reveal the diversification of *P. camemberti* into two
135 varieties, *P. camemberti* var. *camemberti* and var. *caseifulvum*, also displaying genetic and phenotypic
136 differentiation (in terms of growth rate, color, competitive abilities and mycotoxin production),
137 and maybe used for the production of different cheese types.

138

Results and Discussion

Penicillium camemberti, P. biforme and *P. fuscoglaucum* each form separate specific genetic clusters

We collected and sequenced the genomes of 61 strains with Illumina technology, including 36 strains isolated from the rinds of various types of cheese (e.g., tommes, Camembert, blue cheeses), 11 from other food products (e.g. dry sausages, bread), and 14 from non-food environments (e.g., wood or leaf litter), which were attributed by their collectors to the species *P. camemberti, P. biforme, P. commune* or *P. fuscoglaucum* (Supplemental information). We resequenced the reference genome of *P. camemberti* (LCP06093, also known as FM013, and initially sequenced with 454 technology [27]) with PacBio long-read technology and used the ca. 35Mb assembly obtained for mapping. We identified 392,072 SNPs across all strains (Supplemental Data 1). We investigated population structure with two clustering methods based on genetic differences and free from assumptions about mating system and mode of reproduction. Principal component analysis (PCA; Figure 2A) and neighbor-net (splitstree) analysis (Figure 2C) identified three genetic clusters, corresponding to *P. camemberti* (n=19), *P. biforme* (n=28) and *P. fuscoglaucum* (n=14). A population structure analysis implemented in NGSadmix also identified the same three genetic clusters at K=3 (Figure 2D), the K value at which the structure was the strongest and clearest. The genetic differentiation between *P. camemberti, P. biforme* and *P. fuscoglaucum* was further confirmed by the high values of fixation (F_{ST}) and absolute divergence (d_{xy}) indices (Table S1). A recent study described rapid phenotypic change from green-gray to white in a "wild *P. commune*" strain [41], interpreting these changes as evidence that domestication can occur within four weeks. However, we found that this *P. commune* strain actually belonged to the cheese *P. biforme* clade, being genetically almost identical to strains found in commercial cheeses (Figure S1).

The strains present in public collections under the name "*P. commune*" did not form a single cluster nor even a monophyletic clade (red shapes in Figures 2A, 2C and 3A); indeed, the *P. commune* strains isolated from natural environments clustered with *P. fuscoglaucum* strains whereas the *P. commune* strains isolated from cheese clustered with *P. biforme* strains, despite *P. fuscoglaucum* and *P. biforme* not being sister clades (Figure 3A). The *P. commune* neotype strain (LCP05531) and the *P. biforme* type strain (LCP05529) belonged to the same clade (Figures 2C-D and Figure 3A), further indicating that *P. commune* and *P. biforme* overlap. In contrast, the *P. caseifulvum* type strain

171 (LCP05630) and the *P. camemberti* type strain belonged to different clusters at *K* values ≥ 5 in the
172 population structure analysis (red stars on Figure 2C), suggesting that they do correspond to
173 different genetic clusters.

174 *Penicillium camemberti* was found only in food (cheese, sausage, food waste or animal feed; Figure
175 2). The three non-cheese strains were not differentiated from cheese strains. Only three of the 28
176 strains in the *P. biforme* cluster were isolated from environments other than food: the atmosphere,
177 ice and leaf litter. These three strains were not genetically differentiated from those isolated from
178 food (bread, cheese or sausage) and may therefore be feral strains, i.e. escaped from food. The *P.*
179 *fuscoglaucum* cluster included three strains isolated from food environments (two strains from
180 cheese and one from sausage) and 13 strains from natural environments, and was thus likely to
181 constitute a genuine wild population. The three food strains were not differentiated from the wild
182 strains.

183 *Penicillium fuscoglaucum* displayed the highest genetic diversity level, followed by *P. biforme,* as
184 indicated by several diversity indices, such as the number of SNPs, π and Watterson's θ (Table
185 S1), as well as strain scattering in PCA and neighbor-net network (Figures 2A and 2C). The very
186 low genetic diversity detected within *P. camemberti* suggests that it represents a clonal lineage
187 (Table S1, Figures 2A and 2C). The diversity in *P. biforme* was similar to that previously found in
188 the non-cheese *P. roqueforti* population (named the "silage/food spoiler" population in [26], Table
189 S1), while the diversity in *P. camemberti* was similar to that previously found in the domesticated
190 clonal cheese *P. roqueforti* population (named the "non-Roquefort" population in [26]). The
191 average nucleotide identity between *P. camemberti* and *P. biforme* was very high (99.8%) and similar
192 to the 99.5% identity between the domesticated mold *Aspergillus oryzae* and its wild relative *A.*
193 *flavus* [42].

194 The long branches and the lack of reticulation observed in the neighbor-net analysis further
195 confirmed the clonality of *P. camemberti*, contrasting with the footprints of recombination
196 detected within both *P. biforme* and *P. fuscoglaucum* (Figure 2C). In *P. fuscoglaucum,* abundant
197 reticulation was observed, right to the branch edges (Figure 2C), reinforcing the view that *P.*
198 *fuscoglaucum* corresponded to a sexual population thriving in natural environments, while strains of
199 *P. camemberti* and *P. biforme* used in the food industry are replicated clonally.

200 We studied the mating-type genes controlling fungal mating compatibility to determine the
201 likelihood of sexual reproduction occurrence within each species. In heterothallic fungal species,
202 sexual reproduction occurs only between two haploid individuals carrying different alleles
203 (MAT1-1 and MAT1-2 in filamentous ascomycetes) at the mating-type locus. A lack of sexual

204 reproduction leads to relaxed selection on the mating-type genes, as they are no longer used, and
205 a departure from a balanced ratio between mating types. Neither mating-type allele presented any
206 evidence of loss-of-function mutations or SNPs in any of the genomes. In *P. fuscoglaucum,* we
207 found no significant departure from the 1:1 ratio of the two mating-type alleles expected under
208 conditions of regular sexual reproduction ($\chi^2 = 1.14$; df = 1; P = 0.29). By contrast, *P. biforme*
209 displayed a significant departure from the 1:1 ratio ($\chi^2 = 14.29$; df = 1; $P < 1e^{-3}$), with MAT1-2
210 strains predominating. All 19 *P. camemberti* strains carried the MAT1-2 allele, providing further
211 evidence of the clonality of *P. camemberti* [43,44]. The observed clonality in *P. camemberti* may be
212 explained by the recent selection of a white mutant and its use in industrial cheese production
213 [45], whereas *P. biforme* is more widely used in the production of artisanal cheeses and farmers
214 may make use of diverse local strains.

215

216 Two genetic clusters within *Penicillium camemberti*

217 We found two genetic clusters within *P. camemberti*, as indicated by the population structure
218 analysis implemented in NGSadmix at $K = 5$, the PCA and the neighbor-net without *P.*
219 *fuscoglaucum*, as well as the PCA with only *P. camemberti* (Figures 2B, 2D and S2). The two *P.*
220 *camemberti* genetic clusters displayed very few shared SNPs (13% of the 19,510 SNPs, Table S1),
221 but a very high nucleotide identity (99.998%). The two genetic clusters within *P. camemberti*
222 separated strains isolated from soft cheeses, such as Camembert, and strains isolated from other
223 kinds, such as Rigotte de Condrieu and Saint Marcellin. The *P. caseifulvum* type strain (red star on
224 Figure 2D) clustered with *P. camemberti* strains isolated from non-soft cheeses. The genetic cluster
225 including the *P. caseifulvum* type strain will be referred to hereafter as *P. camemberti* var. *caseifulvum,*
226 and the genetic cluster including the *P. camemberti* type strain will be referred to as *P. camemberti*
227 var. *camemberti*. The nucleotide diversity was very low within both varieties (Table S1, Figure 2B).
228 We designed two pairs of primers to identify by a single PCR reaction the *P. camemberti* variety,
229 each yielding a specific amplicon size (Data S5).
230 The population structure analysis also suggested further genetic subdivision within each of *P.*
231 *fuscoglaucum* and *P. biforme*; these subdivisions did not correspond to the environment from which
232 the isolate was obtained, type of cheese, or any other variable for which we had data; it did not
233 correspond to either a strong subdivision in the PCAs or the neighbor-nets (Figures 2 and S1).
234

235 Presence/absence polymorphism of the two horizontally transferred regions
236 likely to be advantageous in cheese

237 By contrast to the first 454-sequenced FM013 reference genome [27], in which the *Wallaby*
238 horizontally transferred region was found to be fragmented, the PacBio-based genome assembly
239 of the same strain showed that the part of *Wallaby* present in the genome formed a single block
240 (scaffold 18, positions 102,711 to 538,265). Two genes in *Wallaby* have been highlighted as
241 putatively important for cheese-making. One, encoding the *Penicillium* antifungal protein (PAF)
242 [27], was absent from FM013, whereas the other, encoding Hce2, a protein with known
243 antimicrobial activities, was present. All *P. camemberti* var. *camemberti* strains harbored the same
244 *Wallaby* region as FM013 (Figure 3C), whereas strains from the *P. camemberti* var. *caseifulvum*
245 lineage carried fragmented and duplicated segments of the *Wallaby* fragment present in FM013, in
246 which the Hce2 gene was systematically present. We also found *Wallaby* fragmented with some
247 duplicated regions in 18 *P. biforme* strains; it was completely absent from the 10 remaining strains
248 (Figure 3C). *Wallaby* appeared to be even more fragmented in *P. fuscoglaucum* strains, always
249 lacking Hce2 but in one strain, and it was completely absent in nine strains. In contrast to *P.*
250 *roqueforti,* in which *Wallaby* was only found in strains isolated from the cheese environment,
251 fragments of *Wallaby* were found in some non-cheese strains of *P. fuscoglaucum*. However, even in
252 *P. fuscoglaucum*, no genetic variability could be detected within the *Wallaby* fragments, indicating
253 that they have also likely been acquired by horizontal gene transfer in this species. So far, we
254 found no species in which *Wallaby* showed genetic variability, indicating that we have not
255 identified the donor species.

256 The other horizontally transferred region thought to be beneficial for growth on cheese [28], the
257 57 kb *CheesyTer* region, was not located at the end of one scaffold in *P. camemberti* (scaffold 17,
258 positions 690,636 to 751,978 in FM013), like its position in *P. roqueforti*. We found no strict
259 association between *CheesyTer* and *Wallaby* in the clades studied here, whereas all strains of *P.*
260 *roqueforti* studied so far carried either both or neither of these regions [28]. In *P. camemberti,* 16 of
261 19 strains carried the whole *CheesyTer* region, with an identity of 100% (Figure 3B). The other
262 three *P. camemberti* strains (LCP00584T, LCP05527 and UBOCC-A-108096), belonging to the *P.*
263 *camemberti* var. *camemberti* cluster, completely lacked *CheesyTer* and an 80 kb-downstream fragment.
264 These three strains were isolated from Camembert cheeses around 1905, suggesting that the
265 horizontal transfer of *CheesyTer* to *P. camemberti* occurred after 1905, in a single strain that has
266 been since clonally cultured. In *P. biforme,* 25 of 28 strains carried *CheesyTer*, with an identity
267 >99.8% between strains. The identity between *P. biforme* and *P. camemberti* was also >99.8%,
268 except in the 5 kb region at the start of *CheesyTer,* in which the similarity between the two species
269 dropped to 90%. The differences between *P. camemberti* and *P. biforme* at the start of *CheesyTer*
270 appeared to be mainly due to C:G to T:A mutations in *P. camemberti,* corresponding to the repeat-

271 induced point (RIP) mutations specific to fungi occurring in repeated sequences during sexual
272 reproduction [46]. As *P. camemberti* was found to be clonal and no RIP footprints were detected
273 within the rest of *CheesyTer*, these findings suggest that *CheesyTer* was transferred to an ancestor of
274 the two *P. camemberti* varieties with this RIPed sequence already present. *CheesyTer* was completely
275 absent from six strains of *P. fuscoglaucum* (Figure 3) and partially present in the other strains. One
276 *P. fuscoglaucum* strain (LCP00218, isolated from rubber) carried two copies of a 30 kb *CheesyTer*
277 terminal region. The two *CheesyTer* genes involved in lactose metabolism [28], one encoding a
278 lactose permease and the other a β-galactosidase, were present in all strains carrying at least one
279 part of *CheesyTer* in *P. camemberti, P. biforme* and *P. fuscoglaucum* (Figure 3B).

280

281 Looking for genomic footprints of adaptation in *P. camemberti* and *P.* 282 *biforme*

283 Not many genomic selection scans could be performed given the high level of clonality in *P.*
284 *camemberti*, as any selective sweep will hitchhike the whole genome. We investigated whether some
285 genomic regions showed greater differentiation than the genomic background when comparing
286 *P. biforme* on the one hand to either *P. camemberti* or *P. fuscoglaucum* on the other, as genomic
287 islands of differentiation can be footprints of divergent adaptation. We used a genomic selection
288 scan approach using 75kb windows based on an absolute measure of differentiation, d_{xy}, instead
289 of the F_{ST} statistic, as the latter is influenced by within-population diversity (Data S3). No
290 particularly striking genomic island of differentiation could be identified (Data S3, S4) and genes
291 within windows in the 1% highest d_{xy} were not enriched in any particular function. We also
292 performed genome scans of diversity based on π and Watterson's θ (Data S3) but no particular
293 genomic islands could be detected either. The genes within windows in the 1% highest d_{xy}
294 however contained significantly less genes with predicted functions than the rest of the reference
295 genome, suggesting that these regions contained more rapidly evolving genes.

296

297 Phenotypes better suited for cheese production in *P. camemberti* and *P.* 298 *biforme* than in the closely related wild species *P. fuscoglaucum*

299 Strains selected by humans are expected to display specific traits beneficial for cheese production,
300 such as faster growth in cheese, colonies of an attractive color, salt tolerance, an absence of
301 harmful toxin production and efficiency at excluding undesirable microorganisms. We performed
302 a set of laboratory experiments to determine whether cheese strains had evolved traits beneficial

303 for cheese production not present in the closely related wild *P. fuscoglaucum* lineage occurring in
304 other environments. In each of *P. camemberti* and *P. biforme,* we considered as a single group the
305 strains from cheese and other food types (e.g. dry sausage, bread) because they were not
306 genetically differentiated and the same strains are sold by producers for cheese and dry sausage
307 maturation.
308
309 We first investigated whether *P. camemberti* and *P. biforme* had acquired traits enabling them to
310 grow more rapidly than *P. fuscoglaucum* in cheese-making conditions. Rapidly growing fungi can be
311 beneficial for cheese-making, as they are more capable of excluding spoiler bacteria, yeasts and
312 molds [11] and promote faster cheese maturation. The process of cheese maturation begins with
313 the transfer of salted milk curd to dark caves at a low temperature and high humidity. We
314 therefore grew 61 strains on various media (unsalted cheese, salted cheese, minimal and malt
315 media), in the dark and at different temperatures and humidity levels (25°C at ambient relative
316 humidity, 10°C at 98% humidity or 8°C at 100% humidity), and we measured colony radial
317 growth after 10 days. The non-parametric Kruskal-Wallis test as well as the ANOVA showed no
318 significant effect on growth of the substrate from which the isolate was originally obtained (i.e.
319 food versus non-food). We found a significant effect on growth of the culture media, species,
320 temperature/humidity conditions and the *P. camemberti* variety. The ANOVA further showed
321 significant effects on growth of the interactions between culture medium and species on the one
322 hand and between culture medium and temperature/humidity on the other hand (Data S1A and
323 S2A, Figure 4A). However, the significance of the interaction between culture medium and
324 temperature/humidity conditions (Data S1A and S2A) reflected the very slow growth at low
325 temperatures, making it difficult to detect significant differences between media. Although the
326 two *P. camemberti* varieties were genetically very closely related, *P. camemberti* var. *caseifulvum* grew
327 similarly to *P. biforme* (post-hoc Tukey honestly significant difference - HSD test $P = 1$), with
328 greater radial growth on unsalted cheese and salted cheese, and slower growth on malt and
329 minimal media than *P. fuscoglaucum*. By contrast, *P. camemberti* var. *camemberti* displayed weaker
330 radial growth than any other lineage on all media (Figure 4A), but was fluffier (Figure 1B).
331 Furthermore, *P. camemberti* and *P. biforme* grew less rapidly than *P. fuscoglaucum* on minimal and
332 malt media, indicating a disadvantage for growth in harsh conditions, as expected for cheese
333 strains due to relaxed selection on functions useful only in wild environments. These findings
334 demonstrated that *P. camemberti* var. *caseifulvum* and *P. biforme* have acquired rapid radial growth in
335 cheese which was likely beneficial for fast maturation. *Penicillium camemberti* var. *camemberti* grew
336 much less radially on cheese than any other lineage, but grew much more vertically (i.e. being

337 fluffier), with the mycelium growing up to the lid of the Petri dishes (Figure 1B). *Penicillium*
338 *camemberti* var. *camemberti* might have been selected for fast vertical growth to make soft cheeses
339 with a fluffy aspect within a short maturation period. Short maturation is beneficial for
340 production when using cow or goat milk available throughout the year, as is the case for
341 Camembert cheeses. In contrast, ewe milk is only available a few months a year, so that
342 Roquefort cheeses need to be stored for several months to be sold and consumed throughout the
343 year, in which case too fast maturation was not beneficial before fridges were available. Such
344 historical long storage can explain a selection for slow growth in the Roquefort *P. roqueforti*
345 population [26], in contrast to a selection for fast growth in *P. camemberti* and *P. biforme.*
346
347 High salt concentrations in cheeses prevent the growth of contaminants and cheese fungi may
348 have adapted to such conditions. We found an effect of salt on *P. biforme* growth, with strains
349 growing more rapidly on salted than unsalted cheese medium (post-hoc Tukey HSD test P =
350 0.006) and more rapidly on salted cheese medium than *P. camemberti* (post-hoc Tukey HSD test P
351 = 0.0001). Salt had no effect on the growth of *P. camemberti* (post-hoc test P = 1) or *P. fuscoglaucum*
352 (post-hoc test P = 1). There may have been stronger selection for salt tolerance in *P. biforme,*
353 which is used to inoculate blue and goat cheeses, both being more salty than soft cheeses, such as
354 Brie or Camembert [47], for which *P. camemberti* var. *camemberti* is used.
355
356 We investigated whether cheese lineages had become whiter, which can be more attractive to
357 some consumers than gray-green mold on cheese, by comparing the opacity of lineages on cheese
358 medium, as opacity increases with the brightness and fluffiness of a colony [48]. We found
359 significant effects of species and variety (Figure 4C and Data S1B and S2B), with *P. camemberti*
360 being significantly more opaque (i.e. brighter and fluffier) than other lineages, and *P. camemberti*
361 var. *camemberti* even more so compared to var. *caseifulvum* (post-hoc Tukey HSD test P = 0.03).
362 This is consistent with the white and fluffy aspect of the crust of Camembert and Brie, made
363 with *P. camemberti* var. *camemberti*, whereas *P. camemberti* var. *caseifulvum* was found here in cheeses
364 with a grayer and less fluffy crust, such as Saint Marcellin or Rigotte de Condrieu (Figures 1 and
365 4B-C). The *P. camemberti* var. *caseifulvum* and *P. biforme* lineages did not differ significantly from
366 each other (post-hoc Tukey HSD test P = 1) and both were brighter than the wild blue-green *P.*
367 *fuscoglaucum* (Figure 4B).
368
369 We also investigated whether cheese lineages produced smaller amounts of cyclopiazonic acid
370 (CPA), which is cytotoxic to humans [38,40], than strains isolated from other environments.

371 *Penicillium camemberti* has been reported to produce CPA on yeast extract sucrose (YES) medium
372 and at very low, non-toxic concentrations in cheese [39]. None of the *P. camemberti* var. *caseifulvum*
373 strains tested here produced CPA on YES (Figure 4D), consistent with findings for the type
374 strain [49]. By contrast, *P. camemberti* var. *camemberti, P. biforme* and *P. fuscoglaucum* produced CPA,
375 the highest levels being obtained with *P. biforme* and *P. fuscoglaucum* (Figure 4D, Data S1C and
376 S2C). The CPA biosynthesis cluster in the genome appeared to be functional, with six genes
377 present (*cpaA*, *cpaD*, *cpaO*, *cpaH*, *cpaM* and *cpaT*) but not the *cpa*R gene encoding a regulatory
378 protein (Figure S3). The only exceptions were the six *P. camemberti* var. *caseifulvum* strains, in which
379 a 2 bp deletion in the *cpaA* gene led to a frameshift. The *cpaA* gene encodes a polyketide
380 synthase/non-ribosomal peptide synthase responsible for the first step of the CPA biosynthetic
381 pathway so that a non-functional protein probably prevents CPA production on all substrates.
382 Humans have often selected fungal strains unable to produce harmful toxins for use in food, as
383 exemplified by the *Aspergillus oryzae* strains used to ferment Asian food products, which do not
384 produce aflatoxins, whereas the wild ancestor, *A. flavus,* does [50]. In the blue cheese fungus *P.*
385 *roqueforti*, strains belonging to the non-Roquefort population were also found unable to produce
386 mycophenolic acid due to a 174 bp deletion in the *mpaC* gene [51].
387
388 Cheese is a nutrient-rich environment in which many microorganisms can thrive, including
389 undesirable food spoilage organisms. We therefore investigated whether the cheese lineages were
390 better able to outcompete challengers. We allowed strains of *P. camemberti* var. *camemberti* (*n*=3), *P.*
391 *camemberti* var. *caseifulvum* (*n*=2) or *P. biforme* (*n*=10) to grow as lawns on the surface of the cheese
392 medium, which were inoculated 24 h later with a single spot, in the middle of the Petri dish, using
393 a challenger: *Geotrichum candidum* (*n*=5), *P. biforme* (*n*=6), *P. fuscoglaucum* (*n*=5), or *P roqueforti* (*n*=12,
394 three strains from each of the four known genetic clusters [26]). The species and the variety used
395 as the lawn had significant effects on the growth of the challenger (ANOVA and non-parametric
396 Kruskal-Wallis tests, Data S1D and S2D, Figure 5). The two lineages of *P. camemberti* prevented
397 the growth of all challengers more effectively than *P. biforme*, and *P. camemberti* var. *camemberti* was
398 the most effective. Furthermore, we observed significant differences between the species used as
399 challengers, with *G. candidum* having the highest growth differential on the various lawns, growing
400 less well on a *P. camemberti* var. *camemberti* lawn than other lawns. *Geotrichum candidum* is a fungus
401 that is also present on the surface of soft cheeses. The exclusion effect may be mediated by the
402 fluffy morphology, resulting in the fungus occupying more space and using more resources,
403 and/or by biochemical interactions.
404

405 # Conclusion

406 Population structure analyses based on whole-genome sequences revealed that *P. biforme, P.*
407 *camemberti* and *P. fuscoglaucum* each formed a separate genetic cluster, with *P. biforme* and *P.*
408 *camemberti* being sister clades. Strains identified as *P. commune* corresponded to either *P. biforme* or
409 *P. fuscoglaucum*, being thus non-monophyletic (Figure 1A), so this name should not be used any
410 longer. Furthermore, considering all *"P. commune"* strains to be wild, as it is currently common in
411 the literature, leads to wrong inferences about trait evolution [41], as some *"P. commune"* strains
412 (those corresponding to *P. biforme*) belong to a domesticated genetic cluster. We found that *P.*
413 *camemberti* and *P. biforme* were sister clades specific to the cheese environment, whereas the more
414 distantly related clade *P. fuscoglaucum* was found mostly in natural environments. These
415 relationships and the lower diversity of *P. biforme* than *P. fuscoglaucum*, with even lower levels of
416 diversity in *P. camemberti*, suggest that an ancient domestication event led to the divergence
417 between *P. biforme* and *P. fuscoglaucum* and that a much more recent domestication event led to the
418 differentiation between *P. camemberti* and *P. biforme*. We used here the *P. biforme, P. camemberti* and
419 *P. fuscoglaucum* Latin species names as they are commonly used and corresponded to differentiated
420 and monophyletic clades based on genome-wide information, harbored contrasted phenotypes
421 and originated from different ecological niches. We did not address here the question of species
422 status or definition, as they are notoriously difficult to elucidate among closely related fungal
423 groups, especially in clonal ascomycetes [52,53].

424 Domestication often leads to bottlenecks [54–57], which can be very severe in clonally multiplied
425 fungi [26,58]. Consistent with this scenario, we found evidence of phenotypic adaptation to
426 cheese-making in *P. biforme* relative to the wild *P. fuscoglaucum* clade, with a whiter color, faster
427 growth on cheese medium under cave conditions and lower levels of toxin production. These
428 signs of domestication were even more marked in *P. camemberti* than in *P. biforme*. We observed a
429 similar evolution of traits as in populations of the domesticated blue-cheese fungus *P. roqueforti*,
430 which grows more rapidly on cheese medium, is more competitive against contaminant
431 microorganisms, and grows less well under harsh conditions than wild strains [26,28]. Such
432 convergent evolution under similar selective pressures suggests that evolution may repeat itself,
433 as already reported in natural populations of *Anolis* lizards [59,60], three-spined sticklebacks [61],
434 Mexican cavefishes [62], and cichlid fishes in African rifts [63] or crater lakes in central America
435 [64,65]. Phenotypic convergent evolution has been also reported in domesticated organisms,
436 particularly in crop plants, for the loss of seed shattering, minimization of seed dormancy and

437 increase in seed size and number, for example, with most of these changes resulting from
438 different genomic alterations in different species [66,67].

439 We identified two genetically differentiated *P. camemberti* varieties, *P. camemberti* var. *camemberti* and
440 *P. camemberti* var. *caseifulvum,* with contrasting phenotypic features, used in the production of
441 different kinds of cheese. The *P. camemberti* var. *camemberti* strains were white and were isolated
442 from Camembert or Brie. Their radial growth was slower but their mycelia were much fluffier
443 than the other *Penicillium* strains tested (Figure 1B) and they produced similar amounts of CPA
444 compared to *P. biforme* strains. They excluded fungal competitors effectively, as previously
445 suggested [68], probably due to their fluffy morphology, taking up the available space and
446 monopolizing resources. *Penicillium camemberti* var. *caseifulvum* strains were gray-green, unable to
447 produce CPA, and were isolated from cheeses other than Camembert, such as St Marcellin or
448 Rigotte de Condrieu (Figure 1A). They displayed more rapid radial growth than *P. camemberti* var.
449 *camemberti* and grew similarly to *P. biforme* strains on cheese medium in cave conditions. The
450 difference in terms of cheese type usage between the two varieties should be confirmed in future
451 studies using more strains, but it seems to correspond well to the contrasted identified
452 phenotypes, and in particular color and fluffiness. The existence of two genetically and
453 phenotypically differentiated lineages with possibly different uses in cheese-making suggested
454 that these lineages emerged as a result of different human selection pressures, as reported for the
455 domesticated blue-cheese fungus *P. roqueforti* [26], wine-making yeasts [12,15], maize, rice,
456 tomatoes, dogs, chickens and horses [1–6].

457 *Penicillium camemberti* var. *camemberti* is the emblematic species used to inoculate soft cheeses, such
458 as Camembert and Brie. According to the technical specifications for Brie de Meaux and Brie de
459 Melun PDOs and Camembert, the crust of the cheese must be white and fluffy, and inoculation
460 with *P. candidum* (a synonym of *P. camemberti*) is even specified for Brie de Meaux. These
461 specifications are recent (20th century) and seem to have had a negative impact on the diversity
462 of the fungi used for making these kinds of cheeses, as a single clonal lineage now predominates.
463 The genetic diversity of *P. roqueforti* has also been greatly reduced by recent industrialization
464 [26,29], although PDO specifications that local strains must be used have protected diversity to
465 some extent in Roquefort cheeses. Camembert and Brie cheeses were gray-green before the 20th
466 century [32,34,45], as illustrated by a 19th century painting by Marie Jules Justin entitled
467 "Symphonie des fromages en Brie majeur - Nature morte au fromage". First mentions of Brie
468 cheeses date from middle ages [32]. These historical records are consistent with our findings,
469 which suggest a first domestication event leading to the emergence of the gray-green mold *P.*

470 *biforme,* subsequently followed, around 1900, by the domestication of *P. camemberti,* with the
471 selection of two different varieties displaying more marked signs of domestication than *P. biforme.*

472 These findings have industrial implications, as they reveal the existence of three closely related
473 but different lineages that have evolved traits beneficial for cheese-making, with different
474 phenotypic traits selected according to usage. This study should foster further research, as it
475 would be interesting to assess other important traits, such as volatile compound production, and
476 the efficiencies of lipolysis and proteolysis. Our findings raise questions about the use of limited
477 numbers of clonal strains for cheese-making, which tends to lead to degeneration, limiting the
478 possibilities for further improvement, which is currently a major concern in the agrofood sector
479 [69], despite the great geneticist Nikolai Vavilov known for having identified the centers of origin
480 of cultivated plants long ago highlighting the importance of genetic diversity in domesticated
481 organisms for variety improvement and diversification [70].

482

483 # Materials and Methods

484 ## Strain collection and DNA extraction

485 We analyzed strains isolated from 37 cheeses from five countries around the world (e.g.
486 Camembert, Saint Marcellin, tomme, Sein de Nounou). We also collected strains from dried
487 sausages (n = 8) and moldy food (e.g. bread, n = 4). Twelve were isolated from other substrates
488 than food. Spores were sampled from the food sources and spread on Petri dishes containing
489 malt-agar medium, which were then incubated for five days at 25°C. A total of 41 strains were
490 obtained from public strain collections (the Laboratoire de Cryptogamie (LCP strains) at the
491 National Museum of Natural History in Paris (France), the Laboratoire Universitaire de
492 Biodiversité et Ecologie Microbienne (LUBEM, UBOCC strains) in Plouzané (France) and the
493 Westerdijk Fungal Biodiversity Institute (CBS strains) in Utrecht (The Netherlands)). The
494 LMA421 strain was kindly given by Steve Labrie (University of Laval, Canada). We obtained
495 strains from natural environments (e.g. wood or natural cave walls) from eight different
496 countries. Detailed information about the strains used in this study can be found in the
497 Supplemental information. For each strain, single-spore cultures were generated by a dilution
498 method, to ensure that only a single haploid genotype was cultured for each strain. We checked
499 the species identification of all strains by Sanger sequencing of the ß-tubulin gene (primers Bt2a -
500 GGTAACCAAATCGGTGCTGCTTTC and bt2b – AACCTCAGTGTAGTGACCCTTGGC
501 [71]) and the PC4 microsatellite flanking regions [36] (primers PC4F –

502 CAAGCTGGCCGATAACCTG and PC4R – CCATCCGCTTGATTTCTCCT), to distinguish
503 between *P. biforme* and *P. camemberti*. All the strains were added to the ESE (Ecology Systematics
504 Evolution Laboratory) collection, under an ESE accession number, and are available from the
505 corresponding collections (ESE, LUBEM, MNHN). For each strain, we used the Nucleospin Soil
506 Kit (Machernary-Nagel, Düren, Germany) to extract DNA from fresh haploid mycelium grown
507 for five days on malt agar.
508

509 **Genome sequencing, assembly and mapping**

510 Sequencing was performed with Illumina HiSeq 2500 paired-end technology (Illumina Inc.), with
511 a mean insert size of 400 bp, at the INRA GenoToul platform, to obtain 10-50X coverage
512 (Supplemental information). In addition, the genome of the *P. camemberti* LCP06093 (also known
513 as FM013) reference strain, initially sequenced with 454 technology [27], was resequenced at 34x
514 using the PacBio's long-read technology, assembled with Canu 1.8 [72] and polished with pilon
515 1.23 [73]. The PacBio assembly and reads for all samples have been deposited at the NCBI
516 Sequence Read Archive under BioProject ID XXXX.
517 Reads were trimmed with Trimmomatic [74] (cut adapter, options PE, LEADING:3
518 TRAILING:3 SLIDINGWINDOW:4:25 MINLEN:36) and mapped onto the high-quality
519 PacBio reference genome of LCP06093 with Bowtie2 [75] (options very-sensitive-local, --
520 phred33 PL:ILLUMINA -X 1000). Mapped reads were filtered for PCR duplicates with picard
521 tools (http://broadinstitute.github.io/picard) and realigned on the PacBio reference genome with
522 GATK [76]. Single-nucleotide polymorphisms (SNPs) were called with GATK HaplotypeCaller,
523 which provides one gVCF per strain (option -ERC GVCF). GVCFs were combined with GATK
524 CombineGVCFs, genotyped with GATK genotype and hard filtered in accordance with GATK
525 best practice workflow recommendations (https://software.broadinstitute.org/gatk/best-
526 practices/).
527

528 **Analysis of the mating-type genes**

529 We looked for the mating type genes in *P. camemberti, P. biforme* and *P. fuscoglaucum* by extracting
530 the region including the flanking genes *APN2* and *SLA2* from all BAM files.
531

532 **Statistics of population genetics and genomic scans**

533 Nucleotide diversity indices (π and Watterson's θ) and genetic differentiation indices (F_{ST} and d_{xy})
534 were calculated using the R package PopGenome [77] for the four genetic clusters *P. biforme, P.*
535 *fuscoglaucum, P. camemberti.* var. *camemberti* and *P. camemberti* var. *caseifulvum*. π and Watterson's θ were

536 calculated per site. We also performed genomic scans along the 20 largest scaffolds of the *P.*
537 *camemberti* PacBio reference using sliding windows of 75 kb and step sizes of 7.5 kb. Genome
538 scans for π and Watterson's θ were performed within each of the four genetic clusters (*P. biforme,*
539 *P. fuscoglaucum, P. camemberti.* var. *camemberti* and *P. camemberti* var. *caseifulvum*). Genome scans based
540 on F_{ST} and d_{xy} indices were performed between all possible pairs of the four clusters (*P. biforme, P.*
541 *fuscoglaucum, P. camemberti.* var. *camemberti* and *P. camemberti* var. *caseifulvum*). We tested whether the
542 genes belonging to windows with the 1% highest d_{xy} values were enriched in particular predicted
543 functions (interpro). We performed gene enrichment analyses using Fisher exact tests ('fisher.test'
544 function in the R package stats) corrected for multiple testing ('p.adjust' function in the R
545 package stats with the false discovery rate, 'fdr', method). We also tested whether the number of
546 genes with predicted functions within windows with the 1% highest d_{xy} values deviated from
547 expectations based on the proportion in the full genome of the PacBio reference using a χ^2 test.
548 We calculated the average pairwise identity between species pairs, as well as the number of fixed,
549 private and shared SNPs. When comparing two populations, private SNPs are sites that are
550 polymorphic in only one population (e.g. A/C for individuals in population 1 and T for
551 individuals in population 2), shared SNPs are polymorphic positions within the two populations
552 (e.g. A/C in population 1 and A/T in population 2) and fixed SNPs are positions that are fixed in
553 each population but different between populations (e.g. A for all individuals of population 1 and
554 G for all individuals of population 2).
555

556 **Genetic structure**

557 We used the dataset of 392,072 SNPs to infer population structure. We inferred individual
558 ancestry from genotype likelihoods based on realigned reads, by assuming a known number of
559 admixing populations, ranging from $K = 2$ to $K = 6$, using NGSadmix from the ANGSD package
560 [78]. A hundred independent runs were carried out for each number of clusters (*K*). We identified
561 clustering solutions in replicated runs for each *K* value and displayed population structure
562 graphically using the R function 'barplot'. We used the R package phangorn [79] for neighbor-net
563 analyses, and the 'prcomp' function of R for principal component analysis (PCA).
564

565 **Phylogenetic analyses**

566 We inferred phylogenetic relationships among the 61 isolates using the dataset of 1,124,983 SNPs
567 in a maximum likelihood framework using IQ-Ttree2 [80]. We also used *P. cavernicola* (LCP05631
568 strain) and *P. palitans* (ESE00153 and LCP04862 strains). The tree has been midpoint rooted. The

569 best-fit model chosen using ModelFinder [81] according to BIC was TVM+F+ASC+R3. Branch
570 supports are SH-aLRT [82] and ultrafast bootstrap support [83].
571

572 **Laboratory experiments**

573 *Sampling and calibration*

574 All experiments were performed on the whole collection of *P. camemberti* (*n* = 19), *P. biforme* (*n* =
575 28) and *P. fuscoglaucum* (*n* = 14) strains, including 35 strains isolated from cheeses, 10 from other
576 food environments (e.g. bread, sausage) and 14 strains isolated from environments other than
577 food. Experiments were initiated with spore suspensions calibrated to 10_7 spores/mL with a
578 hemocytometer, under the constraint of the low rate of sporulation in *P. camemberti.*
579

580 *Growth in different conditions and on different media*

581 We investigated whether strains isolated from cheese displayed faster or slower radial growth
582 than strains isolated from other environments when cultured on cheese or other substrates, and
583 under conditions similar to those in maturation caves or other conditions. We prepared four
584 different culture media: a cheese medium without salt, a salted cheese medium (17g/L,
585 corresponding to a typical cheese), a malt medium and a minimal medium. The cheese media
586 were produced from an unsalted drained cow's milk cheese from Coubertin Farm in Saint Rémy-
587 les-Chevreuse (France), as previously described [68]; we added five drops of blue food coloring
588 to these media, to make it easier to distinguish white fungal colonies from the medium. The
589 cheese media were rich in lipids and proteins, whereas the malt medium (20 g/L) was rich in
590 carbohydrates. The minimal medium [84] contained only the trace elements necessary for fungal
591 survival (i.e. iron sulfate, zinc sulfate, boric acid, magnesium chloride, copper sulfate, ammonium
592 heptamolybdate, cobalt chloride, EDTA, sodium nitrate, potassium chloride, potassium
593 phosphate and magnesium sulfate). All media were sterilized in an autoclave (121°C for 20 min
594 for the malt and minimal media, 110°C for 15 min for the cheese media). Each 90 mm-diameter
595 Petri dish was filled with 20 mL of the appropriate medium.
596

597 We allowed the fungi to grow in the dark, under three different conditions for each strain and
598 each medium: 10°C with 98% humidity (Camembert cave conditions), 8°C with 85% humidity
599 (cave conditions for other cheeses) and 25°C with ambient relative humidity (ambient
600 conditions). We had 244 Petri dishes in total for each set of conditions, and we left the colonies
601 to grow for 10 days. Images of the Petri dishes were obtained with a Scan 1200 from Interscience
602 and analyzed with IRIS [48] for growth and color-opacity measurements. The colony size was

603 estimated as the colony area in pixels. The opacity score was defined as the sum of the brightness
604 values for all the pixels within the colony bounds; this score thus captures both the color and the
605 three-dimensional colony growth (as opacity increases with fluffiness) and can be used as a proxy
606 to assess the fluffiness of the fungal colony.
607
608 *Mycotoxin production*
609 We selected at least six strains from each genetic cluster for this experiment (Data S1). For
610 measurements of mycotoxin production, we used 1 µL of calibrated spore suspension (10_7
611 spores/mL) for each of the 61 strains to inoculate YES (yeast extract sucrose) agar medium
612 buffered at pH 4.5 with phosphate-citrate buffer and characterized by a high C/N ratio to favor
613 mycotoxin production as already described in [51]. Each culture was performed in triplicate for
614 mycotoxin analyses as well as for fungal dry weight measurements. The plates were incubated at
615 25°C in the dark for 10 days and were then stored at -20°C until mycotoxin analysis.
616 For mycotoxin extractions [51], we homogenized the thawed samples with an Ultraturrax T25
617 (IKA, Heidelberg, Germany) before recuperating 4g aliquots. Then, 25 mL of acetonitrile (ACN)
618 supplemented with 0.1% formic acid (v/v) was added, samples were vortexed for 30 sec followed
619 by 15 min sonication. Extracts were then centrifuged for 10 min at 5000g and the recuperated
620 supernatants were directly filtered through 0.2 µm PTFE membrane filters (GE Healthcare Life
621 Sciences, UK) into amber vials. All samples were stored at -20°C until analyses.
622 Mycotoxin detection and quantification were performed using an Agilent 6530 Accurate-Mass
623 Quadropole Time-of-Flight (Q-TOF) LC/MS system equipped with a Binary pump 1260 and
624 degasser, well plate autosampler set to 10°C and a thermostated column compartment. Filtered
625 samples (2µl) were injected into a ZORBAX Extend C-18 column (2.1x50mm and 1.8 µm, 600
626 bar) maintained at 35°C with a flow rate of 0.3 ml/min using mobile phase A (milli-Q water +
627 0.1% formic acid (v/v) and 0.1% ammonium formate (v/v) and mobile phase B (ACN + 0.1%
628 formic acid). Mobile phase B was maintained at 10% for 4 min followed by a gradient from 10 to
629 100% for 16 min. Then, mobile phase B was maintained at 100% for 2 min before a 5 min post-
630 time. Cyclopiazonic acid (CPA) was ionized in electrospray ionization mode ESI+ in the mass
631 spectrometer with the following parameters: capillary voltage 4 kV, source temperature 325°C,
632 nebulizer pressure 50 psig, drying gas 12 l/min, ion range 100-1000 m/z. CPA had a retention
633 time of 15.3 min and was quantified using the $[M+H]_+$ 337.1545 m/z ion and $[M+Na]_+$ 359.1366
634 qualifier ion (ESI+). Other extrolites included in our internal database that were targeted (in
635 ESI+ and ESI- modes) were andrastin A, citreoviridin, citrinin, ermefortins A & B, (iso)-

636 fumigaclavin A, griseofulvin, meleagrin, mycophenolic acid, ochratoxin A, patulin, penitrem A,
637 PR toxin, roquefortin C, sterigmatocystin.
638 We used a matrix matched calibration curve (R_2 >0.99) for reliable mycotoxin quantification with
639 final concentrations ranging from 10 to 10000 ng/ml. Method performance and mycotoxin
640 determination was carried out as previously described in [51]. Specific mycotoxin production was
641 expressed as ng per g fungal dry weight.
642
643 *Competition*
644 We inoculated salted cheese with 150 µL of a calibrated spore solution (10_7 spores/mL), which
645 we allowed to grow as a lawn. We used two *P. camemberti* var. *caseifulvum* strains (LCP05630 and
646 ESE00019), three *P. camemberti* var. *camemberti* strains (FM013, LCP04810 and LCP05527) and 10
647 *P. biforme* strains (ESE00018, ESE00086, ESE00089, ESE00126, ESE00139, LCP05531,
648 LCP06701, ESE00154, LCP06620 and LCP05496). After 24 h of growth, we deposited a single
649 20 µL drop of calibrated spore solution from a challenger species (10_7 spores/mL) in the middle
650 of the Petri dish. The challenger species used were *G. candidum* (*n*=4, ESE00163 and ESE00165,
651 CBS11628 and VTTC4559), *P. biforme* (*n*=6, ESE00126, ESE00089, ESE00018, ESE00154,
652 LCP005531, LCP06701), *P. fuscoglaucum* (*n*=5, ESE00090, LCP06621, LCP000797, LCP04799,
653 LCP03239), and *P roqueforti* (*n*=12, three from each of the four genetic clusters described in [26]:
654 LCP06036, LCP06037, LCP06128, LCP06043, LCP06098, LCP06064, FM164, LCP06137,
655 LCP06131, LCP06173). The colony size of the challenger was calculated as the colony area in
656 pixels.
657

658 **Statistical tests**

659 We performed analyses of variance (ANOVAs) followed by post-ANOVA Tukey's HSD (honest
660 significant difference) tests, with R. The data normality was improved with the bestNormalize R
661 package [85]. For improving distribution normality, we used an ordered quantile (ORQ)
662 transformation (a one-to-one transformation for vectors with an arbitrary distribution that
663 generates normally distributed vectors) for diameter measurements and competition data and a
664 Yeo-Johnson transformation for CPA (mycotoxin) data. Some residuals were still not normally
665 distributed but ANOVA is known to be robust to deviations from normality [86]. When residues
666 deviated from normality (diameter measurements for the growth and competition), we
667 nevertheless also ran non parametric tests, i.e. Kruskal-Wallis tests, using R.

For ANOVAs, we used standard linear models in which all explanatory variables were discrete. The variables common for all analyses (with explained variables being growth, opacity, CPA and competition, respectively) were 'species', '*P. camemberti* variety' and 'substrate of origin' (food versus non-food). The variables 'medium' and 'temperature/hygrometry' were explanatory variables specific to the growth analysis. The 'challenger species' variable was specific to the competition analysis. All variables and all interactions between them were implemented in the ANOVA, and non-significant interactions were subsequently removed before performing post-ANOVA Tukey's HSD tests.

Acknowledgments

We thank everyone who sent us cheese crusts, Steve Labrie for the LMA strain, and Joëlle Dupont for strains from the National Museum of Natural History in Paris.
This work was supported by the Fungadapt ANR-19-CE20-0002-02 ANR grant to TG, JR, EC, MC and a Fondation Louis D. grant from the Institut de France to TG.

Author contributions

JR and TG designed and supervised the study, and obtained funding. JR, ED, AS and SLP generated the data. JR, RRdlV and BB analyzed the genomes. JR, ED and AS performed the experiments. MC, EC and EP performed the mycotoxin and contributed to CPA biosynthetic gene cluster analyses. JR and ED analyzed the data from the laboratory experiments. JR and TG wrote the manuscript, with contributions from all the authors.

References

1. Warmuth, V., Eriksson, A., Bower, M.A., Cañon, J., Cothran, G., Distl, O., Glowatzki-Mullis, M.-L., Hunt, H., Luís, C., do Mar Oom, M., *et al.* (2011). European Domestic Horses Originated in Two Holocene Refugia. PLoS ONE *6*, e18194.
2. Axelsson, E., Ratnakumar, A., Arendt, M.-L., Maqbool, K., Webster, M.T., Perloski, M., Liberg, O., Arnemo, J.M., Hedhammar, Å., and Lindblad-Toh, K. (2013). The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature *495*, 360–364.
3. Frantz, L.A.F., Schraiber, J.G., Madsen, O., Megens, H.-J., Cagan, A., Bosse, M., Paudel, Y., Crooijmans, R.P.M.A., Larson, G., and Groenen, M.A.M. (2015). Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. Nat. Genet. *47*, 1141–1148.
4. Hufford, M.B., Xu, X., van Heerwaarden, J., Pyhäjärvi, T., Chia, J.-M., Cartwright, R.A., Elshire, R.J., Glaubitz, J.C., Guill, K.E., Kaeppler, S.M., *et al.* (2012). Comparative population genomics of maize domestication and improvement. Nat Genet *44*, 808–811.
5. Decroocq, S., Cornille, A., Tricon, D., Babayeva, S., Chague, A., Eyquard, J.-P.,

Karychev, R., Dolgikh, S., Kostritsyna, T., Liu, S., *et al.* (2016). New insights into the history of domesticated and wild apricots and its contribution to Plum pox virus resistance. Mol. Ecol. *25*, 4712–4729.

6. Doebley, J. (2004). The Genetics of Maize Evolution. Annual Review of Genetics *38*, 37–59.
7. Wydooghe, E., Berghmans, E., Rijsselaere, T., and Soom, A.V. (2013). International breeder inquiry into the reproduction of the English bulldog. Vlaams Diergeneeskundig Tijdschrift *82*, 38–43.
8. Perrier, X., Langhe, E.D., Donohue, M., Lentfer, C., Vrydaghs, L., Bakry, F., Carreel, F., Hippolyte, I., Horry, J.-P., Jenny, C., *et al.* (2011). Multidisciplinary perspectives on banana (*Musa* spp.) domestication. PNAS *108*, 11311–11318.
9. Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. Nature *418*, 700–707.
10. Gladieux, P., Ropars, J., Badouin, H., Branca, A., Aguileta, G., de Vienne, D.M., Rodríguez de la Vega, R.C., Branco, S., and Giraud, T. (2014). Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. Molecular Ecology *23*, 753–773.
11. Dupont, J., Dequin, S., Giraud, T., Le Tacon, F., Marsit, S., Ropars, J., Richard, F., and Selosse, M.-A. (2017). Fungi as a Source of Food. In Microbiology Spectrum.
12. Almeida, P., Gonçalves, C., Teixeira, S., Libkind, D., Bontrager, M., Masneuf-Pomarède, I., Albertin, W., Durrens, P., Sherman, D.J., Marullo, P., *et al.* (2014). A Gondwanan imprint on global diversity and domestication of wine and cider yeast *Saccharomyces uvarum*. Nature Communications *5*.
13. Fay, J.C., and Benavides, J.A. (2005). Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. PLoS Genet *1*, e5.
14. Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., Merlevede, A., Roncoroni, M., Voordeckers, K., Miraglia, L., *et al.* (2016). Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. Cell *166*, 1397-1410.e16.
15. Legras, J.-L., Merdinoglu, D., Cornuet, J.-M., and Karst, F. (2007). Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. Molecular Ecology *16*, 2091–2102.
16. Libkind, D., Hittinger, C.T., Valério, E., Gonçalves, C., Dover, J., Johnston, M., Gonçalves, P., and Sampaio, J.P. (2011). Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. Proceedings of the National Academy of Sciences *108*, 14539–14544.
17. Marsit, S., Mena, A., Bigey, F., Sauvage, F.-X., Couloux, A., Guy, J., Legras, J.-L., Barrio, E., Dequin, S., and Galeote, V. (2015). Evolutionary advantage conferred by an eukaryote-to-eukaryote gene transfer event in wine yeasts. Molecular Biology and Evolution *32*, 1695–1707.
18. Novo, M., Bigey, F., Beyne, E., Galeote, V., Gavory, F., Mallet, S., Cambon, B., Legras, J.-L., Wincker, P., Casaregola, S., *et al.* (2009). Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. PNAS *106*, 16333–16338.
19. Peter, J., Chiara, M.D., Friedrich, A., Yue, J.-X., Pflieger, D., Bergström, A., Sigwalt, A., Barre, B., Freel, K., Llored, A., *et al.* (2018). Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. Nature *556*, 339–344.
20. Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., James, S.A., Davey, R.P., Roberts, I.N., Burt, A., Koufopanou, V., *et al.* (2009). Population genomics of domestic and wild yeasts. Nature *458*, 337–341.

21. Spor, A., Kvitek, D.J., Nidelet, T., Martin, J., Legrand, J., Dillmann, C., Bourgais, A., De Vienne, D., Sherlock, G., and Sicard, D. (2014). Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. Evolution *68*, 772–790.
22. Dunn, B., Paulish, T., Stanbery, A., Piotrowski, J., Koniges, G., Kroll, E., Louis, E.J., Liti, G., Sherlock, G., and Rosenzweig, F. (2013). Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression. PLoS Genetics *9*, e1003366.
23. Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.-J., Wortman, J.R., Batzoglou, S., Lee, S.-I., Baştürkmen, M., Spevak, C.C., Clutterbuck, J., *et al.* (2005). Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature *438*, 1105–1115.
24. Gibbons, J.G., Salichos, L., Slot, J.C., Rinker, D.C., McGary, K.L., King, J.G., Klich, M.A., Tabb, D.L., McDonald, W.H., and Rokas, A. (2012). The evolutionary imprint of domestication on genome variation and function of the filamentous fungus *Aspergillus oryzae*. Current Biology *22*, 1403–1409.
25. Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., Kusumoto, K.-I., Arima, T., Akita, O., Kashiwagi, Y., *et al.* (2005). Genome sequencing and analysis of *Aspergillus oryzae*. Nature *438*, 1157–1161.
26. Dumas, E., Feurtey, A., Vega, R.C.R. de la, Prieur, S.L., Snirc, A., Coton, M., Thierry, A., Coton, E., Piver, M.L., Roueyre, D., *et al.* (2020). Independent domestication events in the blue-cheese fungus *Penicillium roqueforti*. Molecular Ecology *00*, 1–22.
27. Cheeseman, K., Ropars, J., Renault, P., Dupont, J., Gouzy, J., Branca, A., Abraham, A.-L., Ceppi, M., Conseiller, E., Debuchy, R., *et al.* (2014). Multiple recent horizontal transfers of a large genomic region in cheese making fungi. Nat Commun *5*, 2876.
28. Ropars, J., Rodríguez de la Vega, R.C., López-Villavicencio, M., Gouzy, J., Sallet, E., Dumas, É., Lacoste, S., Debuchy, R., Dupont, J., Branca, A., *et al.* (2015). Adaptive horizontal gene transfers between multiple cheese-associated Fungi. Current Biology *25*, 2562–2569.
29. Gillot, G., Jany, J.-L., Coton, M., Le Floch, G., Debaets, S., Ropars, J., López-Villavicencio, M., Dupont, J., Branca, A., Giraud, T., *et al.* (2015). Insights into *Penicillium roqueforti* morphological and genetic diversity. PLOS ONE *10*, e0129849.
30. Legras, J.-L., Galeote, V., Bigey, F., Camarasa, C., Marsit, S., Nidelet, T., Sanchez, I., Couloux, A., Guy, J., Franco-Duarte, R., *et al.* (2018). Adaptation of *S. cerevisiae* to fermented food environments reveals remarkable genome plasticity and the footprints of domestication. Mol Biol Evol *35*, 1712–1727.
31. Pitt, J.I. (1979). The genus Penicillium and its teleomorphic states *Eupenicillium* and *Talaromyces.* (London, UK: Academic Press).
32. Delfosse, C. (2008). Histoires de bries (Saint-Cyr-sur-Morin: Illustria Librairie des Musées).
33. Pitt, J.I., Cruickshank, R.H., and Leistner, L. (1986). *Penicillium commune*, *P. camembertii*, the origin of white cheese moulds, and the production of cyclopiazonic acid. Food Microbiology *3*, 363–371.
34. Pouriau, A.-F.A. du texte (1895). La laiterie : art de traiter le lait, de fabriquer le beurre et les principaux fromages français et étrangers (5e éd.) / par A.-F. Pouriau.
35. Garnier, L., Valence, F., Pawtowski, A., Auhustsinava-Galerne, L., Frotté, N., Baroncelli, R., Deniel, F., Coton, E., and Mounier, J. (2017). Diversity of spoilage fungi associated with various French dairy products. International Journal of Food Microbiology *241*, 191–197.
36. Giraud, F., Giraud, T., Aguileta, G., Fournier, E., Samson, R., Cruaud, C., Lacoste, S.,

Ropars, J., Tellier, A., and Dupont, J. (2010). Microsatellite loci to recognize species for the cheese starter and contaminating strains associated with cheese manufacturing. International Journal of Food Microbiology *137*, 204–213.

37. Visagie, C.M., Houbraken, J., Frisvad, J.C., Hong, S.-B., Klaassen, C.H.W., Perrone, G., Seifert, K.A., Varga, J., Yaguchi, T., and Samson, R.A. (2014). Identification and nomenclature of the genus *Penicillium*. Studies in Mycology *78*, 343–371.
38. Hymery, N., Vasseur, V., Coton, M., Mounier, J., Jany, J.-L., Barbier, G., and Coton, E. (2014). Filamentous Fungi and Mycotoxins in Cheese: A Review. Comprehensive Reviews in Food Science and Food Safety *13*, 437–456.
39. Le Bars, J. (1979). Cyclopiazonic acid production by Penicillium camemberti Thom and natural occurrence of this mycotoxin in cheese. Appl Environ Microbiol *38*, 1052–1055.
40. Hymery, N., Masson, F., Barbier, G., and Coton, E. (2014). Cytotoxicity and immunotoxicity of cyclopiazonic acid on human cells. Toxicol In Vitro *28*, 940–947.
41. Bodinaku, I., Shaffer, J., Connors, A.B., Steenwyk, J.L., Biango-Daniels, M.N., Kastman, E.K., Rokas, A., Robbat, A., and Wolfe, B.E. (2019). Rapid Phenotypic and Metabolomic Domestication of Wild Penicillium Molds on Cheese. mBio *10*, e02445-19.
42. Gibbons, J.G., and Rokas, A. (2013). The function and evolution of the Aspergillus genome. Trends in Microbiology *21*, 14–22.
43. Dyer, P.S., and O'Gorman, C.M. (2012). Sexual development and cryptic sexuality in fungi: insights from Aspergillus species. FEMS Microbiol. Rev. *36*, 165–192.
44. Barve, M.P., Arie, T., Salimath, S.S., Muehlbauer, F.J., and Peever, T.L. (2003). Cloning and characterization of the mating type (MAT) locus from Ascochyta rabiei (teleomorph: Didymella rabiei) and a MAT phylogeny of legume-associated Ascochyta spp. Fungal Genetics and Biology *39*, 151–167.
45. Boisard, P. (2007). Le camembert, mythe français Odile Jacob.
46. Selker, E.U., Cambareri, E.B., Jensen, B.C., and Haack, K.R. (1987). Rearrangement of duplicated DNA in specialized cells of Neurospora. Cell *51*, 741–752.
47. Hashem, K.M., He, F.J., Jenner, K.H., and MacGregor, G.A. (2014). Cross-sectional survey of salt content in cheese: a major contributor to salt intake in the UK. BMJ Open *4*, e005051–e005051.
48. Kritikos, G., Banzhaf, M., Herrera-Dominguez, L., Koumoutsi, A., Wartel, M., Zietek, M., and Typas, A. (2017). A tool named Iris for versatile high-throughput phenotyping in microorganisms. Nature Microbiology *2*, 17014.
49. Lund, F., Filtenborg, O., and Frisvad, J.C. (1998). *Penicillium caseifulvum*, a new species found on Penicillium roqueforti fermented cheeses. Journal of Food Mycology *1*, 95–101.
50. Barbesgaard, P., Heldt-Hansen, H.P., and Diderichsen, B. (1992). On the safety of Aspergillus oryzae: a review. Appl Microbiol Biotechnol *36*, 569–572.
51. Gillot, G., Jany, J.-L., Poirier, E., Maillard, M.-B., Debaets, S., Thierry, A., Coton, E., and Coton, M. (2017). Functional diversity within the *Penicillium roqueforti* species. Int. J. Food Microbiol. *241*, 141–150.
52. Giraud, T., Refrégier, G., Le Gac, M., de Vienne, D.M., and Hood, M.E. (2008). Speciation in fungi. Fungal Genetics and Biology *45*, 791–802.
53. Le Gac, M., and Giraud, T. (2008). Existence of a pattern of reproductive character displacement in Homobasidiomycota but not in Ascomycota. Journal of Evolutionary Biology *21*, 761–772.
54. Zhu, Q., Zheng, X., Luo, J., Gaut, B.S., and Ge, S. (2007). Multilocus analysis of nucleotide variation of Oryza sativa and its wild relatives: severe bottleneck during domestication of rice. Mol. Biol. Evol. *24*, 875–888.
55. Marsden, C.D., Ortega-Del Vecchyo, D., O'Brien, D.P., Taylor, J.F., Ramirez, O., Vilà,

855 C., Marques-Bonet, T., Schnabel, R.D., Wayne, R.K., and Lohmueller, K.E. (2015).
856 Bottlenecks and selective sweeps during domestication have increased deleterious genetic
857 variation in dogs. Proceedings of the National Academy of Sciences, 201512501.
858 56. Eyre-Walker, A., Gaut, R.L., Hilton, H., Feldman, D.L., and Gaut, B.S. (1998).
859 Investigation of the bottleneck leading to the domestication of maize. PNAS *95*, 4441–
860 4446.
861 57. Gaut, B.S., Díez, C.M., and Morrell, P.L. (2015). Genomics and the Contrasting
862 Dynamics of Annual and Perennial Domestication. Trends in Genetics *31*, 709–719.
863 58. Duan, S.-F., Han, P.-J., Wang, Q.-M., Liu, W.-Q., Shi, J.-Y., Li, K., Zhang, X.-L., and
864 Bai, F.-Y. (2018). The origin and adaptive evolution of domesticated populations of yeast
865 from Far East Asia. Nat Commun *9*.
866 59. Mahler, D.L., Ingram, T., Revell, L.J., and Losos, J.B. (2013). Exceptional convergence
867 on the macroevolutionary landscape in island lizard radiations. Science *341*, 292–295.
868 60. Thorpe, R.S., Barlow, A., Malhotra, A., and Surget-Groba, Y. (2015). Widespread
869 parallel population adaptation to climate variation across a radiation: implications for
870 adaptation to climate change. Mol Ecol *24*, 1019–1030.
871 61. Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A., and Cresko, W.A.
872 (2010). Population Genomics of Parallel Adaptation in Threespine Stickleback using
873 Sequenced RAD Tags. PLOS Genetics *6*, e1000862.
874 62. Duboué, E.R., Keene, A.C., and Borowsky, R.L. (2011). Evolutionary convergence on
875 sleep loss in cavefish populations. Current Biology *21*, 671–676.
876 63. Salzburger, W. (2009). The interaction of sexually and naturally selected traits in the
877 adaptive radiations of cichlid fishes. Molecular Ecology *18*, 169–185.
878 64. Elmer, K.R., Fan, S., Kusche, H., Spreitzer, M.L., Kautt, A.F., Franchini, P., and Meyer,
879 A. (2014). Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel
880 routes. Nature Communications *5*, 5168.
881 65. Elmer, K.R., Kusche, H., Lehtonen, T.K., and Meyer, A. (2010). Local variation and
882 parallel evolution: morphological and genetic diversity across a species complex of
883 neotropical crater lake cichlid fishes. Philos Trans R Soc Lond B Biol Sci *365*, 1763–
884 1782.
885 66. Martínez-Ainsworth, N.E., and Tenaillon, M.I. (2016). Superheroes and masterminds of
886 plant domestication. Comptes Rendus Biologies *339*, 268–273.
887 67. Gross, B.L., and Olsen, K.M. (2010). Genetic perspectives on crop domestication. Trends
888 in Plant Science *15*, 529–537.
889 68. Decker, M., and Nielsen, P.V. (2005). The inhibitory effect of *Penicillium camemberti*
890 and *Geotrichum candidum* on the associated funga of white mould cheese. International
891 Journal of Food Microbiology *104*, 51–60.
892 69. Gouyon, P.-H., Leriche, H., Civard, A., Reeves, H., and Hulot, N. (2010). Aux origines
893 de l'environnement (Fayard).
894 70. Vavilov, N.I., Vavylov, M.I., Vavílov, N.Í., and Dorofeev, V.F. (1992). Origin and
895 Geography of Cultivated Plants (Cambridge University Press).
896 71. Glass, N.L., and Donaldson, G.C. (1995). Development of primer sets designed for use
897 with the PCR to amplify conserved genes from filamentous ascomycetes. Appl Environ
898 Microbiol *61*, 1323–1330.
899 72. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M.
900 (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting
901 and repeat separation. Genome Res. *27*, 722–736.
902 73. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo,
903 C.A., Zeng, Q., Wortman, J., Young, S.K., *et al.* (2014). Pilon: An Integrated Tool for
904 Comprehensive Microbial Variant Detection and Genome Assembly Improvement.

905 PLOS ONE *9*, e112963.
906 74. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for
907 Illumina sequence data. Bioinformatics *30*, 2114–2120.
908 75. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2.
909 Nat. Methods *9*, 357–359.
910 76. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A.,
911 Garimella, K., Altshuler, D., Gabriel, S., Daly, M., *et al.* (2010). The Genome Analysis
912 Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.
913 Genome Res. *20*, 1297–1303.
914 77. Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E., and Lercher, M.J. (2014).
915 PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R.
916 Mol Biol Evol *31*, 1929–1936.
917 78. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next
918 Generation Sequencing Data. BMC Bioinformatics *15*, 356.
919 79. Schliep, K.P. (2011). phangorn: phylogenetic analysis in R. Bioinformatics *27*, 592–593.
920 80. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von
921 Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods
922 for Phylogenetic Inference in the Genomic Era. Mol Biol Evol *37*, 1530–1534.
923 81. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S.
924 (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nature
925 Methods *14*, 587–589.
926 82. Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C., and Gascuel, O. (2011). Survey
927 of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast
928 Likelihood-based Approximation Schemes. Syst Biol *60*, 685–699.
929 83. Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013). Ultrafast approximation for
930 phylogenetic bootstrap. Mol. Biol. Evol. *30*, 1188–1195.
931 84. Hill, T.W., and Kafer, E. (2001). Improved protocols for Aspergillus minimal medium:
932 trace element and minimal medium salt stock solutions. Fungal Genetics Reports *48*, 20–
933 21.
934 85. Peterson, R.A., and Cavanaugh, J.E. (2019). Ordered quantile normalization: a
935 semiparametric transformation built for the cross-validation era. Journal of Applied
936 Statistics *0*, 1–16.
937 86. Porcher, E., Giraud, T., Goldringer, I., and Lavigne, C. (2004). Experimental
938 demonstration of a causal relationship between heterogeneity of selection and genetic
939 differentiation in quantitative traits. Evolution *58*, 1434–1445.
940

941 # Figure Legends

942 **Figure 1: Cheeses and fungi used for cheese-making.**
943 A) Schematic representation of relationships between *Penicillium camemberti* and its closest
944 relatives, *P. biforme* and *P. fuscoglaucum*. Pictures of *P. biforme, P. camemberti* var. *camemberti* and *P.*
945 *camemberti* var. *caseifulvum* on salted cheese medium at 25°C after 10 days of growth. B) Petri dish
946 with a white and fluffy *P. camemberti* var. *camemberti* strain (FM013) and a green rough *P.*
947 *fuscoglaucum* strain (LCP06617) on malt medium at 25°C.
948

949 **Figure 2: Population structure of 61 strains from the *Penicillium camemberti* species**
950 **complex, based on whole-genome data.**
951 A) Principal component analysis (PCA) based on the 61 strains, B) Principal component analysis
952 (PCA) based on only *P. camemberti* strains (n = 19). C) Neighbor-net analysis based on SNP data.
953 Cross-linking indicates the likely occurrence of recombination. Branch lengths are shown and the
954 scale bar represents 0.07 substitutions per site. D) Population subdivision inferred from $K = 2$ to

955 *K* = 6. Each column represents a strain and colored bars represent their coefficients of
956 membership for the various gene pools. For all panels, genetic clusters are represented by the
957 same colors: dark purple for *P. fuscoglaucum,* light blue for *P. biforme,* dark green for *P. camemberti*
958 var. *camemberti* and light green for *P. camemberti* var. *caseifulvum.* The strains identified as *P. commune*
959 (non-monophyletic group) by collectors are indicated in red. For PCA and neighbor-net analysis,
960 symbols correspond to the environment of collection: circles for cheese, triangles for dried
961 sausages and squares for other environments. On the structure analysis, letters below the plots
962 indicate the origin of the strains: C for cheese, F for food strains other than cheese and E for
963 other environments. Red asterisks indicated the type strains of *P. biforme, P. camemberti and P.*
964 *caseifulvum.*
965 See also Table S1, Figures S1, S2, Data S5 and Supplemental Information.
966
967 **Figure 3: Phylogenetic relationships and presence/absence of the *Wallaby* and**
968 ***CheesyTer* horizontally transferred regions in the 61 studied strains of the *Penicillium***
969 ***camemberti* species complex and the three outgroups.**
970 A) Maximum likelihood tree for the 64 strains based on 1,124,983 SNPs (586,131 parsimony
971 informative sites). Branch supports, represented as SH-aLRT/ultrafast bootstrap in percentage,
972 are only shown for clades but were always > 75%. The strain name color indicates the
973 environment of collection, black for food and blue for other environments. Asterisks after strains
974 indicate strains previously named *P. commune.*) Mapping coverage of the *CheesyTer* region, as
975 defined in the reference genome *P. camemberti* var. *camemberti* FM013. C) Mapping coverage of the
976 *Wallaby* region, as defined in the reference genome *P. camemberti* var. *camemberti* FM013. On panels
977 B and C, no read mapping in a given position is represented in white, grey showing regions with
978 at least a read mapped.
979 See also Table S1, Figures S1, S2, Data S3, S4 and S5, Supplemental Information.
980
981 **Figure 4: Phenotypic traits distinguishing the four lineages *Penicillium biforme, P.***
982 ***camemberti* var. *caseifulvum, P. camemberti* var. *camemberti* and *P. fuscoglaucum.*** A)
983 Mean radial growth of the four lineages on unsalted cheese, salted cheese, malt and minimal
984 media. B) Pictures of colonies on unsalted cheese, salted cheese, malt and minimal media; *P.*
985 *fuscoglaucum*: LCP05471; *P. camemberti* var. *camemberti*: UBOCC-A-113052; *P. biforme*: ESE00211; *P.*
986 *camemberti* var. *caseifulvum*: LCP06622. C) Difference in opacity between the four lineages on salted
987 cheese at 25°C. D) Difference in production of cyclopiazonic acid between the four lineages on
988 yeast extract sucrose. C-D) Horizontal lines of the boxplots represent the upper quartile, the
989 median and the lower quartile. Each dot represents a strain and p-values are given for the tests of
990 differences between the lineages (Wilcoxon test). The color indicates assignments to the four
991 lineages as in other figures.
992 See also Data S1, S2, Figure S3 and Supplemental Information.
993
994 **Figure 5: Competitive abilities of *Penicillium biforme, P. camemberti* var. *caseifulvum***
995 **and *P. camemberti* var. *camemberti* against *Geotrichum candidum, P. biforme, P.***
996 ***fuscoglaucum* and *P. roqueforti* competitors.** Top: Pictures of *P. roqueforti* FM164, *Geotrichum*
997 *candidum* ESE00163, *P. biforme* ESE00154 and *P. fuscoglaucum* ESE00221 on no lawn (first line), *P.*
998 *biforme* ESE00086 lawn (second line), *P. camemberti* var. *camemberti* LCP04810 (third line) and *P.*
999 *camemberti* var. *caseifulvum* ESE00019 (fourth line) lawns, on salted cheese. Bottom: Boxplots
1000 representing the differences of growth abilities of the competitors on different species lawn.
1001 Horizontal lines of the boxplots represent the upper quartile, the median and the lower quartile.
1002 Each dot represents a strain and p-values are given for the tests of differences between the
1003 lineages (Wilcoxon test). The color indicates assignments to the three lineages as in other figures.
1004 Abbreviations: *P. camemberti* var. *camemberti:* var. *cam; P. camemberti* var. *caseifulvum:* var. *casei.*
1005 See also Data S1, S2, supplemental Information.

1006 # Supplemental Data Legends

1007

1008 **Table S1:** Population genetics statistics in the *P. fuscoglaucum, P. biforme, P. camemberti* var.
1009 *camemberti* and var. *caseifulvum* populations. F_{ST} and d_{xy} values were calculated on pairwise
1010 comparisons between populations, as well as proportion of shared, private and fixed SNPs.
1011 See also Figures 2 and 3.

1012

1013 **Figure S1: Population structure of 61 strains from the *Penicillium camemberti* cheese-**
1014 **making fungal species complex, based on whole-genome data with the addition of the**
1015 **genome of the *P. commune* strain published in [41].**
1016 A) Principal component analysis; B) Neighbor-net analysis. Cross-linking indicates the likely
1017 occurrence of recombination Branch lengths are shown and the scale bar represents 0.04
1018 substitutions per site. For both panels, the genetic clusters are represented by the same colors:
1019 light blue for *P. biforme,* dark green for *P. camemberti* var. *camemberti* and light green for *P. camemberti*
1020 var. *caseifulvum.* The strain added in this analysis, 162 3FA WT8, identified as *P. commune* (non-
1021 monophyletic group) is shown in red. The symbols correspond to the environment of collection:
1022 circles for cheese, triangles for dried sausages and squares for other environments.
1023 See also Figures 2 and 3A.

1024

1025 **Figure S2: Population structure of 47 strains from the two *Penicillium camemberti***
1026 **varieties and *P. biforme* lineages, based on whole-genome data.**
1027 A) Principal component analysis; B) Neighbor-net analysis. Reticulation indicates the likely
1028 occurrence of recombination Branch lengths are shown and the scale bar represents 0.01
1029 substitutions per site. For both panels, the genetic clusters are represented by the same colors:
1030 light blue for *P. biforme,* dark green for *P. camemberti* var. *camemberti* and light green for *P. camemberti*
1031 var. *caseifulvum.* The strains identified as *P. commune* (non-monophyletic group) by collectors are
1032 indicated in red. The symbols correspond to the environment of collection: circles for cheese,
1033 triangles for dried sausages and squares for other environments.
1034 See also Figures 2 and 3A.

1035

1036 **Figure S3: Cyclopiazonic acid (CPA) biosynthesis cluster in the *Penicillium camemberti***
1037 **species complex.** A) Schematic representation of the CPA biosynthesis cluster in the *P.*
1038 *camemberti* reference genome FM013. B) Amino-acid alignment of the *CpaA* gene showing a 2bp
1039 deletion in strains belonging to the *P. camemberti* var. *caseifulvum* genetic cluster, leading to a
1040 truncated protein.
1041 See also Figure 4D.

1042

1043 **Supplemental Information:** Description of the origin, donor and genome coverage of the 61
1044 strains of the *Penicillium camemberti* species complex used in this study.
1045 See also Figures 2 - 5.

1046

1047 **Data S1:** Data for measured traits in *Penicillium camemberti* var. *camemberti, P. camemberti* var.
1048 *caseifulvum, P. biforme* and *P. fuscoglaucum.* A tab: growth; B tab: opacity; C tab: Cyclopiazonic acid
1049 production; D tab: competition.

1050

1051 **Data S2:** Results of statistical analyses performed for testing differences in traits between
1052 *Penicillium camemberti* var. *camemberti, P. camemberti* var. *caseifulvum, P. biforme* and *P. fuscoglaucum.*
1053 Significant P-values are indicated with an asterisk and highlighted in gray. A tab: growth; B tab:
1054 opacity; C tab: Cyclopiazonic acid production; D tab: competition.
1055 See also Figures 4-5.

Data S3: Genomic scans along the first 20 scaffolds of d_{xy}, F_{ST}, π and Watterson's θ. Dots show windows in the 1% highest d_{xy} between *Penicillium biforme* and *P. caememberti* var. *caseifulvum*, var. *camemberti* and *P. fuscoglaucum*.
See also Figures 2 and 3A, Data S4.

Data S4: Gene annotation within windows in the 1% highest d_{xy} between *Penicillium biforme* and *P. camemberti* var. *caseifulvum* (A tab), var. *camemberti* (B tab) and *P. fuscoglaucum* (C tab).
See also Figures 2 and 3A, Data S3.

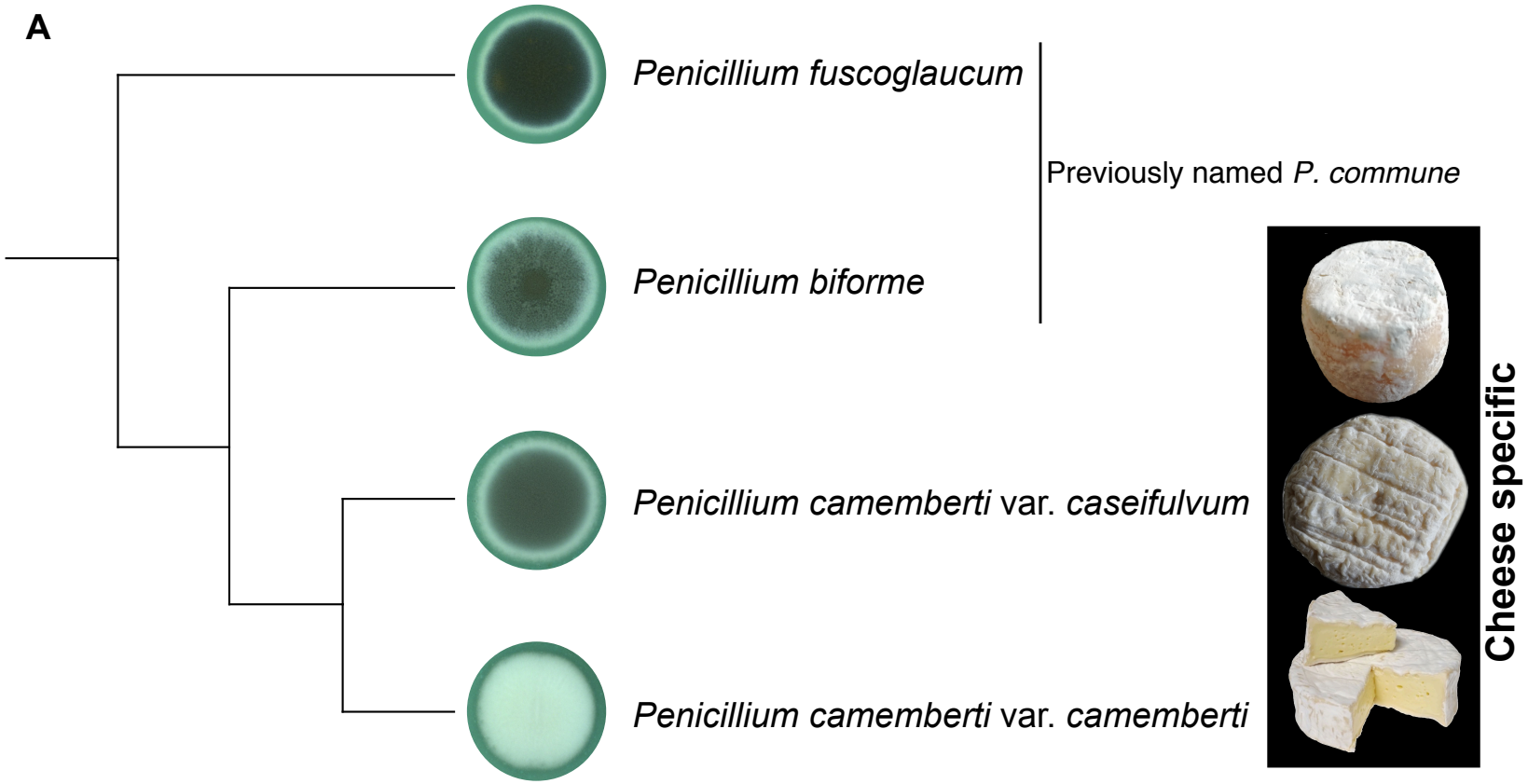
Data S5: Oligonucleotide primer sets, PCR reaction, thermal cycler program for amplification, electrophoresis gel and Sanger sequences (PCR products were purified and sequenced by GENEWIZ) of *P. camemberti* var. *camemberti* and var. *caseifulvum*.
See also Figures 2 and 3A.

Materials availability

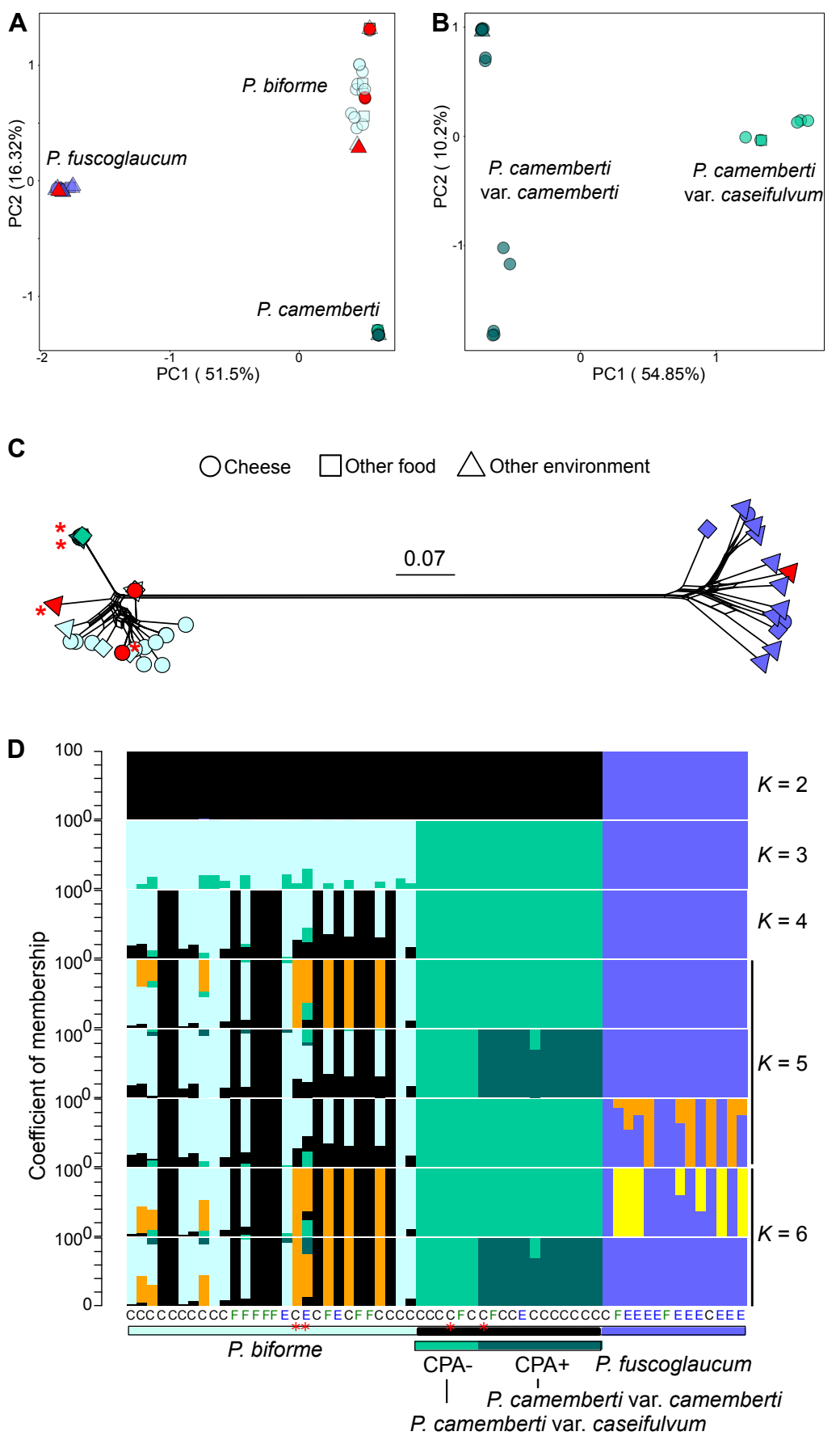
Strains used in this study are publicly available: strains with LCP identifiers are from the public collection of the National Museum of Natural History in Paris; strains with UBOCC identifiers are from the public collection of the Université de Bretagne Occidentale Culture Collection; other strains with ESE numbers are conserved at the Ecology, Systematics and Evolution laboratory in Orsay.

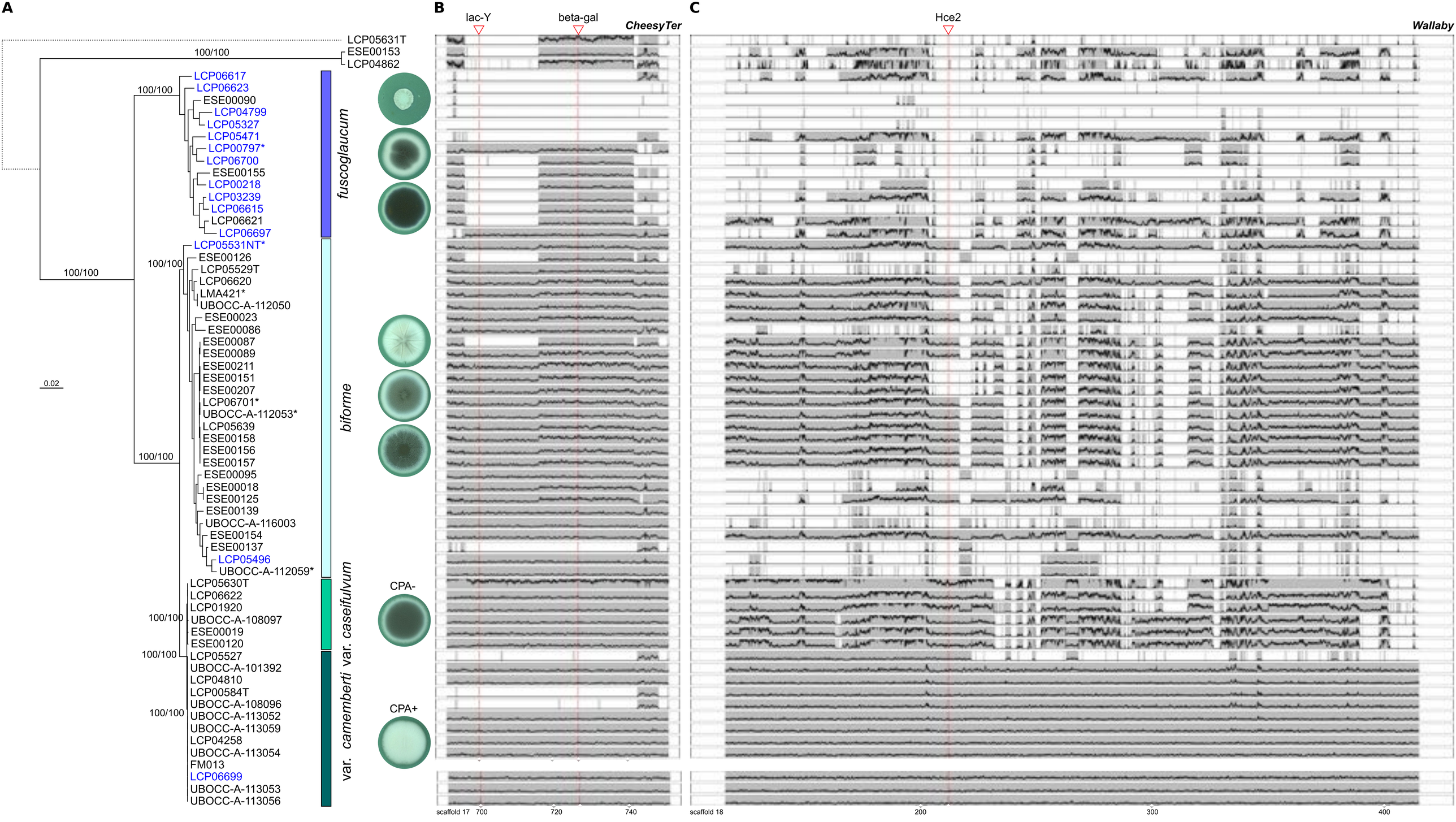
Data and code availability

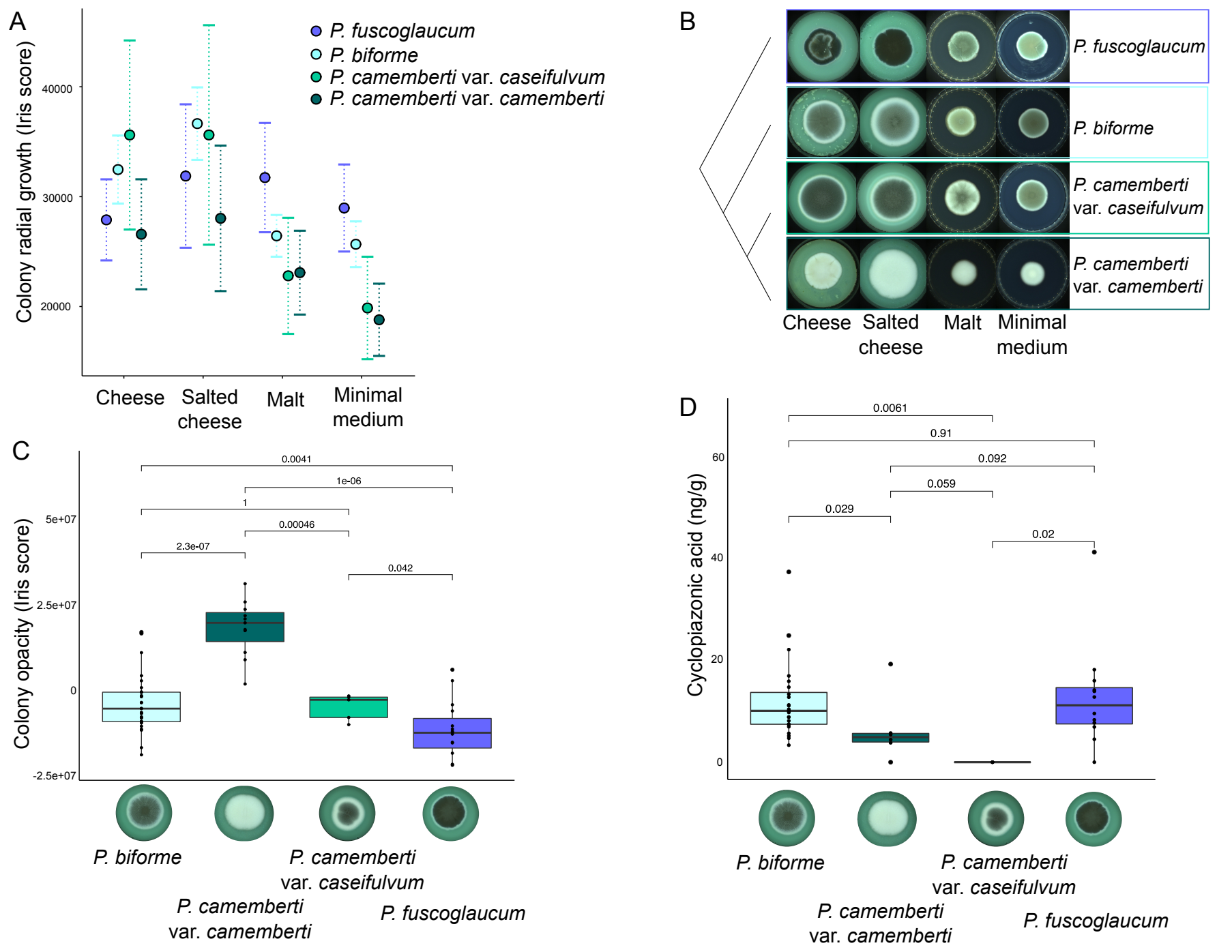
Raw reads have been deposited to NCBI under the project number XXX.
The codes generated during this study are available at XXX.

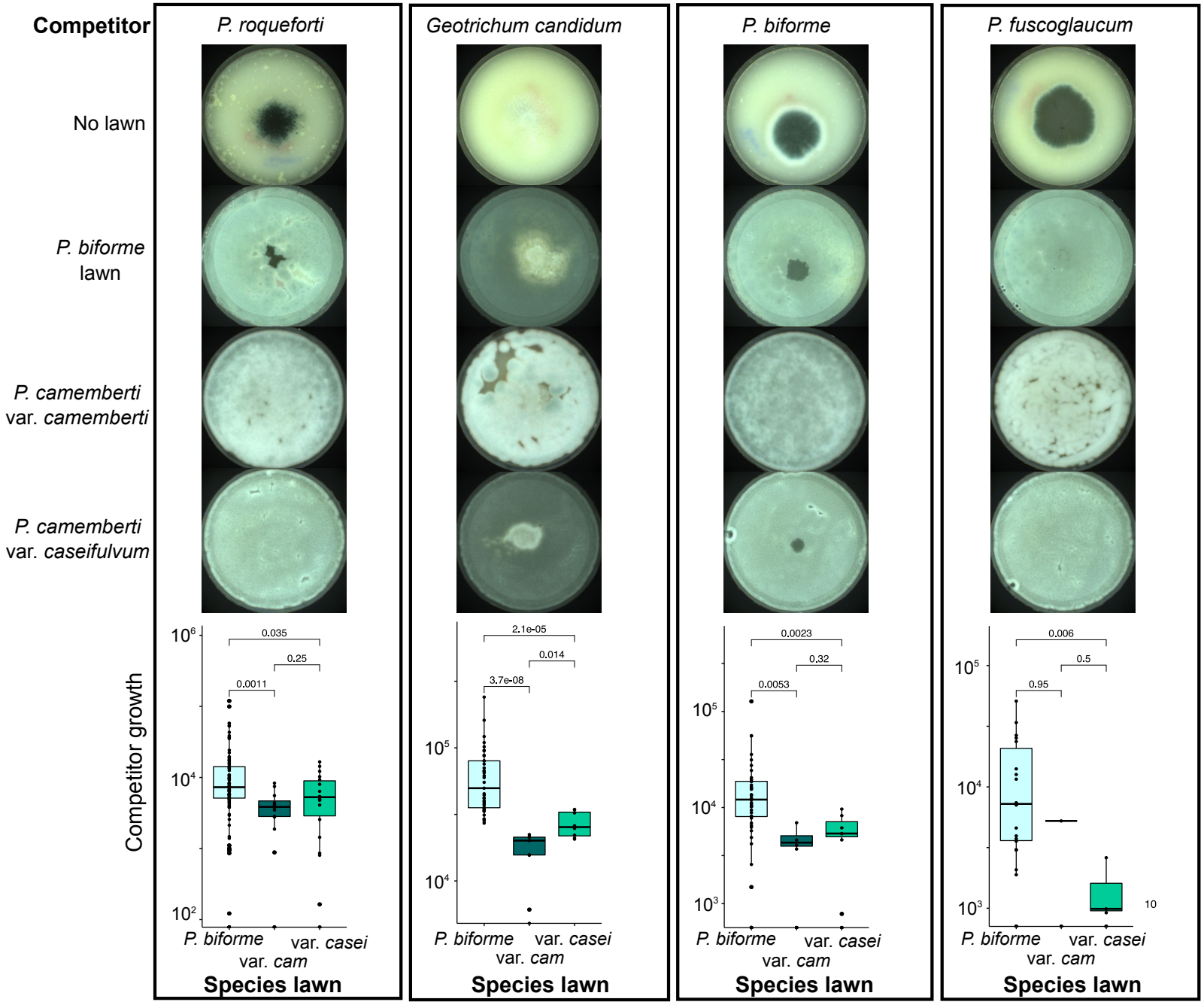

A
Penicillium fuscoglaucum
Previously named P. commune
Penicillium biforme
Penicillium camemberti var. caseifulvum
Penicillium camemberti var. camemberti
Cheese specific

B


A
P. biforme
P. fuscoglaucum
P. camemberti
PC2 (16.32%)
PC1 (51.5%)
B
P. camemberti
var. camemberti
P. camemberti
var. caseifulvum
PC2 (10.2%)
PC1 (54.85%)
C
Cheese
Other food
Other environment
0.07
D
Coefficient of membership
K = 2
K = 3
K = 4
K = 5
K = 6
P. biforme
CPA-
CPA+
P. fuscoglaucum
P. camemberti var. camemberti
P. camemberti var. caseifulvum


A
B
C
LCP05631T
ESE00153
LCP04862
100/100
LCP06617
LCP06623
ESE00090
LCP04799
LCP05327
LCP05471
LCP00797*
LCP06700
ESE00155
LCP00218
LCP03239
LCP06615
LCP06621
LCP06697
fuscoglaucum
LCP05531NT*
ESE00126
LCP05529T
LCP06620
LMA421*
UBOCC-A-112050
ESE00023
ESE00086
ESE00087
ESE00089
ESE00211
ESE00151
ESE00207
LCP06701*
UBOCC-A-112053*
LCP05639
ESE00158
ESE00156
ESE00157
ESE00095
ESE00018
ESE00125
ESE00139
UBOCC-A-116003
ESE00154
ESE00137
LCP05496
UBOCC-A-112059*
biforme
0.02
LCP05630T
LCP06622
LCP01920
UBOCC-A-108097
ESE00019
ESE00120
var. caseifulvum
LCP05527
UBOCC-A-101392
LCP04810
LCP00584T
UBOCC-A-108096
UBOCC-A-113052
UBOCC-A-113059
LCP04258
UBOCC-A-113054
FM013
LCP06699
UBOCC-A-113053
UBOCC-A-113056
var. camemberti
CPA-
CPA+
lac-Y
beta-gal
CheesyTer
Hce2
Wallaby
scaffold 17
700
720
740
scaffold 18
200
300
400


A
P. fuscoglaucum
P. biforme
P. camemberti var. caseifulvum
P. camemberti var. camemberti
Colony radial growth (Iris score)
40000
30000
20000
Cheese
Salted cheese
Malt
Minimal medium
B
P. fuscoglaucum
P. biforme
P. camemberti var. caseifulvum
P. camemberti var. camemberti
Cheese
Salted cheese
Malt
Minimal medium
C
Colony opacity (Iris score)
5e+07
2.5e+07
0
-2.5e+07
0.0041
1e-06
1
0.00046
2.3e-07
0.042
P. biforme
P. camemberti var. caseifulvum
P. camemberti var. camemberti
P. fuscoglaucum
D
Cyclopiazonic acid (ng/g)
60
40
20
0
0.0061
0.91
0.092
0.059
0.029
0.02
P. biforme
P. camemberti var. caseifulvum
P. camemberti var. camemberti
P. fuscoglaucum


Competitor
P. roqueforti
Geotrichum candidum
P. biforme
P. fuscoglaucum
No lawn
P. biforme lawn
P. camemberti var. camemberti
P. camemberti var. caseifulvum
Competitor growth
0.035
0.25
0.0011
2.1e-05
0.014
3.7e-08
0.0023
0.32
0.0053
0.006
0.5
0.95
P. biforme var. cam
var. casei
Species lawn