



**HAL**  
open science

# Domestication des champignons du fromage

Bastien Bennetot

► **To cite this version:**

Bastien Bennetot. Domestication des champignons du fromage. Evolution [q-bio.PE]. Université Paris-Saclay, 2022. Français. NNT : 2022UPASB066 . tel-04027173

**HAL Id: tel-04027173**

**<https://theses.hal.science/tel-04027173>**

Submitted on 13 Mar 2023

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

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

# Domestication des champignons du fromage

*Domestication of cheese fungi*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n°567 : sciences du végétal : du gène à l'écosystème  
(SEVE)

Spécialité de doctorat : Biologie

Graduate School : BioSphERA - Biologie, Société, Ecologie & Environnement,  
Ressources, Agriculture & Alimentation. Référent : Faculté des Sciences d'Orsay

Thèse préparée dans l'unité de recherche **Ecologie Systématique et Evolution**  
(Université Paris-Saclay, CNRS, AgroParisTech), sous la direction  
de **Jeanne ROPARS**, Chargée de recherche CNRS-Université Paris-Saclay

Thèse soutenue à Paris-Saclay, le 25 novembre 2022,  
par

**Bastien BENNETOT**

### Composition du Jury

Membres du jury avec voix délibérative

<b>Cécile FAIRHEAD</b> Professeure, Université Paris-Saclay	Présidente	
<b>Warren ALBERTIN</b> Maîtresse de conférence (HDR), École Nationale Supérieure de Chimie, de Biologie et de Physique, Bordeaux INP	Rapporteuse Examinatrice	&
<b>Joseph SCHACHERER</b> Professeur, Université de Strasbourg	Rapporteur Examineur	&
<b>Elisabeth FOURNIER</b> Directrice de recherche, INRAE - Université de Montpellier	Examinatrice	
<b>Jean-Luc LEGRAS</b> Ingénieur de recherche, INRAE - Université de Montpellier	Examineur	

**Titre : Domestication des champignons du fromage**

**Mots clés :** Évolution, Domestication, Fromage, Champignon, Génomique, Adaptation

**Résumé :** Durant ma thèse, j'ai étudié la domestication chez deux champignons inoculés dans le fromage, *Penicillium camemberti* et *Geotrichum candidum*. *Penicillium camemberti* est la moisissure blanche et cotonneuse des fromages à pâte molle type camembert; *Geotrichum candidum* est présent naturellement dans le lait cru, et est aussi utilisé comme ferment dans de nombreux fromages. Nous avons montré que *G. candidum* et *P. camemberti* étaient issus de plusieurs événements de domestication successifs.

Grâce à la comparaison de génomes de souches du fromage et sauvages et d'expériences au laboratoire, nous avons mis en évidence plusieurs variétés de *P. camemberti* et *G. candidum* aux phénotypes contrastés, en termes de croissance, couleur, production de mycotoxines et d'inhibition de contaminants. La comparaison de plusieurs lignées phylogénétiquement distinctes a révélé des convergences phénotypiques entre ces deux champignons du fromage telle que la croissance plus dense verticalement.

**Title :** Domestication of cheese fungi

**Keywords :** Evolution, Domestication, Cheese, Fungi, Genomic, Adaptation

**Abstract :** During my thesis, I studied the domestication of two inoculated cheese fungi, *Penicillium camemberti* and *Geotrichum candidum*. *Penicillium camemberti* is the white, fluffy mold inoculated on soft cheeses such as camembert. *Geotrichum candidum* occurs naturally in raw milk, but is also used as a starter in many kinds of cheeses. We have shown that the domestication of *G. candidum* and *P. camemberti* occurred in several successive steps.

Comparing the genomes of cheese and wild strains and using laboratory experiments, we identified several varieties of *P. camemberti* and *G. candidum* with contrasting phenotypes, in terms of growth, color, mycotoxin production and food spoiler inhibition. The comparison of two phylogenetically distinct cheese lineages revealed phenotypic convergence, such as denser and fluffier growth.

# Remerciements

Tout d'abord, je tiens à remercier le ministère de l'agriculture, l'université Paris Saclay, l'ANRT (Association Nationale Recherche Technologie), le CNRS pour m'avoir permis de travailler sur cette thèse. Je remercie également les membres de mon jury de thèse, Warren Albertin, Joseph Schacherer, Cécile Fairhead, Elisabeth Fournier et Jean-Luc Legras d'avoir accepté d'analyser mes travaux de recherches. Je souhaite aussi remercier les membres de mes comités de thèse, Jacqui Shykoff, Sophie Landaud-Liautaud, Pierre Gladieux, Eric Dugat-Bony, Tatiana Giraud et Steve Labrie dont les précieux conseils m'ont guidé vers de nouvelles idées et parfois remis sur le bon chemin.

À mon encadrante préférée de thèse, Jeanne Ropars, un grand merci pour ces trois ans d'encadrement. Je me souviens encore de notre appel, deux jours avant la fin des candidatures, nous nous sommes tout de suite entendus. Je n'ai pas été tout le temps facile (euphémisme), parfois entêté, mais j'ai beaucoup apprécié notre duo qui fonctionnait très bien. J'espère ne pas lui avoir trop enlevé d'espérance de vie après les maintes relectures de phrases n'ayant aucun sens et truffées de fautes. J'ai trouvé en elle, un modèle de rigueur, de connaissances larges et de dynamisme

dont je suis encore loin. Une expérience enrichissante aussi bien sur le plan scientifique que personnel.

Je me dois aussi de remercier, ce que certains appellent la Ropars team, Jean-Philippe Vernadet et Samuel O'Donnell pour leur aide et leur dépannage informatique. Un immense merci à Jean-Philippe qui m'a évité de rester à l'âge de pierre sur ce qui est l'analyse de données génomiques. Si j'ai pu bosser à fond sur les analyses, c'est parce qu'il était toujours derrière moi pour me transmettre le bon fichier.

En 2019, j'emménageais avec mon colloc de bureau et je n'en suis jamais reparti. Qu'aurait été mon bureau sans Thibault Caron, qui à ses heures perdues marmonnait quelques malédictions envers ses divers travaux. J'ai trouvé en lui une oreille attentive et une validation de tous mes projets farfelus ! Sans lui, mon expérience ici aurait été bien différente, bien moins agréable. Une petite pensée à la team pizza/croziflette comptant Marine Duhamel et Elizabeth Chahine. Merci à Alodie qui ne m'a pas assassiné pour avoir mis le bazar au laboratoire. Je dois d'ailleurs lui confesser ici ne pas avoir signé sa charte de labo. Merci à Jacqui pour avoir été à la fois un soutien pour ce qui est de la méthodologie scientifique et une source de créativité dans le survivalisme/DIY (do it yourself). Merci à Lili pour le superbe manguier de mon bureau. Merci à Ricardo pour nos discussions

intéressantes et sa technicité sans faille. Merci à Tatiana d'être une figure inspirante du féminisme dans un milieu malheureusement encore trop masculin. Merci à Xavier qui a rempli avec beaucoup de dévotion sa mission de mentor. Merci à tous les stagiaires que j'ai encadrés, au moins en partie, et qui m'ont permis d'accomplir des tâches bien fastidieuses Sophie Hautefeuille, Thomas Mari, Laura Prugneau, Jérémy Raynaud, Véronique Bougie et Jules Larouche.

Je remercie Bouygues pour la qualité de leurs efforts, l'innovation et la communication dont ils ont fait preuve. Quel défi de savoir construire un bâtiment neuf avec plus de défaut qu'un ancien bâtiment datant de la guerre froide. Mon seul regret étant qu'ils n'aient pas pensé à rajouter un formulaire de remerciement.

Merci au Covid-19 pour m'avoir donné deux beaux mois de télétravail et pour les multiples surprises de parcours. J'aurais presque réussi à rester vierge nasalement, mais cet été 2022 a eu raison de moi.

Je remercie enfin tous ceux qui m'ont poussé jusqu'ici et dont le récit serait bien long avec parmi eux, mes encadrants de stages précédent, mes professeurs, amis scientifiques, collègues, parents et le(a) meilleur(e) pour la fin me direz-vous, ma compagne, Émilienne Grzegorzcyk dont le soutien m'a toujours été indéfectible bien que je ne sois pas facile au quotidien.

Merci à tous, on se revoit sur le chemin, à bientôt les amis...

## Résumé en français

Aujourd'hui, la biodiversité est plus que jamais menacée par une destruction massive des environnements naturels, l'agriculture intensive et le dérèglement climatique. Il est donc plus qu'urgent et indispensable de mieux comprendre comment la biodiversité est générée et se maintient, pour être en mesure d'agir contre l'effondrement de la biodiversité.

La domestication fournit un cadre intéressant pour comprendre les mécanismes d'adaptation et le rôle de la diversité génétique dans la génération et le maintien de la biodiversité. Darwin se sert d'ailleurs de la domestication comme modèle pour comprendre comment les espèces apparaissent et établir les bases de la théorie de l'évolution. En effet, la domestication est un cas spécifique d'adaptation, impliquant une sélection forte et récente et une divergence rapide en lignées évolutives séparées. L'être humain a domestiqué un grand nombre d'animaux et de plantes. Certains micro-organismes, dont les champignons, ont aussi été domestiqués pour fermenter les aliments, produire des antibiotiques, et pour leur qualité gustative. Les champignons sont de bons modèles eucaryotes pour étudier l'évolution, car des espèces phylogénétiquement proches occupent des niches écologiques différentes, et ils possèdent de nombreux avantages pour les études expérimentales (e.g. faciles à cultiver au laboratoire, temps de génération courts) et génomiques (petits génomes souvent haploïdes).

*Penicillium camemberti* est la moisissure blanche et duveteuse typique en surface des fromages tels que le camembert et le brie. Nous avons montré que *P. camemberti* est le fruit d'une domestication en plusieurs étapes, menant à différentes variétés. Un premier événement de domestication a conduit à

l'émergence de la moisissure domestiquée bleue-grise nommée *P. biforme*, qu'on retrouve par exemple sur les fromages frais de chèvre. Un deuxième évènement de domestication, plus récent, a donné la lignée clonale *P. camemberti*. Les deux espèces domestiquées *P. camemberti* et *P. biforme* montrent des caractères avantageux pour l'affinage des fromages comparé à l'espèce sauvage proche. Nous avons aussi identifié deux variétés de *P. camemberti* aux phénotypes contrastés en termes de croissance, de couleur, de production de mycotoxines et de capacité d'inhibition des contaminants. La variété *camemberti*, blanche et duveteuse, est inoculée sur les fromages type camembert, alors que la variété *caseifulvum*, grise et moins duveteuse, est inoculée sur d'autres fromages comme le Saint Marcellin.

*Geotrichum candidum* est un champignon retrouvé naturellement dans le lait. Les analyses génomiques ont révélé trois clades monophylétiques, un clade comprenant uniquement des souches sauvages, un clade d'origine mixte comprenant des souches sauvages et du fromage et un clade comprenant uniquement des souches isolées de fromage. Les souches du fromage ont des phénotypes différents des souches sauvages, comme une croissance plus lente sur tous les milieux, même fromage, une production importante de composés volatils souhaitable et une activité protéolytique plus faible. Au sein du clade fromage, nous avons identifié trois populations aux phénotypes contrastés. Les populations du fromage sont génétiquement diversifiées et montrent des traces de recombinaison ; il n'y a donc pas eu sélection d'une lignée clonale utilisée comme ferment mais de plusieurs souches génétiquement et phénotypiquement différentes. Une des populations du fromage montre des traces d'un état plus avancé de domestication, avec une diversité génétique moins importante que les autres populations, un mycélium plus



dense et une capacité à inhiber plus efficacement des contaminants, comme *P. camemberti*.

*Penicillium camemberti* est le résultat de la sélection d'un unique mutant alors que l'espèce *G. candidum* possède encore une grande diversité. On a donc, avec ces deux espèces, une vision contrastée de la domestication des champignons du fromage.

## Résumé en anglais

Biodiversity is more than ever threatened by several global phenomena such as the continued destruction of natural habitats, intensive agriculture and climate disturbance. It is therefore urgent to better understand how biodiversity is generated and maintained, in order to be able to act against the collapse of biodiversity.

Domestication provides an interesting framework for understanding the mechanisms of adaptation and the role of genetic diversity in the generation and maintenance of biodiversity. Darwin used domestication as a model to understand how species arise and to establish the basis of the theory of evolution. Indeed, domestication is a specific case of adaptation, involving strong and recent selection and rapid divergence into separate evolutionary lineages. Although we are not the only species at the root of domestication events, humans have domesticated a large number of both animals and plants. Additionally, some microorganisms, including fungi, have also been domesticated for a variety of tasks such as controlled fermentation, the production of antibiotics, and consumed directly. Fungi are good eukaryotic models for studying evolution, because phylogenetically related species occupy different ecological niches, and they have many experimental and genomics assets (e.g. relatively easy manipulation within a lab setting, short generation times, small and often haploid genomes). One such setting, where fungi domestication has been championed, is in the production of cheese.

*Penicillium camemberti* is the typical white and fluffy mold on the surface of cheeses such as camembert and brie. We have shown that *P. camemberti* is the result of domestication in several steps, leading to different varieties. The first domestication event led to the emergence of the domesticated blue-gray mold named *P. biforme*, found on fresh goat cheese. A second, more recent domestication event resulted in the clonal lineage *P. camemberti*. Both domesticated species *P. camemberti* and *P. biforme* show advantageous traits for cheese ripening compared to the closest related wild species. Within the *P. camemberti* species, we identified two varieties (*camemberti* and *caseifulvum*) with contrasting phenotypes in terms of growth, color, mycotoxin production and ability of contaminant inhibition. The white and fluffy *camemberti* variety is inoculated on camembert-type cheeses, while the gray and less fluffy *caseifulvum* variety is inoculated on other cheeses such as Saint Marcellin.

*Geotrichum candidum* is a dimorphic fungus (yeast and filamentous form) found naturally in milk. Genomic analyses revealed three monophyletic clades, one comprising only wild strains, a mixed-origin clade comprising wild and cheese strains, and a clade containing only cheese strains. The cheese strains have different phenotypes from the wild strains, such as slower growth on all media, including cheese, high production of desirable volatile compounds and lower proteolytic activity. Within the cheese clade, we identified three populations with contrasting phenotypes. In contrast to *P. camemberti*, the *G. candidum* cheese populations are genetically diverse and show evidence of recombination; thus, there is not one clonal lineage used as a ferment but several genetically and phenotypically different strains. However, one of the cheese populations shows

evidence of a more advanced state of domestication, with less genetic diversity than the other populations, a denser mycelium and an ability to inhibit contaminants more efficiently, similar to that of *P. camemberti*.

*Penicillium camemberti* is the result of the selection of a single mutant, whereas the species *G. candidum* still has a great diversity. Thus, with these two species, we have a contrasting view of cheese fungi domestication.

# Sommaire

<b>Remerciements</b>	<b>3</b>
<b>Résumé en français</b>	<b>6</b>
<b>Résumé en anglais</b>	<b>9</b>
<b>Sommaire</b>	<b>12</b>
<b>Introduction</b>	<b>14</b>
I. La Biodiversité et la sixième crise d'extinction massive	15
II. Les théories de l'évolution	18
III. Mieux comprendre l'évolution grâce à l'adaptation parallèle	22
IV. La domestication, un modèle d'étude de l'adaptation	25
A) Domestication des mammifères et des plantes	25
B) Domestication des champignons	30
1) Le règne des champignons	30
2) Les champignons, de bons modèles pour étudier l'évolution	32
3) Les champignons impliqués dans la fermentation	32
4) Domestication de <i>Brettanomyces bruxellensis</i>	33
5) Domestication des <i>Saccharomyces</i>	35
6) Domestication de <i>Penicillium roqueforti</i>	38
V. Le fromage	44
A) Les modifications physico-chimiques	45
1) Catabolisme des glucides	46
2) Lipolyse et protéolyse	47
B) L'écosystème et la succession écologique	48
1) L'impact de la flore bactérienne en début de fabrication du fromage	48
2) Première flore d'affinage : les levures et les bactéries	49
3) Seconde flore d'affinage : les champignons filamenteux	49
4) Les différentes flores protègent le fromage des contaminants alimentaires	50
<b>Objectifs de la thèse</b>	<b>51</b>
<b>Communications scientifiques</b>	<b>53</b>
Valorisation des résultats sous forme d'article et mes contributions	53
Valorisation des résultats lors de séminaire	54
<b>Chapitre 1 Domestication de <i>Penicillium camemberti</i></b>	<b>55</b>
I. Résumé	56
II. Article : Domestication of the Emblematic White Cheese-Making Fungus <i>Penicillium camemberti</i> and Its Diversification into Two Varieties	59
III. Article : La domestication des champignons <i>Penicillium</i> du fromage	78

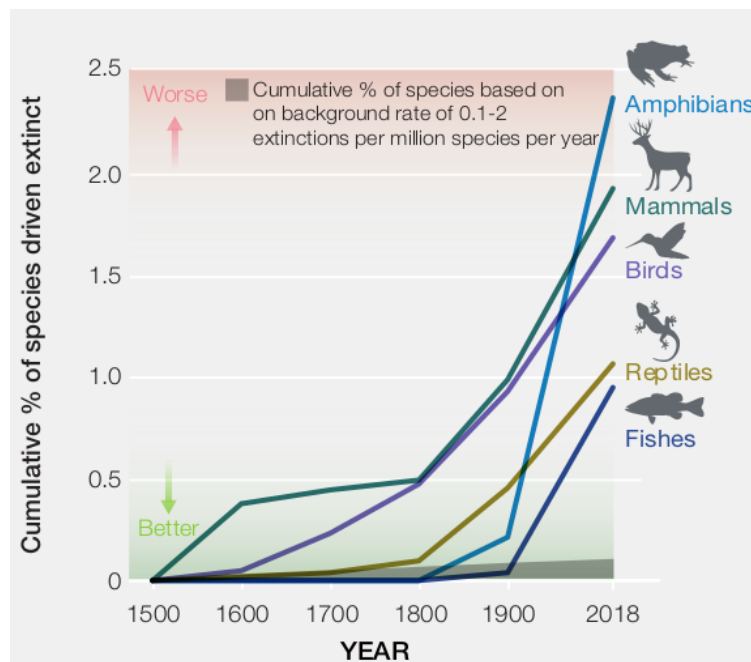
IV. Synthèse de la domestication de <i>Penicillium camemberti</i>	102
<b>Chapitre 2 Domestication de <i>Geotrichum candidum</i></b>	<b>104</b>
I. Résumé	105
II. Article : Domestication of different varieties in the cheese-making fungus <i>Geotrichum candidum</i>	108
III. Synthèse de la domestication de <i>Geotrichum candidum</i>	131
<b>Discussion générale</b>	<b>134</b>
I. Les conséquences de la domestication des champignons du fromage	135
A) Des niveaux de diversité génétique variés entre les espèces domestiquées	137
B) Une diversité phénotypique des souches qui reflète plusieurs usages	137
C) La population Cheese_2 de <i>G. candidum</i> : une trajectoire évolutive similaire à <i>P. camemberti</i>	138
II. Processus génétiques impliqués dans l'adaptation	140
A) Transferts horizontaux	140
B) Réarrangements chromosomiques	140
C) Variations du nombre de copie de gènes	141
D) Mutations ponctuelles	141
III) Phénotypes sélectionnés pour la production des fromages	143
A) Phénotypes acquis au cours de la domestication	143
B) Caractéristiques facilitant l'adaptation à la niche écologique du fromage	144
<b>Conclusion &amp; Perspectives</b>	<b>146</b>
<b>Références</b>	<b>152</b>

# Introduction

## I. La Biodiversité et la sixième crise d'extinction massive

En 2021, on recense plus de deux millions d'espèces identifiées (Figure 1). Cette estimation, pourtant époustouflante, n'est que la sous-estimation du nombre d'espèces, la biodiversité spécifique. En effet, ce nombre est limité par les moyens techniques et humains mis en place pour découvrir de nouvelles espèces et indiquer celles qui se sont éteintes.

Avec l'accroissement de l'impact de l'homme sur son environnement par une destruction massive des environnements naturels, l'agriculture intensive et le dérèglement climatique, le nombre d'individus par espèce tend à diminuer au point que les taux d'extinction sont de plus en plus élevés (Figure 1) (Ceballos et al. 2015). Le consensus scientifique actuel est que nous sommes en train d'assister à la sixième crise d'extinction massive, avec des taux d'extinction nettement plus forts depuis le début du XXème siècle.



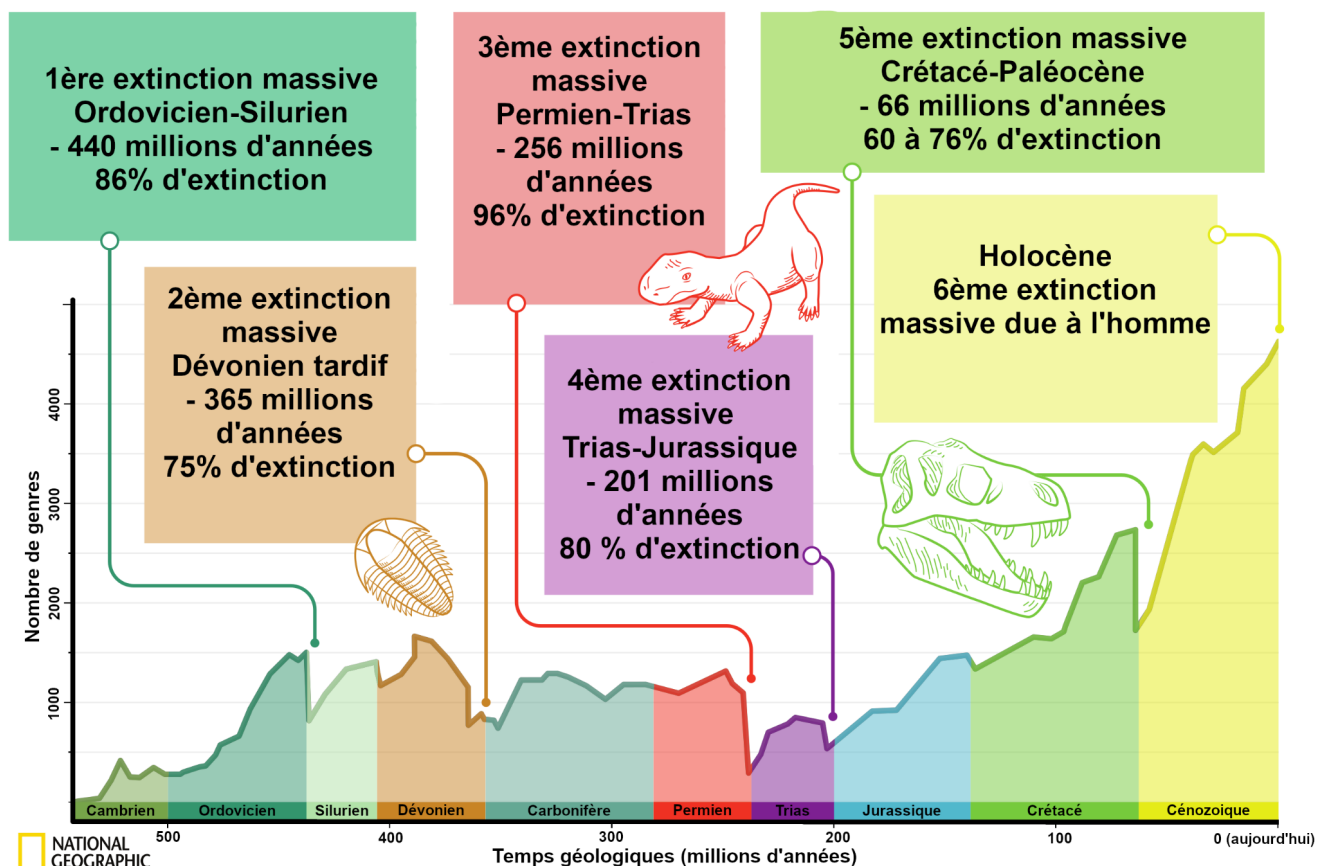
**Figure 1. Courbe cumulative du pourcentage d'espèces éteintes au cours des années.** Le taux attendu sans intervention de l'homme est indiqué en gris. Plus la pente de la courbe est forte, plus le taux d'extinction est fort. On peut voir que dans toutes les familles de



vertébrés le taux d'extinction a drastiquement augmenté depuis le XXème siècle. Source : rapport IPBES 2022.

Cette crise de la biodiversité a pour conséquence directe la perte irrémédiable d'espèces et *in fine* de services écosystémiques essentiels. A long terme, cela aura un impact non négligeable sur nos modes de vie et le bien-être des générations futures.

En analysant les données paléontologiques, il a été possible de déterminer cinq grandes extinctions de la biodiversité au cours des temps géologiques où le nombre d'espèces mondiales a chuté considérablement. Lors de la crise du Dévonien (-358 millions d'années), c'est 75% des espèces sur Terre qui ont disparu. Bien que ces crises aient réduit drastiquement le nombre d'espèces, on peut voir que de nouvelles espèces sont apparues après les crises (Figure 2).

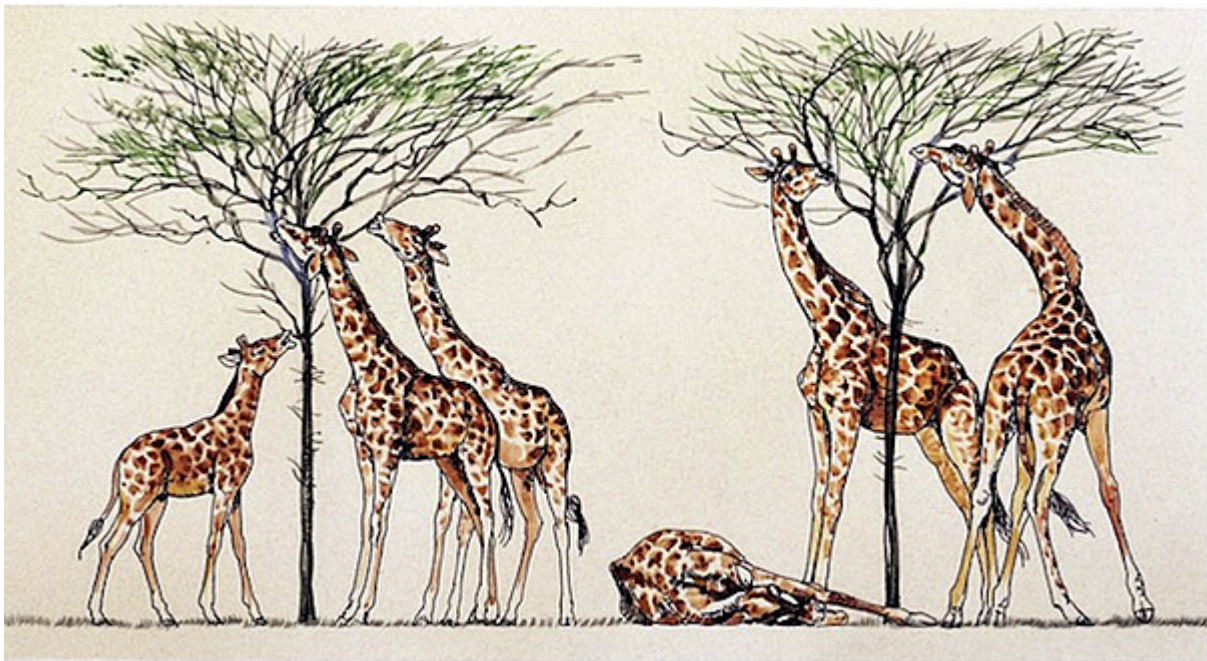


**Figure 2. Évolution du nombre de genres dans le registre fossile au cours du temps et crise majeure d'extinction.** Source : National Geographic.

La biodiversité spécifique est donc en évolution constante et c'est dans le cadre de la sixième crise de la biodiversité qu'il est plus que jamais nécessaire de comprendre comment les espèces évoluent au cours du temps.

## II. Les théories de l'évolution

En 1859, Charles Darwin publie son ouvrage "*On the Origin of Species*", qui remet en cause la conception fixiste des organismes. Le fixisme est une théorie selon laquelle les organismes vivants ne changent pas au cours des générations. Or pour Darwin, les organismes présentent des variations individuelles, qui à l'échelle des populations subissent la sélection naturelle : les individus les mieux adaptés au milieu actuel se reproduiront plus et si ces caractères adaptatifs sont héréditaires, alors au fil des générations ces variations seront fixées (tous les individus porteront cet allèle, cette version d'un gène) dans les populations. Chez la girafe, il existe une variation naturelle de la longueur du cou au sein de la population de girafe mais celles qui ont un long cou vont pouvoir manger les feuilles plus hautes alors que celles avec un cou court mourront lorsque les feuilles basses se seront taries (figure 3). De ce fait, au fil des générations, la longueur du cou, un caractère héréditaire, a subi la sélection naturelle.

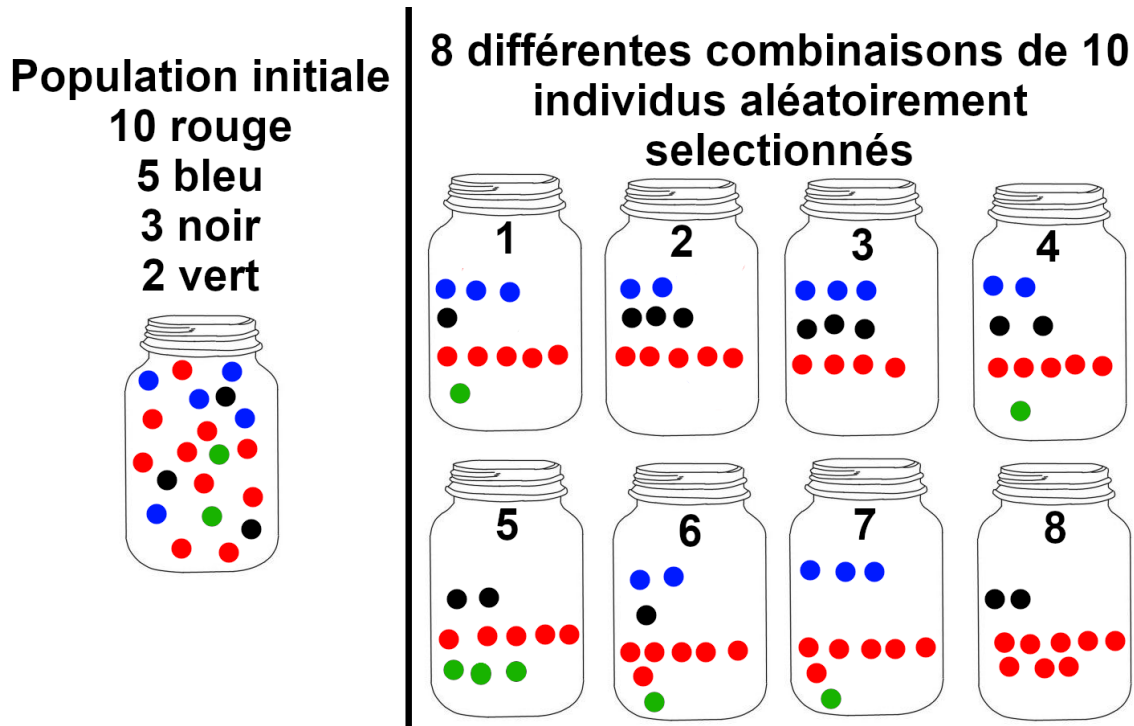


**Figure 3. Illustration de la sélection naturelle.** Dans la population de girafe, il peut y avoir de la variation interindividuelle au niveau de la taille du cou. Si jamais la nourriture devient

limitante et ne se trouve plus qu'en hauteur, alors seules les girafes au long cou pourraient se nourrir. Les girafes à long cou seraient les seuls à survivre. Avoir un long cou serait donc soumis à la sélection naturelle.

Cependant Darwin ne connaissait pas les mécanismes permettant de produire ces variations phénotypiques. Gregor Mendel aida à mieux comprendre les lois de l'hérédité en 1850 mais c'est bien plus tard, avec le développement de la génétique des populations que l'on aboutit à la théorie synthétique de l'évolution. Au sein de cette nouvelle théorie, l'évolution est régie par quatre grandes forces évolutives : la sélection naturelle, la dérive génétique, les mutations et la migration. Tout d'abord comme l'avait théorisé Darwin, la sélection naturelle élimine les individus les moins bien adaptés à un temps  $t$ , qu'ils survivent moins bien ou font moins de descendants, cela a pour effet de modifier les fréquences des allèles dans la population. Deuxièmement, la dérive génétique agit à chaque génération, il s'agit d'un effet stochastique dans le choix des individus qui vont se reproduire. La dérive génétique est d'autant plus intense que le nombre d'individus est faible. On peut facilement reproduire le mécanisme en jeu en prenant l'exemple d'un pot (représentant une population) rempli de billes de différentes couleurs (représentant des individus possédant des variations génétiques) (figure 4). On mime une génération en piochant aléatoirement dix billes (individus) avec reprise dans la population initiale. Ces individus piochés représentent les génotypes qui participent à la reproduction. En reproduisant cette pioche 8 fois on obtient des populations où certains variants génétiques (couleur) ont disparu bien que la pioche était aléatoire : les verts ont disparu dans les bocaux 2, 3 et 8; les bleus des bocaux 5 et 8; les noirs du bocal 7. On remarque aussi que le variant qui a eu le risque de disparition le plus fort est le vert qui était peu présent dans la population. De plus, si l'on reproduit la

même expérience en piochant moins de billes, la stochasticité est plus grande. La dérive génétique est d'autant plus grande que peu d'individus se reproduisent.



**Figure 4. Exemple de sous-échantillonnage dans une population initiale et de dérive génétique.** En partant d'un pot de 20 billes (représentant la population) de couleurs différentes (représentant des génotypes différents), on sélectionne aléatoirement avec reprise 10 billes (variants génétiques). Cette sélection aléatoire pourrait représenter le passage d'une génération à une autre via la reproduction de certains individus seulement. A la seconde génération, certaines couleurs (génotypes) disparaissent dans certains tirages. Dans les pots 2, 3 et 8 il n'y a plus le variant vert et le variant bleu a disparu des bocaux 5 et 8.

Les deux dernières forces agissant sur l'évolution des espèces sont les mutations génétiques, pouvant générer des variations phénotypiques des individus, et la migration, influant sur l'interconnection des populations entre elles.

A partir de ces différentes forces évolutives, on comprend que la biodiversité spécifique (nombre d'espèces) ne permet pas de se faire une vision complète de la dynamique de la biodiversité mondiale. Une espèce dont les individus n'auraient que peu de variation génétique n'a pas beaucoup de potentiel adaptatif et si un jour les conditions de vie changent trop, alors les risques d'extinction sont forts. La

biodiversité génétique est donc aussi importante. Une population avec très peu d'individus est sujette à une forte dérive génétique, menant à une baisse de la diversité génétique, phénomène que l'on appelle un vortex d'extinction (Blomqvist et al. 2010). De plus, les espèces interagissent entre elles, occupent des fonctions, des niches écologiques au sein d'un écosystème et l'extinction d'une espèce influence potentiellement la survie d'une autre. Pour conclure, dans l'idéal et pour avoir une image la plus complète de la biodiversité, il faut s'intéresser à toute sa complexité : génétique, populationnelle, spécifique, fonctionnelle et écosystémique. Comprendre l'évolution des organismes est une des clés pour faire face aux enjeux actuels.

### III. Mieux comprendre l'évolution grâce à l'adaptation parallèle

Dans son ouvrage "*Wonderful Life – The Burgess Shale and the Nature of History*", en 1989, Stephen Jay Gould repousse un peu plus loin les concepts de l'évolution en questionnant leur répétabilité. Pour illustrer son propos, Gould utilise la métaphore d'une cassette : si l'on rembobine les temps géologiques, le déroulé de l'évolution des espèces serait-il à nouveau le même ?

Dans un environnement similaire, observe-t-on forcément des convergences adaptatives ? Est-ce que ces adaptations mettent en jeu des mécanismes moléculaires similaires ? L'évolution est-elle répétable, et dans quelle mesure ?

Bien que la métaphore de Gould soit intéressante, il ne nous est pas possible de remonter le temps. Cependant, dans certains cas, des lignées phylogénétiquement indépendantes s'adaptent à une même niche écologique, c'est ce qu'on appelle l'adaptation parallèle. Cette situation est particulièrement puissante pour détecter quels sont les caractères adaptatifs dans une niche spécifique. Ces adaptations peuvent dans certains cas reposer sur des mécanismes moléculaires similaires (mêmes réarrangements chromosomiques, mêmes gènes impliqués, mêmes mutations, etc). L'adaptation parallèle est donc un excellent cadre pour étudier la répétabilité de l'évolution.

Certains cas d'adaptations parallèles en milieu naturel ont été étudiés, comme les poissons cichlidés des grands lacs africains. Les poissons cichlidés sont un des groupes de poissons téléostéens ayant la diversité spécifique la plus riche, avec au moins 700 espèces principalement présentes dans les grands lacs africains (Turner et al. 2008). Lors de la formation des grands lacs africains, des cichlidés ont migrés depuis le fleuve Congo vers ces nouveaux espaces et ont subi une radiation

évolutive (Schwarzer et al. 2011). Une radiation évolutive consiste en une grande diversification d'un groupe, ici des cichlidés, permis par une nouvelle niche écologique. Lors de la radiation évolutive, différentes lignées se sont spécialisées pour des régimes alimentaires particuliers et des profondeurs de vies différentes. On a pu observer des convergences phénotypiques entre les cichlidés d'espèces différentes habitant les différents lacs. Certains poissons vivent uniquement dans les tranches d'eau profondes (limnétique) et d'autres en surface (pélagique), or la luminosité de ces deux milieux est différente. Les poissons limnétiques ont tendance à avoir une sensibilité rétinienne différente de ceux qui sont pélagiques car l'atténuation de la lumière en profondeur dépend de la longueur d'onde. A partir d'un ancêtre commun ayant une même sensibilité rétinienne, différentes lignées se sont spécialisées en fonction du milieu de vie indépendamment dans les lacs Malawi, Victoria et Tanganyika chez des espèces différentes (O'Quin et al. 2010). La même convergence phénotypique a aussi été observée chez les espèces limnétiques ou pélagiques (Kassam et al. 2003). En revanche, des études génétiques ont montré que l'adaptation parallèle des cichlidés pour la sensibilité rétinienne et la morphologie n'a pas mis en jeu les mêmes mécanismes moléculaires (Elmer et al. 2010; 2014).

Il existe d'autres exemples d'adaptations parallèles, après la dernière glaciation. L'épinoche à trois épines est passée d'un milieu marin à un milieu lacustre par évolution répétée et indépendante. Le phénotype lacustre se distingue du phénotype marin par la perte de la plaque osseuse pelvienne, la couleur, l'halotolérance et le nombre d'épines dorsales (Elmer et Meyer 2011; Bell et A 1994). Chez les lézards du genre *Anolis*, présents sur différentes îles des Caraïbes, des études ont montré que certaines adaptations morphologiques en lien avec la vie en altitude sont



survenues de manières récurrentes (Mahler et al. 2013; Rodríguez et al. 2017). Certaines de ces convergences impliquent des mutations répétées dans certains gènes ou réseaux de gènes comme chez l'épinoche (Cresko et al. 2004; Cole et al. 2003; Corbett-Detig et al. 2020). Chez l'épinoche, l'analyse comparée des populations marines et lacustres a mis en évidence que l'adaptation parallèle a impliqué l'évolution répétée d'inversions de certaines régions génomiques. En empêchant la recombinaison, les inversions chromosomiques permettent au génotype de se maintenir, dans une sous-population, avant que ne se mette en place l'isolement reproducteur (Jones et al. 2012).

Les phénomènes d'adaptations parallèles ont aussi été décrits chez les plantes. Par exemple, les fleurs attirent les hyménoptères pollinisateurs grâce aux couleurs de leur pétales. Or ces hyménoptères sont capables de discriminer seulement certaines parties du spectre de la lumière. En analysant la couleur des fleurs d'Australie et celles de l'hémisphère nord, on retrouve des occupations du spectre de la lumière très similaires car contraints par la vision des hyménoptères (Dyer et al. 2012).

Les exemples d'adaptations parallèles en milieu naturel sont néanmoins rares.

## IV. La domestication, un modèle d'étude de l'adaptation

La domestication est un excellent cadre pour étudier l'adaptation étant donné que les caractères sélectionnés sont connus et que la sélection est forte et récente, impliquant une divergence rapide (Larson, 2014). De nombreuses espèces ont été domestiquées pour des utilisations allant de l'agriculture au transport ou encore la médecine. Des espèces très distantes phylogénétiquement ont été domestiquées pour répondre à un même besoin, dans un environnement similaire, préparant ainsi un terreau fertile pour l'émergence d'adaptations parallèles.

La domestication se définit selon plusieurs critères :

- Une différenciation génétique et phénotypique entre les populations vivant dans l'environnement anthropique et l'environnement naturel.
- Le contrôle de la reproduction par l'être humain.
- Un bénéfice pour l'être humain. On tend à exclure les pathogènes qui satisferaient aux autres critères, car bien qu'ils s'adaptent aussi au milieu anthropique, les caractères acquis ne sont pas bénéfiques pour l'être humain; on préférera donc le terme de pestification (De Gracia et al. 2015; Gladieux et al. 2018; Feurtey et al. 2020).

### A) Domestication des mammifères et des plantes

On retrouve des organismes domestiqués aussi bien chez les animaux (chien, bovidés, équidés, caprins, etc.) que chez les végétaux (blé, maïs, choux, etc.).

Chez les animaux de compagnie, on retrouve des caractéristiques communes tels que le maintien de phénotypes juvéniles, des couleurs de robes particulières, un comportement social exacerbé (Svartberg et Forkman 2002; Larson et al. 2012; Plassais et al. 2022) alors que chez les animaux agricoles, c'est davantage la force

musculaire, la taille, la production de lait ou de viande, ou encore une croissance rapide qui ont été sélectionnées (Diamond 2002; Frantz et al. 2015; Petersen et al. 2013; Warmuth et al. 2011; Liu et al. 2022; Qanbari et al. 2014). La domestication implique une divergence plus rapide que celle des populations en milieu naturel parce que la pression de sélection anthropique est plus forte. Les traces de la domestication sont donc plus susceptibles d'être visibles dans les génomes car plus récentes. Dans le cas des moutons et des chèvres, les études génomiques ont mis en évidence des traces de sélection dans des gènes contrôlant le système nerveux, le système immunitaire, la production de lait et la prise de muscles (Alberto et al. 2018). Le développement conjoint des méthodes de séquençage et de l'extraction de l'ADN ancien a permis de revisiter l'histoire de la domestication de certains organismes domestiqués. Par exemple, des études récentes de génomique ont révélé la présence d'une lignée de chevaux (datant de 4.700 à 4.000 ans) aujourd'hui éteinte en Ibérie qui ne serait pas l'ancêtre des chevaux domestiques modernes (Fages et al. 2019). Les chevaux domestiques modernes résultent d'un événement de domestication lié à un besoin de locomotion alors que les premières populations de chevaux domestiqués l'avaient été pour la production de viande et de lait (Librado et al. 2021). A partir des données génomiques, le foyer de la domestication des chevaux modernes a pu être situé dans l'ouest des steppes eurasiennes. D'autres analyses de génomique ont permis d'affiner l'histoire de la domestication des chiens (Bergström et al. 2022). Jusqu'à maintenant, les chiens et les loups d'Eurasie (*Canus lupus*) étaient considérés comme des lignées distantes phylogénétiquement. Or, des études récentes suggère que la diversité génétique des loups d'Eurasie serait sous-estimé et que ces deux espèces seraient plus proche que considéré jusqu'à maintenant. L'étude génétique des populations de

chiens laissent penser qu'il y a eu plusieurs populations ancestrales. Une première population proviendrait d'Eurasie orientale et une deuxième d'Eurasie occidentale. Les chiens de Sibérie, d'Amérique, d'Asie orientale sont affiliés à cette population d'Eurasie orientale. Les chiens du Proche-Orient et africains sont affiliés à la population d'Eurasie occidentale. Enfin on retrouve, dans une moindre mesure, des traces génétiques de ces ancêtres d'Eurasie occidentale parmi les chiens européens, ce qui laisse à penser qu'il y a eu des introgressions, des flux génétiques entre différentes populations. Ces deux exemples de domestication (chien et cheval) montrent l'intérêt des données génomiques pour retracer les histoires de la domestication. De multiples événements de domestication sont possibles et les introgressions des différentes populations peuvent compliquer la détection de ces événements.

Dans le cadre du développement de l'agriculture, la domestication des plantes a elle aussi été essentielle (Purugganan et Fuller 2009). De très nombreuses plantes (blé, riz, maïs, choux, abricot, etc) ont fait l'objet de domestication pour améliorer le rendement, la facilité de récolte, le goût, la maturité des fruits et la comestibilité (Hufford et al. 2012; Mabry et al. 2021; Peng, Sun, et Nevo 2011; Groppi et al. 2021). La perte de dispersion des graines à maturité résulte d'une modification d'un seul et même locus, *Sh1*, dans le cas du sorgho, millet, blé, maïs et de la séttaire d'Italie (Purugganan 2019). Cependant, les convergences phénotypiques des plantes domestiquées exploitent pour la plupart des modifications génétiques différentes (Gross et Olsen 2010; Martínez-Ainsworth et Tenailon 2016). Par exemple, plusieurs gènes peuvent être impliqués dans la dispersion des graines à maturité chez l'orge (Komatsuda et al. 2004).

Dans le cas du blé tendre, le génome a subi plusieurs épisodes successifs de différenciation et d'hybridation, afin de générer de nouvelles variétés (figure 5). Aux prémices de la domestication du blé, l'être humain cultivait le petit épeautre (*T. monococcum*) qui possède un génome diploïde de quatorze chromosomes, alors qu'aujourd'hui on cultive principalement le blé tendre (*T. aestivum*) possédant un génome hexaploïde composé de 42 chromosomes. C'est grâce à ces hybridations que le rendement du blé a été amélioré, au point qu'en 2021/2022, c'est 778 millions de tonnes de blé qui ont été produites dans le monde (source Organisation pour l'alimentation et l'agriculture FAO) (Peng, Sun, et Nevo 2011).

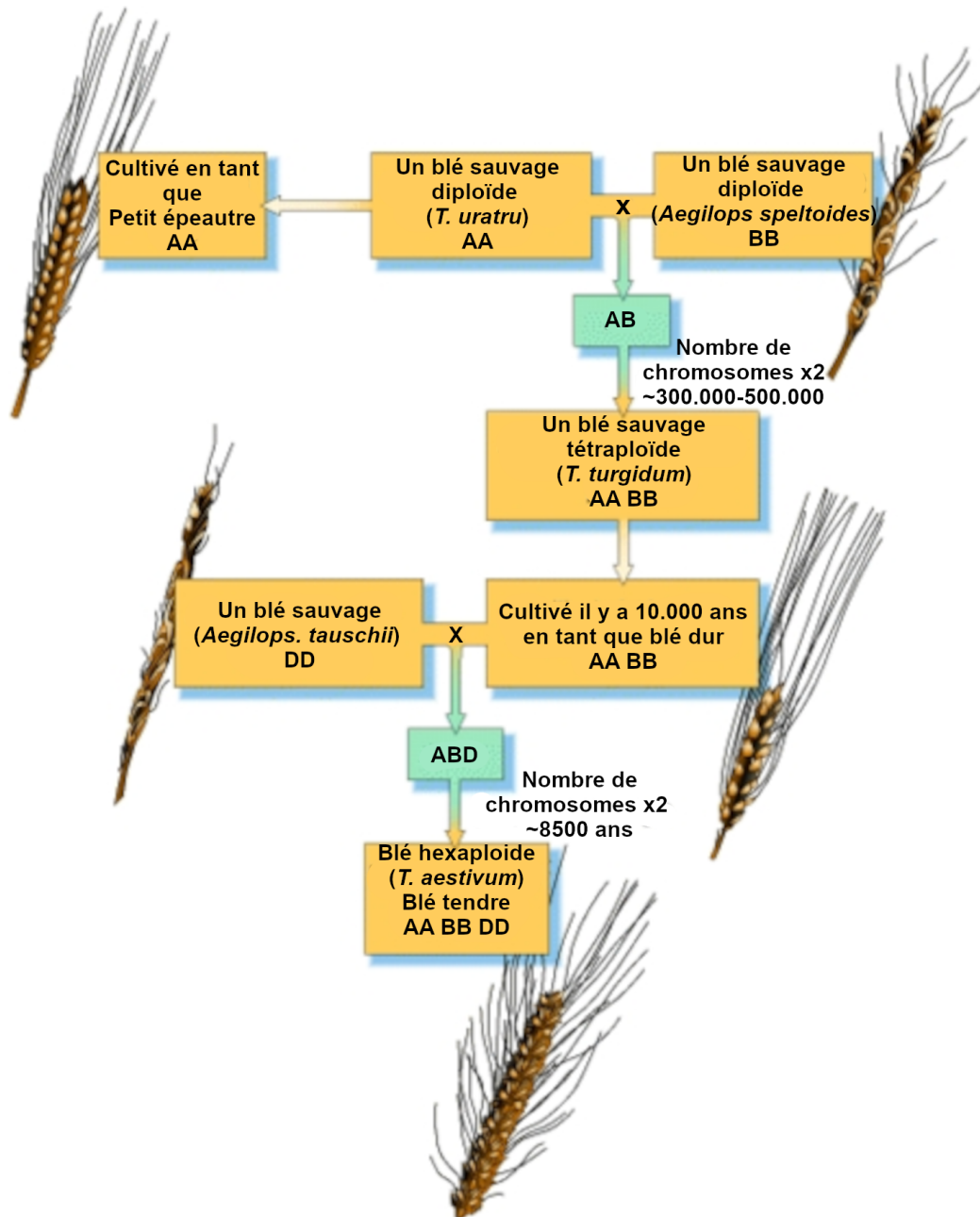


Figure 5. Évolution du génome du blé via de multiples hybridations. Source : Ben Himme.

L'amélioration des souches domestiquées ne se fait pas sans un certain coût. La sélection forte pour des caractères d'intérêt dans la population sauvage résulte souvent en un sous-échantillonnage drastique de la diversité : c'est ce qu'on appelle un goulet d'étranglement (Diamond, 2002 ; Meyer & Purugganan, 2013). De plus,

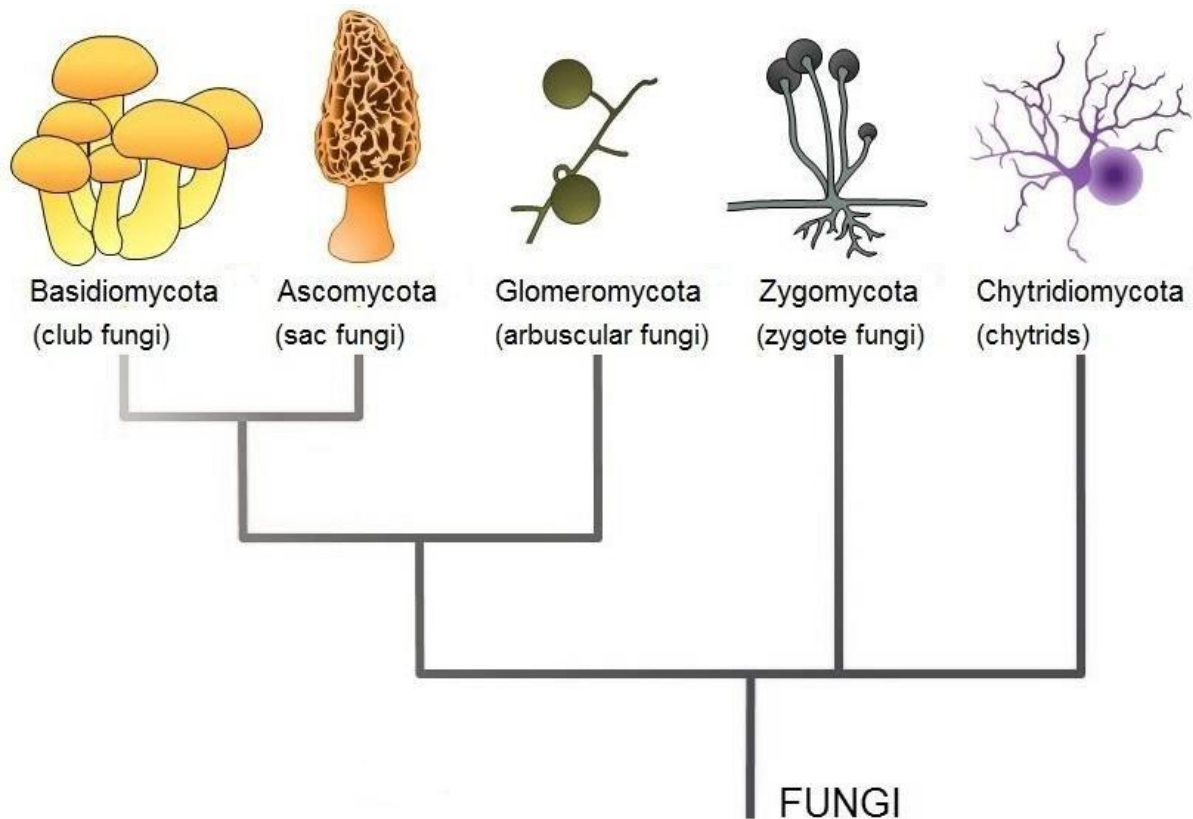
d'autres caractères, indispensables pour survivre dans l'environnement sauvage, ne le seront plus dans le nouvel environnement anthropique. Ces caractères vont ainsi dégénérer petit à petit, d'autant plus rapidement que la dérive génétique liée aux goulets d'étranglements est intense. Ce relâchement de la sélection s'illustre bien chez les plantes dont les mécanismes de défense dégénèrent dans les populations domestiquées (Cornille et al. 2014; Hufford et al. 2012).

## B) Domestication des champignons

### 1) Le règne des champignons

On retrouve une grande diversité morphologique et trophique au sein du règne des Fungi (champignons). On estime qu'il y aurait plus de deux millions d'espèces de champignons sur Terre et seulement une centaine de milliers décrites (Hawksworth et Lücking 2017). Les champignons sont des organismes eucaryotes hétérotrophes, possédant une paroi composée de chitine. C'est cette paroi qui conditionne le mode de vie des champignons. Qu'ils soient sous forme de levure (unicellulaire) ou mycélienne (multicellulaire en filaments), il leur est difficile d'ingérer de grosses particules par endocytose. De ce fait, ils doivent sécréter des exoenzymes pour dégrader les molécules nutritives du milieu en molécules plus simples, avant de pouvoir les assimiler.

Le règne des Fungi comprend cinq phyla : basidiomycètes, ascomycètes, gloméromycètes, zygomycètes et chytridiomycètes (figure 6).



**Figure 6. Les cinq phyla du règne Fungi.** De gauche à droite : basidiomycètes, ascomycètes, glomérormycètes, zygomycètes et chytridiomycètes. Figure modifiée à partir de Bear et.al, 2016.

Ces différents phyla regroupent des organismes macro et microscopiques aux modes de vie extrêmement divers. Les glomérormycètes vivent en symbiose avec des plantes, sous forme mycorhizienne, échangeant sel minéraux et eau en échange de glucides photosynthétiques provenant de la plante. D'autres champignons, moins bénéfiques aux plantes, sont parasites ou saprophytes (se nourrissent de bois mort). Parmi les parasites, on trouve *Batrachochytrium dendrobatidis* (chytridiomycètes), un champignon parasite de batraciens qui provoque le chytridiomycose, une maladie mortelle ravageant les populations de batraciens dans le monde (Berger et al. 2016). C'est au sein des ascomycètes que l'on retrouve les champignons utilisés par l'être humain pour la fermentation des aliments.



## 2) Les champignons, de bons modèles pour étudier l'évolution

Les champignons sont des modèles eucaryotes. Les connaissances acquises sur des modèles fongiques peuvent, en partie, être élargies à d'autres modèles eucaryotes (Stajich et al. 2009; Tirosh et al. 2006). Contrairement aux autres eucaryotes, les champignons ont des cycles de vie courts. Les champignons ascomycètes présentent aussi d'autres avantages tels que la possibilité de congeler leur spores et d'être cultivables en boîte de Pétri au laboratoire. Du point de vue génomique, les champignons ascomycètes ont de petits génomes, souvent haploïdes, ce qui facilite la production et l'analyse des données (Gladieux et al. 2014).

## 3) Les champignons impliqués dans la fermentation

Les exemples de plantes et animaux domestiqués sont nombreux, mais l'être humain a aussi domestiqué des microorganismes à des fins alimentaires (p.e. aliments fermentés, champignons de Paris, Shiitaké) et médicales (p.e. production de la pénicilline par *Penicillium rubens*) (figure 7).



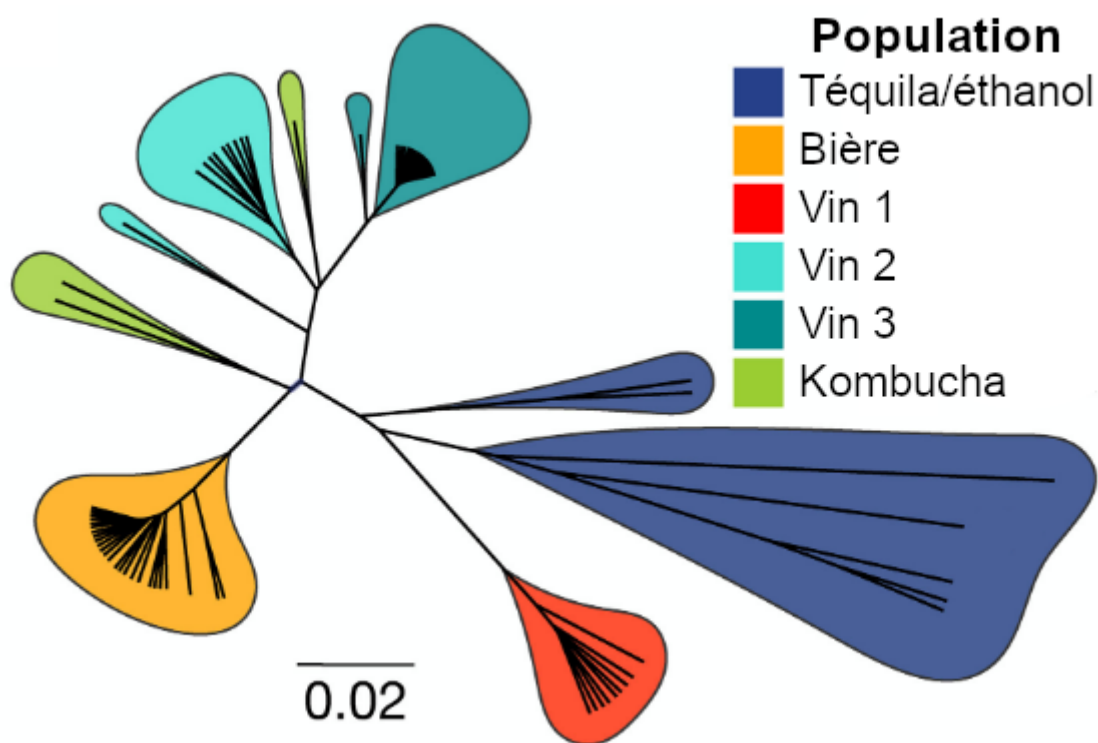
Figure 7. Exemples de microorganismes utilisés dans les fermentations alimentaires.

Bien qu'il y ait aussi des exemples de procaryotes utilisés à des fins similaires, je préfère ne pas les aborder dans ce manuscrit, car les procaryotes ont une dynamique évolutive particulière (Baptiste et al. 2009; Koonin et Wolf 2009).

La fermentation est utilisée depuis des millénaires partout dans le monde : c'est une des plus anciennes techniques de transformation et de conservation des aliments. Parmi les milliers d'aliments fermentés consommés quotidiennement, on peut citer le pain, les laits fermentés, le fromage, le vin, le saké ou le miso (Aidoo, Rob Nout, et Sarkar 2006). Au sein des champignons impliqués dans ces fermentations, on trouve les levures *Brettanomyces bruxellensis*, *Saccharomyces cerevisiae* et des champignons filamenteux tel que *Penicillium camemberti*, *Penicillium roqueforti* et *Aspergillus oryzae*. *Aspergillus oryzae* est utilisé pour fermenter les produits à base de riz et de soja en Asie comme la sauce soja et le miso shoyu (Galagan et al. 2005; Gibbons et al. 2012; Machida et al. 2005). Les souches domestiquées d'*A. oryzae* ne produisent plus d'aflatoxines, mycotoxines toxiques pour l'homme, et elles présentent des expansions de gènes du métabolisme des glucides et des acides aminés (Machida et al. 2005).

#### 4) Domestication de *Brettanomyces bruxellensis*

*Brettanomyces bruxellensis* est une levure retrouvée dans certains vins et bières (trappiste et bière aigre) mais aussi dans le kéfir et le kombucha (Tran et al. 2020; Lynch et al. 2021). Dans le vin et la bière, la présence de *B. bruxellensis* résulte en la production de composés phénoliques donnant des arômes épicés (Holt et al. 2018). Dans certains cas, *B. bruxellensis* est considérée comme une levure contaminante car ces composés phénoliques sont négativement perçus.



**Figure 8 :** Arbre de distance de *B. bruxellensis* produit à partir de 24.110 polymorphismes (SNP). On distingue six sous-populations correspondant à des niches écologiques spécifiques : tequila/éthanol, bière, vin 1–3 (trois populations) et kombucha (Eberlein et al. 2021).

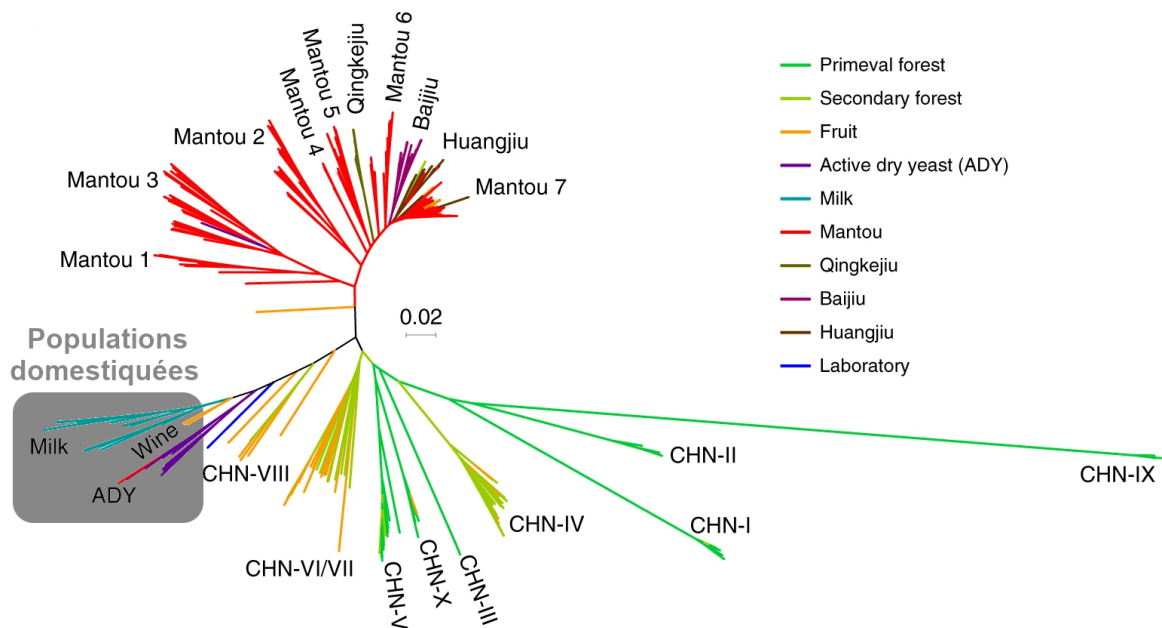
Les études de génétique de populations ont permis de montrer l'existence de plusieurs sous-populations, chacune attribuée à un usage spécifique (figure 8; Harrouard et al. 2022; Eberlein et al. 2021). La ploïdie varie en fonction des populations de *B. bruxellensis* entre diploïdie et triploïdie. Les populations des environnements tequila/éthanol, du vin et de la bière sont toutes exclusivement triploïdes. Cela résulte d'évènements d'allo-polypléidisation (polypléidie résultant d'une hybridation entre différentes espèces) et d'auto-polypléidisation. Ces évènements de polypléidisation peuvent avoir un rôle adaptatif à court terme. En effet, la polypléidisation permet de faire varier le nombre de copies de certains gènes (et donc leur transcription) mais aussi d'avoir la possibilité d'avoir une plus grande diversité de génotypes au sein d'un même individu (Morrow et Fraser 2013;

Steensels, Gallone, et Verstrepen 2021). La prévalence de la polyploïdie au sein de *B. bruxellensis* suggère que celle-ci a un rôle adaptatif bien que d'autres études sont nécessaires pour l'affirmer. La polyploïdie a un coût énergétique non négligeable, elle est contre sélectionnée si elle n'est pas adaptative (Neiman, Kay, et Krist 2013; Comai 2005).

D'autres adaptations à la niche anthropiques ont pu être mis en évidence telles que des duplications de gènes. Ces duplications impliquent des glucosidases (impliquées dans le métabolisme de l'amidon ou du galactose) et des gènes en lien avec l'assimilation de l'azote. Des transferts de gènes horizontaux (translocation de régions génomiques d'une espèce à une autre) sont aussi à l'origine de la capacité à métaboliser le sucrose de *B. bruxellensis* (Roach et Borneman 2020).

## 5) Domestication des *Saccharomyces*

*Saccharomyces cerevisiae* est à la fois l'organisme fongique le plus étudié et la levure dont l'impact économique est le plus important. *Saccharomyces cerevisiae* est utilisé pour la fermentation du vin, du pain, de la bière ou du fromage (Fay et Benavides 2005; Gallone et al. 2016; Legras et al. 2018; Libkind et al. 2011; Liti et al. 2009; Peter et al. 2018; Lahue et al. 2020; Bai et al. 2022; Lee et al. 2022). Les populations définies à partir des génomes montrent l'existence de populations spécifiques à chaque usage (figure 9; Duan et al. 2018).



**Figure 9 : Arbre phylogénétique des populations sauvages et domestiquées de *S. cerevisiae* à partir d'une matrice de 923.479 sites génomique (SNP) (Duan et al. 2018).**

Plus récemment, grâce à un meilleur échantillonnage de la population présente dans le milieu naturel, des études ont montré que les populations naturelles provenaient majoritairement d'Asie (Lee et al. 2022; Bai et al. 2022; Lahue et al. 2020). Le foyer de domestication de *S. cerevisiae* serait donc oriental. Les populations domestiquées de *S. cerevisiae* ont, seulement pour certaines, une diversité génétique réduite comparé aux populations sauvages. Cependant, la diversité à l'échelle globale de toutes les souches domestiquées est moins importante que celles des environnements naturels (Bai et al. 2022).

Dans le cadre de la boulangerie, ce sont les capacités de fermentation alcoolique de *S. cerevisiae*, résultant en la production de CO<sub>2</sub>, qui permettent de faire gonfler la pâte. De plus, l'hydrolyse de certains glucides complexes améliore leur digestibilité (L. De Vuyst et al. 2009; Gobbetti et al. 2014). Traditionnellement les pains étaient ensemencé à partir d'un levain, une mixture de farine et d'eau naturellement contaminée par des levures et des bactéries lactiques de l'environnement ambiant

(Luc De Vuyst et Neysens 2005). Le levain est prélevé pour faire du pain puis réapprovisionné régulièrement en eau et farine. Grâce aux travaux de Pasteur et Hansen au XIX<sup>ème</sup> siècle, il a été possible d'identifier la levure responsable de la fermentation du pain et de produire industriellement la levure de boulanger sous forme déshydratée. Au XX<sup>ème</sup> siècle, la levure lyophilisée (contenant uniquement *Saccharomyces cerevisiae*) s'impose dans une grande partie du monde occidental car cela permet d'uniformiser les processus (Mattanovich, Sauer, et Gasser 2014). Les souches utilisées actuellement pour faire du pain font partie de deux populations génétiques distinctes, une population industrielle caractérisée par une triploidie et une phase de démarrage de fermentation plus courte, et une population du levain dont les gènes du métabolisme du maltose sont dupliqués (Bigey et al. 2021).

Lors de la fabrication de la bière ou du vin, la maîtrise de la production d'alcool par fermentation alcoolique est cruciale. On retrouve naturellement *S. cerevisiae* sur la peau des raisins dont l'aspect ressemble à une poudre blanche. Malgré tout, il est souvent d'usage d'ajouter des souches industrielles dont on connaît bien les capacités (temps de fermentation, composés volatils). Les souches isolées de vin et de bière font partie de deux populations génétiques distinctes. Les souches utilisées pour la bière sont caractérisées par des duplications de gènes impliqués dans le métabolisme du maltotriose et ne produisent plus de composés aromatiques non souhaités dans la bière. Comme pour le pain, de telles adaptations sont accompagnées de variations au niveau de la structure chromosomique et des hybridations (Gallone et al. 2016; 2018).

L'hybridation inter- ou intra-spécifique est fréquente au sein du genre *Saccharomyces*, ce qui est un moyen d'adaptation rapide pour s'adapter à un nouvel environnement (Gabaldón 2020). Par exemple, *Saccharomyces pastorianus*, la

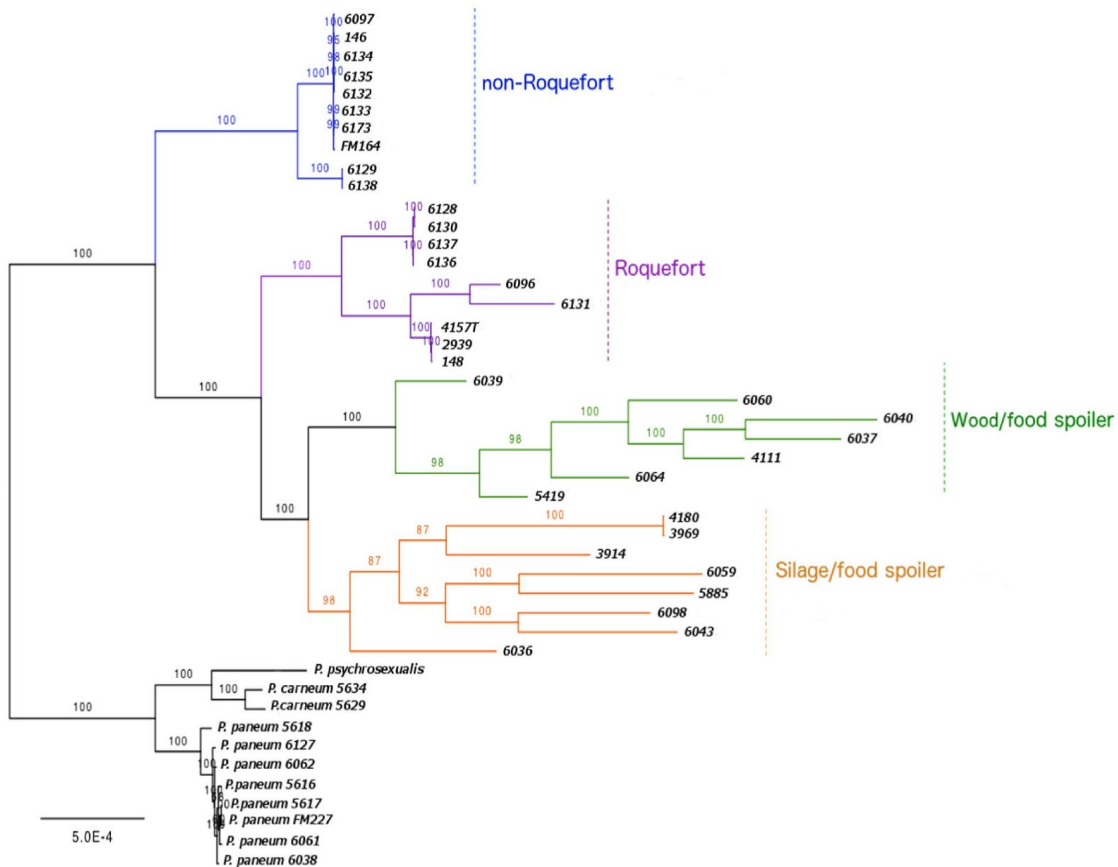
levure utilisée pour la production de la bière de type lager (fermentation à basse température) est le fruit de l'hybridation entre *S. cerevisiae* et *S. eubayanus* (Monerawela et Bond 2018). L'hybridation se produit aussi dans le milieu naturel, comme décrit chez *S. paradoxus* dans les forêts d'Amérique du nord (Leducq et al. 2016). Pour certains caractères, les hybrides peuvent avoir un phénotype plus avantageux que celui des deux parents, c'est ce qu'on appelle la vigueur hybride. La vigueur hybride est d'ailleurs grandement utilisée en agriculture avec ce que l'on appelle les lignées F1 (résultant de l'hybridation de deux parents F0 de génotypes différents).

Enfin, *S. cerevisiae* est retrouvé dans l'environnement laitier et plus particulièrement le fromage (Roostita 1996; Legras et al. 2018). Les souches fromagères forment une population spécifique. Les souches de la population du fromage ont subi de nombreux réarrangements chromosomiques. Le gène DLD3, une enzyme avec une activité de D-lactate déshydrogénase, a été dupliquée tandis que des gènes du métabolisme du galactose et des enzymes codant pour des transporteurs d'hexose ont subi des délétions. Les gènes du métabolisme du galactose ayant subi une délétion ont été remplacés par des orthologues acquis par transfert horizontal. La perte de transporteurs d'hexose résulte de la fusion de plusieurs transporteurs d'hexose entre eux. Ainsi, on retrouve à nouveau le fait que des variations structurales du génome de *S. cerevisiae* sont sources d'adaptation lors de la domestication.

#### 6) Domestication de *Penicillium roqueforti*

*Penicillium roqueforti* est la moisissure bleue utilisée pour fabriquer tous les fromages à pâte persillée du monde entier comme le Roquefort, le Gorgonzola ou le Stilton. Bien que sa niche écologique naturelle soit encore inconnue, il est

fréquemment retrouvé en tant que saprophyte (décompose de la matière organique morte). Lors du stockage de la paille (ensilage), *P. roqueforti* se développe lorsque les conditions d'aération sont insuffisantes. La paille devient alors impropre à la consommation par le bétail à cause de mycotoxines (O'Brien et al. 2006; Malekinejad et al. 2015).



**Figure 10 : Arbre de maximum de vraisemblance de *P. roqueforti*.** L'arbre est raciné à partir des espèces proches de *P. roqueforti* : *P. psychrosexualis*, *P. carneum* et *P. paneum*. Le bootstrap est indiqué à partir de 1000 répétitions. La barre d'échelle indique le nombre de substitutions par site. Source : Dumas et al. 2020

L'analyse génomique de souches provenant de divers environnements (fromage et autres environnements), a permis d'identifier quatre populations génétiques chez *P. roqueforti* (figure 10). Les souches du fromage et les souches isolées d'autres environnements (ensilage, contaminants alimentaires, bois, etc) forment des populations distinctes (Dumas et al. 2020). Les souches isolées de fromage font



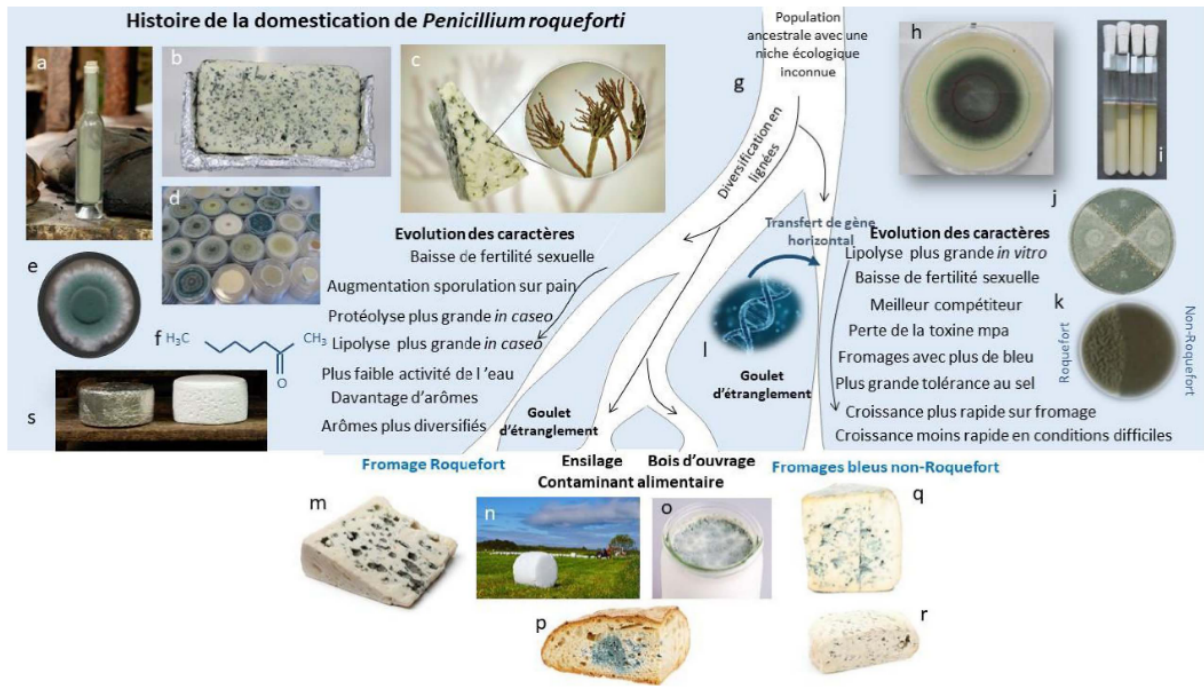
partie de deux populations génétiques distinctes, une première population rassemble uniquement les souches isolées de fromages Roquefort AOP (population Roquefort) alors que la seconde population rassemble toutes les souches isolées de fromages bleus du monde entier (population non-Roquefort). La population Roquefort présente une diversité génétique supérieure à la population non-Roquefort, qui est clonale. L'appellation d'origine protégée (AOP) a donc permis de conserver la diversité unique au Roquefort alors que la seconde population, fruit d'une sélection plus forte pour une production à une échelle industrielle et mondiale, a résulté en l'utilisation de la même souche.

Ces deux populations du fromage de *P. roqueforti* sont aussi phénotypiquement différenciées car elles ont été sélectionnées par l'être humain pour des caractères différents. Les souches de la population Roquefort sporulent plus fortement sur du pain, milieu de culture sur lequel était maintenu traditionnellement le *P. roqueforti* durant les périodes où il n'y avait plus de lait de brebis. Les souches de la population Roquefort présentent aussi une protéolyse et une lipolyse plus intenses in caseo et des arômes plus diversifiés (Caron et al. 2021). Au contraire, pour des besoins industriels, la population non-Roquefort a été sélectionnée pour une croissance plus rapide, une densité de bleu plus forte, une meilleure tolérance au sel et une inhibition plus importante des contaminants (figure 11; Ropars et al. 2015, Dumas et al. 2020).

Les souches isolées d'autres environnements que fromage forment les deux autres populations génétiques, l'une regroupe les souches de l'ensilage et l'autre du bois d'ouvrage et les contaminants alimentaires. Les goulets d'étranglement des populations sauvages sont moins importants que ceux des populations du fromage.

D'un point de vue génomique, on retrouve de nombreux segments génomiques acquis par transfert horizontal dans la lignée non-Roquefort. Parmi ces régions, deux possèdent des gènes d'intérêt pour l'adaptation au milieu fromage, appelées *Wallaby* et *CheesyTer* (d'une longueur de 500 kb et 80 kb, respectivement) (Cheeseman et al. 2014; Ropars et al. 2015). Dans *CheesyTer* se trouvent des gènes d'intérêt pour le métabolisme du lactose telles qu'une bêta-galactosidase et une lactose perméase, tandis que l'on retrouve dans *Wallaby* des gènes codant des protéines anti-microbiennes.

Les populations du fromage ont subi des goulets d'étranglements importants suite à la sélection de quelques individus, et sont propagées clonalement par les industriels, ce qui peut mener à une dégénération des fonctions reproductives sexuées. Les populations fromagères sont moins fertiles (Ropars et al. 2014). Soit les souches ne parviennent pas à former de structure reproductrice, les cléistothèces, soit ceux-ci sont présents mais ils sont vides (pas d'asques et donc d'ascospores). En revanche, les croisements entre souches d'autres environnements étaient fertiles (Ropars et al. 2016). La domestication de *P. roqueforti* a donc mené à une baisse de la fertilité.



**Figure 11. Histoire de la domestication du *Penicillium roqueforti* pour la fabrication des fromages bleus.** D'une population ancestrale avec une niche écologique jusqu'alors inconnue, les quatre populations identifiées de *P. roqueforti* se sont diversifiées en s'adaptant aux niches écologiques créées par l'homme, avec deux populations domestiquées indépendamment pour la fabrication de fromage, et deux populations contaminant la nourriture et l'ensilage. L'une des deux populations ne se trouve que dans les fromages appartenant à l'appellation d'origine protégée (AOP) Roquefort et abrite une certaine diversité génétique, tandis que l'autre population est une lignée clonale unique utilisée pour tous les autres fromages bleus du monde. Les goulets d'étranglement se sont produits dans toutes les populations, étant les plus forts dans les populations de fromages. Les populations fromagères ont développé des caractères spécifiques bénéfiques pour la fabrication du fromage, tels que les arômes, la coloration bleue des cavités du fromage, la lipolyse, la protéolyse et la tolérance au sel, avec des spécificités différentes dans les deux populations. La population de Roquefort produit une plus grande quantité de spores lorsqu'elle est cultivée sur du pain, le milieu de culture ancestral de Roquefort sur Souzlon. a : spores de *P. roqueforti* prélevées sur du pain inoculé, cuit auparavant en produisant des croûtes brûlées, comme on l'a longtemps fait pour l'inoculation du fromage bleu. b : tranche de fromage expérimental ressemblant au Roquefort. c : Morceau de fromage bleu avec des spores de *P. roqueforti* grossies en incrustation, portées dans les "pinceaux" typiques des champignons *Penicillium*. d : Boîtes de Pétri telles qu'elles sont utilisées aujourd'hui pour la culture stérile de *P. roqueforti*. e : Colonie de *P. roqueforti* sur une boîte de Pétri avec un milieu malté. f : Molécule de 2-heptatone responsable de l'un des arômes typiques du fromage bleu de Roquefort. g : Illustration de l'histoire de la divergence déduite des quatre populations de *P. roqueforti* identifiées, deux étant utilisées pour la fabrication du fromage et ayant souffert de forts goulets d'étranglement, une contaminant les aliments et le bois d'œuvre et une l'ensilage et également les aliments. h : Colonie de *P. roqueforti* sur une boîte de Pétri avec un milieu de fromage, avec des caractéristiques de mesures de taux de croissance. i : Tubes avec tributyrine utilisés pour mesurer le taux de lipolyse. j : Boîte de Pétri avec deux souches de *P. roqueforti* et structures sexuées (appelées cléistothèces, en marron) induites dans la zone de confrontation. k : Expérience de compétition sur une boîte de Pétri avec deux souches de *P. roqueforti*, montrant une plus grande capacité de compétition et une couleur plus foncée de la souche non-Roquefort. l : Acquisition horizontale de gènes dans la lignée non-Roquefort. m : Morceau de fromage roquefort. n :

Balle d'ensilage dans un champ. o : Yaourt contaminé par *P. roqueforti*. p : Pain contaminé par *P. roqueforti*. q et r : Fromages bleus non Roquefort (Fourme d'Ambert et Bleu d'Auvergne). s : Fromages de Roquefort sur leurs étagères en bois dans les caves, l'une étant enveloppée de la typique feuille de plomb. (Ropars, Caron, et al. 2020)

## V. Le fromage

Le fromage est un aliment fermenté dont la plus vieille trace de production date d'il y a 7200 ans (McClure et al. 2018). La fabrication du fromage nécessite de nombreuses étapes et variantes, influençant les caractéristiques physico-chimiques du produit final et donc les potentiels micro-organismes qui pourront s'y développer.

Le lait est une émulsion de matière grasse (87% d'eau) contenant du lactose et de la caséine. La première étape de fabrication du fromage est le caillage du lait. La coagulation du lait consiste à faire précipiter les micelles de caséines du lait. Pour cela, deux méthodes sont possibles : baisser le pH ou utiliser des enzymes telle que la présure. La présure est un cocktail d'enzymes extraites de la caillette (4ème estomac) de jeunes ruminants contenant des chymosines et des pepsines, des enzymes pouvant dégrader les protéines. L'intérêt de cette étape est de pouvoir retirer une bonne partie de l'eau du lait en séparant le lait en caillé (bloc contenant les protéines) et en lactosérum (aussi appelé petit lait). C'est avec le caillé que l'on fabriquera la plupart des fromages. Les différents fromages se différencient tout d'abord par le fait qu'on égoutte plus ou moins le caillé et du fait qu'il soit pressé ou cuit (figure 12). Le fromage est salé directement ou par bain de saumure. Le sel préserve le fromage de la croissance de certains pathogènes ou microbes non voulus mais il donne aussi du goût.

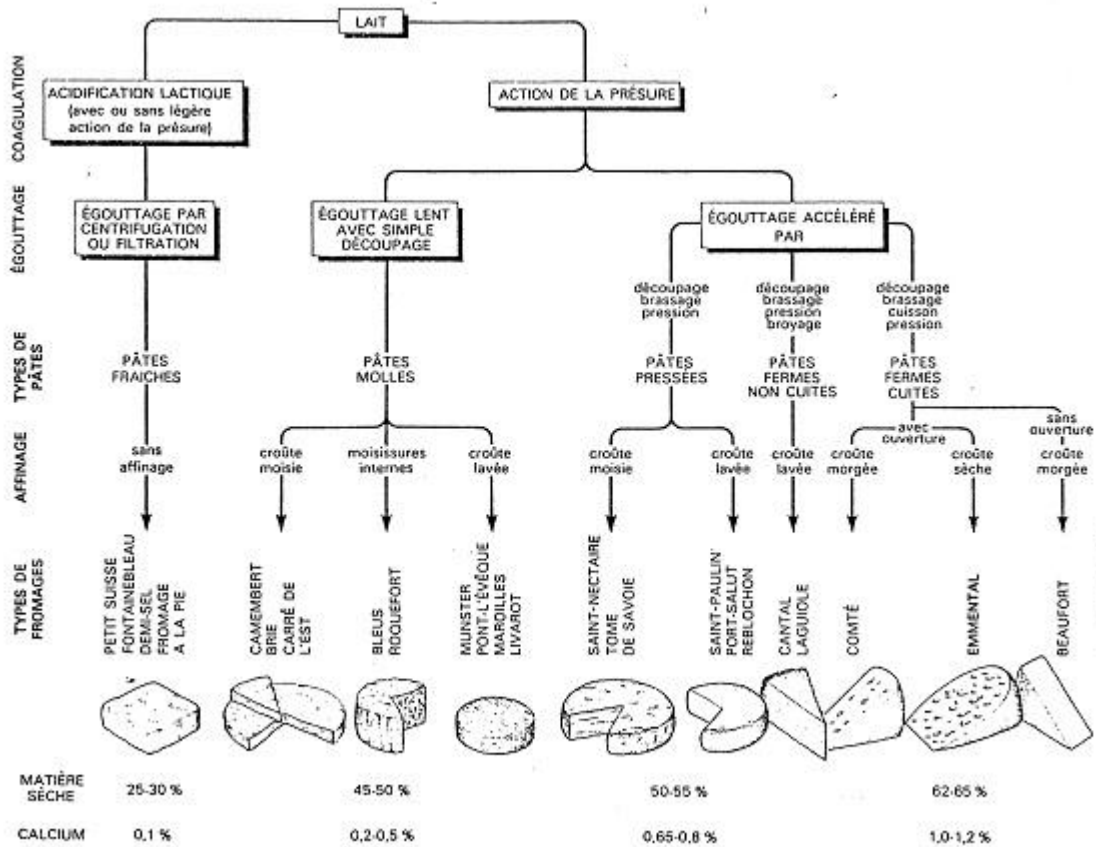


Figure 12. Technologies fromagères et fromages associés. Source : genie-alimentaire.com.

L'étape suivante est l'affinage. L'affinage développe la texture, le goût et les arômes du fromage. Ce sont les nombreuses réactions chimiques, réalisées par la communauté fongique et bactérienne, qui vont produire les composés volatils du fromage et en changer les propriétés physico-chimiques (McSweeney 2004).

### A) Les modifications physico-chimiques

Au cours de la fabrication puis de l'affinage d'un fromage, de nombreuses modifications physico-chimiques vont influencer la texture et les arômes du fromage. Trois grandes voies sont mises en jeu : le catabolisme des glucides, la lipolyse et la

protéolyse, reflétant les trois principaux substrats métaboliques que sont le lactose (et citrate en moindre quantité), les triglycérides et les caséines, respectivement.

### 1) Catabolisme des glucides

Dans le lait, on retrouve deux types de glucides : le lactose et le citrate. Le lactose est rapidement métabolisé par les bactéries lactiques en lactate lors des premières étapes de fabrication du fromage. Cette production de lactate va influencer la texture du caillé en rendant plus acide le lait. Plus le lait est acide, plus les caséines vont précipiter. Dans un second temps, lors de l'affinage, la communauté bactérienne et fongique va métaboliser le lactate en divers composés (acétate, acétoïne, éthanol, CO<sub>2</sub>). Cela a pour conséquence une désacidification, la formation de bulle dans le cas des fromages type emmental et la formation de composés volatils odorants.

Le citrate, présent dans une moindre mesure, sera dégradé par les bactéries lactiques. Les différents produits de dégradation (figure 13; acétate, 2,3 butanediol, acétoïne) sont des composés importants dans la flaveur (sensations olfactives et gustatives ressenties lors de la dégustation d'un aliment) du fromage affiné. Par exemple, le 2,3 butanediol rappelle l'odeur du beurre.

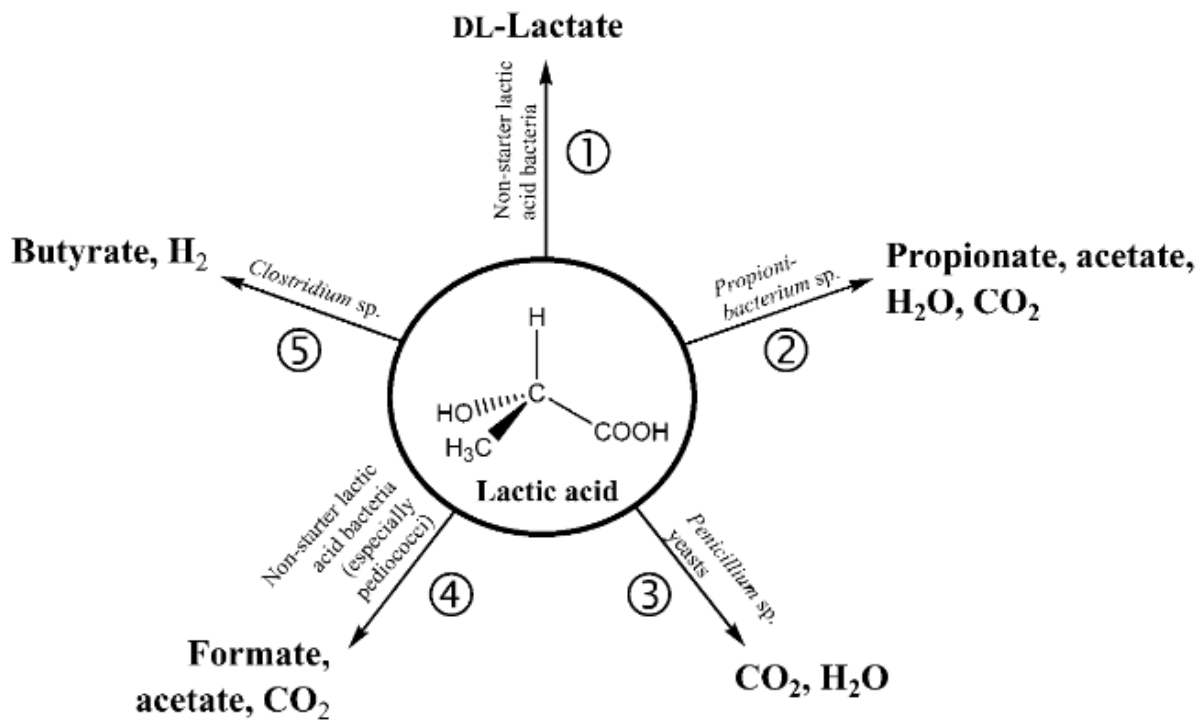


Figure 13. Principales voies de dégradation du lactose selon les microorganismes mis en jeu. Source : McSweeney 2004.

## 2) Lipolyse et protéolyse

La lipolyse et la protéolyse sont des processus métaboliques essentiels à la texture et l'arôme du fromage (McSweeney 2004).

Les triglycérides contenus dans le lait sont hydrolysés par diverses lipases produites par les microorganismes mais aussi par des lipases présentes dans le lait. Sous l'action des lipases, les triglycérides sont hydrolysés en acide gras libre à chaîne courte dont la contribution gustative est essentielle. Une lipolyse trop forte peut résulter en un goût rance.

La protéolyse contribue à la production de molécules olfactives mais elle a aussi un rôle structural important. Plus la protéolyse sera forte, plus le fromage sera coulant (comme le Chaource). Cet effet provient du fait qu'en hydrolysant les caséines du fromage, les protéases déstructurent la matrice de protéine et produisent des acides



aminés libres ayant un fort pouvoir attracteur de l'eau. Il y a aussi production de petits peptides et d'acides aminés souvent amers, que l'on remarque assez aisément dans des fromages à pâte molle laissés trop longtemps au réfrigérateur, car le froid n'empêche pas la protéolyse de se poursuivre. Certains de ces acides aminés pourront être métabolisés dans de nouvelles voies biochimiques pour mener à des composés volatils majeurs pour le goût du fromage.

Ces différentes réactions chimiques sont le fruit du métabolisme de divers microorganismes du fromage, qui ont chacun leur spécificité et vont selon leurs enzymes, donner des fromages uniques.

## B) L'écosystème et la succession écologique

Le fromage évolue au cours du temps, sa croûte se différencie du cœur à travers différentes réactions métaboliques. C'est un substrat dynamique où bactéries, levures et moisissures coexistent, entrent en compétition et se succèdent au cours du temps. Parmi les champignons du fromage, on retrouve notamment des levures telles que *Geotrichum candidum*, *Debaryomyces hansenii*, *Saccharomyces cerevisiae* et *Kluyveromyces lactis* ainsi que des champignons filamenteux des genres *Mucor*, *Scopulariopsis*, *Sporendonema* et *Penicillium* (Dugat-Bony et al. 2016; Ropars et al. 2012; Hermet et al. 2012).

### 1) L'impact de la flore bactérienne en début de fabrication du fromage

Tout commence avec du lait, une matrice liquide où seuls les microorganismes pouvant se développer en condition liquide se multiplient. Le lait est contaminé par de nombreuses bactéries et champignons de l'environnement de traite, certains se développant de suite, d'autres restant sous forme latente pour se développer plus tard quand les conditions sont favorables. Le fromager a le choix de pasteuriser,

stériliser ou garder un lait cru puis d'ajouter des ferments lactiques. Nous retrouvons donc principalement des bactéries lactiques des familles Enterococcaceae, Lactobacillaceae, Leuconostocaceae et Streptococcaceae qui vont acidifier le milieu. Durant la première étape de fermentation lactique, c'est principalement la flore bactérienne qui va dominer l'écosystème.

## 2) Première flore d'affinage : les levures et les bactéries

Une fois en fromage frais, le fromage est devenu une matrice solide, salée et acide. Le fait d'avoir un fromage solide a un impact direct sur la distribution des organismes car l'intérieur du fromage ne sera pas suffisamment oxygéné pour des organismes aérobies. C'est d'ailleurs pourquoi pour fabriquer un roquefort, il est nécessaire de percer le fromage pour avoir la moisissure bleue (*Penicillium roqueforti*) de manière homogène à l'intérieur du fromage. Dans certains fromages tels que l'emmental, le parmesan et le cheddar, c'est principalement à l'intérieur du fromage que se passent la plupart des réactions physico-chimiques par des microorganismes anaérobies (bactéries propioniques du genre *Propionibacterium*). Des levures telles que *Debaryomyces hansenii* ou *Geotrichum candidum*, voire la bactérie *brevibacterium linens*, vont s'installer sur la croûte du fromage (Dugat-Bony et al. 2015; Boutrou et Gueguen 2005), tolérant le sel et l'acidité. Ces microorganismes, en métabolisant le lactate, vont notamment faire remonter le pH de la croûte, ce qui va permettre à une nouvelle flore de s'installer (Irlinger et Mounier 2009).

## 3) Seconde flore d'affinage : les champignons filamenteux

La désacidification de la croûte permet aux champignons filamenteux d'y croître. Les plus connus sont ceux des genres *Penicillium*, *Mucor* et *Scopulariopsis*. Les pratiques du fromager sont cruciales lors de cette étape. Le fromager peut frotter la

croûte avec un chiffon plus ou moins humide, une brosse, ne pas frotter du tout, envelopper dans une étamine, dans de la cire, etc. La communauté microbienne d'affinage sur la croûte est fortement impactée par ces différentes pratiques. Le fromage est donc un écosystème, lieu de diverses successions écologiques avec diverses niches écologiques. Dans certains cas, les champignons responsables de l'affinage proviennent de l'environnement de la cave d'affinage.

#### 4) Les différentes flores protègent le fromage des contaminants alimentaires

Des contaminants alimentaires peuvent aussi croître sur le fromage conduisant à des défauts de fabrication ou des risques sanitaires. Pour éviter cela, les fromagers prennent un maximum de précaution sanitaire tout au long du processus de fabrication, mais ils peuvent aussi compter sur les différentes bactéries et champignons du fromage qui entrent en compétition indirectement ou directement avec les pathogènes et les contaminants. Par exemple, *G. candidum* est capable d'inhiber la bactérie responsable de la listériose, *Listeria monocytogenes* (Dieuleveux et al. 1998; Nielsen, Frisvad, et Nielsen 1998).

# Objectifs de la thèse

Durant ma thèse, je me suis intéressé à deux champignons inoculés en fromagerie, *Geotrichum candidum* et *Penicillium camemberti*. Le laboratoire possède une collection de souches isolées de différents types de fromage et d'autres environnements, et les données génomiques de ces souches.

Ma thèse s'articule autour de deux questions :

- 1) Est-ce que les champignons *P. camemberti* et *G. candidum* sont adaptés au fromage ?
  - a) Est-ce que les souches sauvages sont génétiquement différenciées des souches isolées d'autres environnements ?
  - b) S'il y a une différenciation génétique entre souches sauvages et souches du fromage, quels sont les processus génomiques impliqués dans l'adaptation au fromage ?
  - c) Est-ce que les populations du fromage ont développé des caractères spécifiques, bénéfiques pour la fabrication de fromage? comme des arômes spécifiques, une activité lipolytique et protéolytique importante, une tolérance au sel, une croissance plus rapide ou plus lente sur milieu fromage, une capacité plus importante à inhiber des contaminants alimentaires ? Pour cela, j'ai réalisé des expériences au laboratoire.
- 2) Est-ce que l'évolution est répétable ? Pour cela, j'ai comparé les résultats obtenus durant ma thèse sur l'adaptation de *G. candidum* et *P. camemberti* avec les résultats obtenus sur *P. roqueforti* et *S. cerevisiae*. Ces quatre espèces phylogénétiquement éloignées sont adaptées à une même niche

écologique, le fromage, créant ainsi une situation idéale d'adaptation parallèle. On peut tester si l'adaptation au fromage s'est réalisée de manière indépendante et répétée, c'est-à-dire si les mécanismes génomiques impliqués dans l'adaptation sont les mêmes et/ou si les mêmes caractères/gènes/régions génomiques ont été les cibles de la sélection.

Pour répondre à ces différents questionnements, cette thèse sera organisée en deux chapitres présentant successivement les deux organismes modèles, *Penicillium camemberti* et *Geotrichum candidum* puis d'une discussion générale de ces résultats.

# Communications scientifiques

## Valorisation des résultats sous forme d'article et mes contributions

### **Domestication of different varieties in the cheese-making fungus *Geotrichum candidum***

1er auteur, bioRxiv DOI [doi.org/10.1101/2022.05.17.492043](https://doi.org/10.1101/2022.05.17.492043), en révision dans *PCI evol biol.*

Conception, supervision et financement de l'étude : Jeanne Ropars

Production des données génétiques : **Bastien Bennetot**, Vincent Perkins, Jeanne Ropars et Alodie Snirc

Prêt de souches biologiques : Cécile Grondin du CIRM-Levure INRAE

Analyse des données génomique : **Bastien Bennetot**, Jean-Philippe Vernadet, Ricardo C. Rodríguez de la Vega et Samuel O'Donnel

Expérience de phénotypage : **Bastien Bennetot**, Sophie Hautefeuille, Jeanne Ropars et Alodie Snirc

Collaboration pour les analyses de lipolyse et de protéolyse : Marie-Hélène Lessard et Steve Labrie

Collaboration pour l'analyse des composés volatils émis sur milieu fromage : **Bastien Bennetot**, Sophie Landaud et Anne-Claire Péron

Interprétation, rédaction et correction du manuscrit : **Bastien Bennetot**, Tatiana Giraud et Jeanne Ropars ainsi que tous les autres auteurs

### **Domestication of the Emblematic White Cheese-Making Fungus *Penicillium camemberti* and Its Diversification into Two Varieties**

4ème auteur, *Current Biology* 2020, DOI : [doi.org/10.1016/j.cub.2020.08.082](https://doi.org/10.1016/j.cub.2020.08.082)

Conception, supervision et financement de l'étude : Jeanne Ropars et Tatiana Giraud

Production des données génétiques : Jeanne Ropars, Estelle Didiot, Alodie Snirc et Stéphanie le Prieur

Analyse des données génomique : **Bastien Bennetot**, Jeanne Ropars et Ricardo C. Rodríguez de la Vega.

Expérience de phénotypage : Jeanne Ropars, Estelle Didiot et Alodie Snirc

Collaboration analyse des mycotoxines et gènes de biosynthèse : Monika Coton, Emanuel Coton et Elisabeth Poirier

Analyse des données : Jeanne Ropars et Estelle Didiot

Rédaction et correction du manuscrit : Jeanne Ropars, Tatiana Giraux et Ricardo C. Rodríguez de la Vega

### **La domestication des champignons *Penicillium* du fromage**

4ème auteur, Comptes Rendus Biologies 2020, DOI : [doi.org/10.5802/crbiol.15](https://doi.org/10.5802/crbiol.15)

Rédaction de la partie correspondant à la domestication de *P. camemberti* : **Bastien Bennetot**, Jeanne Ropars et Tatiana Giraud

Rédaction de la partie correspondant à la domestication de *P. roqueforti* : Jeanne Ropars, Thibault Caron, Ying-Chu Lo et Tatiana Giraud

### **Valorisation des résultats lors de séminaire**

- Présentation orale au “2nd Symposium on Fungal Genetics, Ile de France”, novembre 2021
- Présentation d’un poster au “Congress of the European Society for Evolutionary Biology” (ESEB) à Prague, août 2022

Chapitre 1 Domestication de *Penicillium*  
*camemberti*



## I. Résumé

*Penicillium camemberti* est un champignon filamenteux de couleur blanche et à l'aspect duveteux que l'on retrouve sur les fromages à pâte molle tel que le camembert, le brie mais aussi la rigotte de Condrieu et le Saint-Marcellin. Le camembert de Normandie est un fromage protégé par une appellation d'origine protégée (AOP) qui indique dans son cahier des charges : "La croûte est fine, elle présente des moisissures superficielles constituant un feutrage blanc (*Penicillium candidum* [*camemberti*]) pouvant selon le degré d'affinage laisser apparaître des pigmentations rouges (*Brevibacterium linens*)". Jusqu'au XX<sup>ème</sup> siècle, le brie avait une croûte orangée recouverte d'une moisissure gris-verte. On peut le voir sur la peinture du XIX<sup>ème</sup> siècle intitulée "Symphonie des fromages en brie majeur ; Nature morte au fromage" de Marie Jules Justin (figure 14; 1847).



**Figure 14 :** "Symphonie des fromages en brie majeur ; Nature morte au fromage", peinture de Marie Jules Justin (1847). On y observe un brie orangé et recouvert d'une moisissure gris-verte.

Pourtant aujourd'hui, les camemberts et les bries sont blancs. C'est au début du XX<sup>ème</sup> siècle que Georges Roger isole le "*Penicillium candidum*" (*P. camemberti*) à partir d'un brie avant de revendre cette souche aux fromagers. Il le fait à la demande des fromagers de la brie qui ont des problèmes de contamination par des champignons non voulus et des productions inconsistantes. En leur conseillant de stériliser leur cave avant d'ensemencer avec du *Penicillium camemberti*, Georges Roger garantit une meilleure fiabilité de la production de brie. Au fur et à mesure, la méthode est si efficace qu'elle se démocratise jusqu'en Normandie pour la production du camembert. C'est ainsi que progressivement, par l'utilisation d'un mutant blanc, nommé *P. camemberti*, le camembert et le brie deviennent blancs.

Les données historiques mentionnent que *P. camemberti* a été isolé à partir d'un champignon gris-vert nommé *P. commune*. Les souches indiquées comme étant *P. commune* ont aujourd'hui été identifiées comme appartenant à deux espèces distinctes : *P. biforme* et *P. fuscauglaucum* (Giraud et al. 2010). *Penicillium camemberti* provient très certainement d'un *P. biforme* car ce champignon en plus d'être spécifique du milieu fromager est l'espèce la plus proche de *P. camemberti* (Pitt 1979). *Penicillium biforme* est une moisissure gris-verte parfois inoculée ou qui colonise naturellement la croûte de certains fromages comme les crottins de chèvre. *Penicillium fuscoglaucum* est la seconde espèce la plus proche de *P. camemberti* mais contrairement à *P. biforme*, celle-ci est retrouvée dans le milieu naturel. *Penicillium fuscoglaucum* a une niche écologique encore peu connue bien que cette espèce ait été isolée à partir de caoutchouc, de racines, de bois, de noix, d'eau ou d'aliments contaminés. L'analyse de phénotypage et de génomique des populations de ce complexe d'espèce a mené à la rédaction d'un article (participation aux

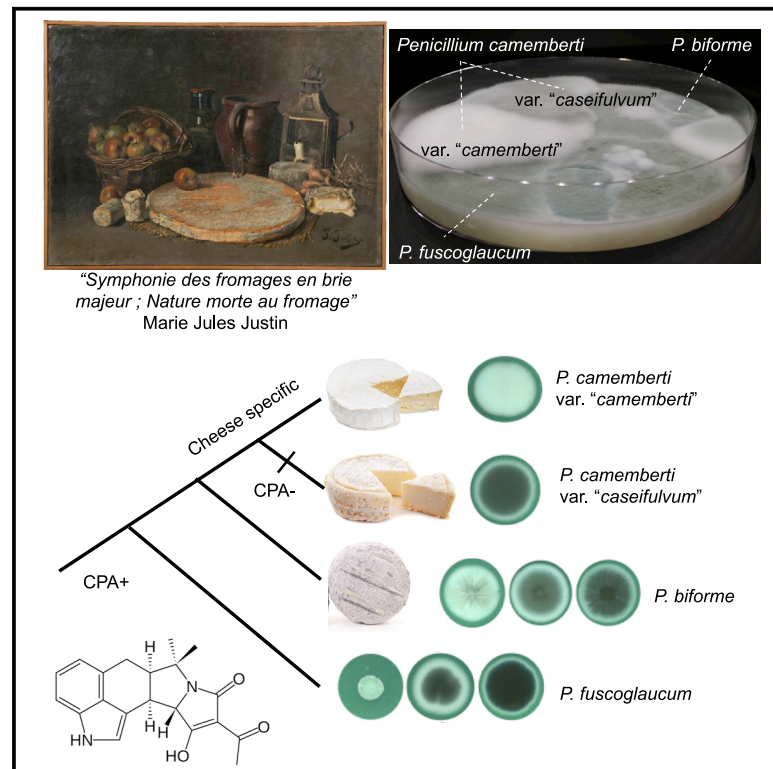
analyses bioinformatiques) présenté en partie II et d'une sous partie d'une revue (participation à l'écriture de la sous-partie sur *P. camemberti*) en partie III.

II. Article : Domestication of the Emblematic White  
Cheese-Making Fungus *Penicillium camemberti* and Its  
Diversification into Two Varieties

# Current Biology

## Domestication of the Emblematic White Cheese-Making Fungus *Penicillium camemberti* and Its Diversification into Two Varieties

### Graphical Abstract



### Authors

Jeanne Ropars, Estelle Didiot, Ricardo C. Rodríguez de la Vega, ..., Alodie Snirc, Stéphanie Le Prieur, Tatiana Giraud

### Correspondence

jeanne.ropars@universite-paris-saclay.fr

### In Brief

Ropars et al. show that the white and fluffy mold *Penicillium camemberti* used for soft cheese making results from domestication by divergence from its gray-green, cheese-specific relative *P. biforme*. Two *P. camemberti* varieties occur and display contrasted phenotypic features in terms of color, growth, toxin production, and competitive ability.

### Highlights

- The molds used in soft cheeses result from multiple domestication steps
- The clonal and fluffy *Penicillium camemberti* displays strong signs of domestication
- Two varieties occur in *P. camemberti*, with contrasted phenotypic features

Article

# Domestication of the Emblematic White Cheese-Making Fungus *Penicillium camemberti* and Its Diversification into Two Varieties

Jeanne Ropars,<sup>1,3,\*</sup> Estelle Didiot,<sup>1</sup> Ricardo C. Rodríguez de la Vega,<sup>1</sup> Bastien Bennetot,<sup>1</sup> Monika Coton,<sup>2</sup> Elisabeth Poirier,<sup>2</sup> Emmanuel Coton,<sup>2</sup> Alodie Snirc,<sup>1</sup> Stéphanie Le Prieur,<sup>1</sup> and Tatiana Giraud<sup>1</sup>

<sup>1</sup>Ecologie Systematique et Evolution, Université Paris-Saclay, CNRS, AgroParisTech, 91400 Orsay, France

<sup>2</sup>Univ Brest, Laboratoire Universitaire de Biodiversité et Écologie Microbienne, 29280 Plouzané, France

<sup>3</sup>Lead Contact

\*Correspondence: [jeanne.ropars@universite-paris-saclay.fr](mailto:jeanne.ropars@universite-paris-saclay.fr)

<https://doi.org/10.1016/j.cub.2020.08.082>

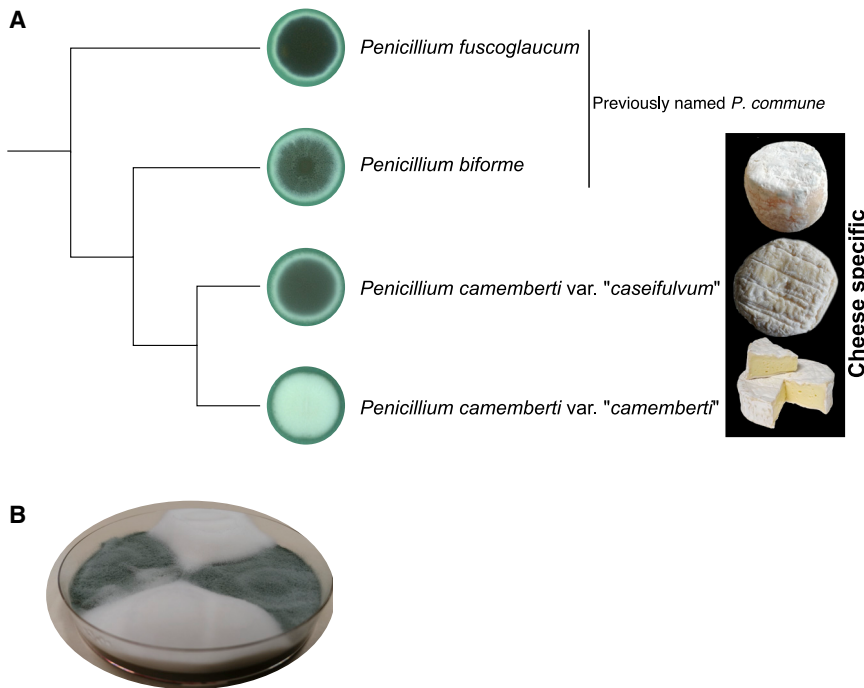
## SUMMARY

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. bifforme* 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. bifforme*. *Penicillium bifforme* 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 amounts 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, bottlenecks, 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 effects 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, and pig) [1–3] and plants (e.g., maize and apricot) [4, 5] have improved our understanding of adaptive divergence. Maize, for example, has undergone major changes in phenotype compared with its wild relative (teosinte), including a decrease in tillering and the development of larger, non-dehiscent grains [6]. 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. Such notable adaptation is often associated with a decrease in fitness in natural environments, for example, a decrease in or loss of sexual reproduction ability in bulldogs [7] and bananas [8].

Fungi are excellent models for studying evolution and adaptation in eukaryotes, given their many experimental assets [9], including their small genomes and tractability for laboratory experiments. Humans have domesticated several fungi for the fermentation of foods (e.g., for beer, bread, wine, and cheese) [10]. Despite their economic importance, fungi used by humans have been little studied, with the exception of the budding yeast *Saccharomyces cerevisiae* used for beer, wine, and bread production [11–21]; the filamentous fungus *Aspergillus oryzae* used to ferment soy and rice products in Asia [22–24]; and the blue-cheese mold *Penicillium roqueforti* [25–27]. Whole-genome analyses have revealed that *P. roqueforti* has been domesticated twice, 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 [25–28]. The Roquefort population displays some genetic diversity and harbors beneficial traits for cheese production before the industrial era. 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



**Figure 1. Cheeses and Fungi used for Cheese Making**

(A) Schematic representation of relationships between *Penicillium camemberti* and its closest relatives, *P. bifforme* and *P. fuscoglaucum*. Pictures of *P. bifforme*, *P. camemberti* var. "camemberti," and *P. camemberti* var. "caseifulvum" on salted cheese medium at 25°C after 10 days of growth are shown.

(B) Petri dish with a white and fluffy *P. camemberti* var. "camemberti" strain (FM013) and a green rough *P. fuscoglaucum* strain (LCP06617) on malt medium at 25°C.

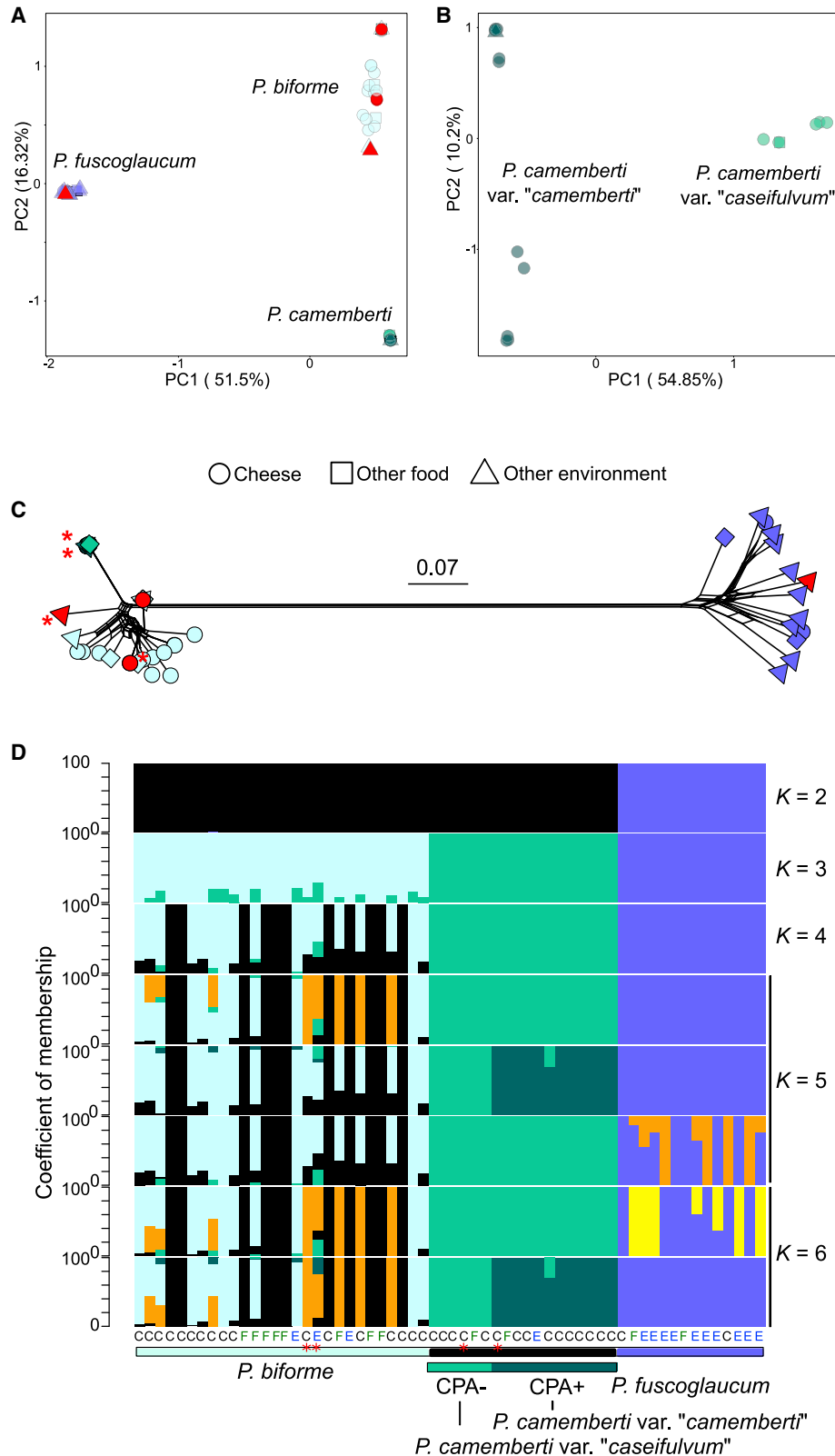
*P. commune* name [36]. *Penicillium caseifulvum* has also sometimes been advocated to constitute a separate species in this clade [36], on the basis of colony morphology and the lack of production of cyclopiazonic acid (CPA), a mycotoxin known to be produced by *P. camemberti*. However, a study based on a few genetic markers was unable to differentiate these putative *P. caseifulvum* strains from *P. camemberti* [35].

traits have been conferred by two large horizontally transferred genomic regions, *Wallaby* and *CheesyTer* [26, 27], in the non-Roquefort cheese population. A cheese-specific population also occurs in *S. cerevisiae*, differentiated from the populations used for alcohol or bread production, assimilating galactose more rapidly than *S. cerevisiae* populations thriving in other food environments (e.g., beer, bread, and wine) or in natural environments (oak) [29].

The white mold *Penicillium camemberti* is used for the maturation of soft cheeses, such as Camembert, Brie, and Neufchatel, and 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 20<sup>th</sup> century [30] and cultured clonally ever since. The first records of Brie cheeses date from Middle Ages [31], being of blue aspect until the middle of the 20<sup>th</sup> century [32, 33], as illustrated by a 19<sup>th</sup> century painting by Marie Jules Justin entitled "Symphonie des Fromages en Brie Majeur - Nature Morte au Fromage" and a picture from the record "Fromages de France" published in 1953 [31]. 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. bifforme*, *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 [34] 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. bifforme* and *P. fuscoglaucum* [35]. However, the most recent taxonomic reference study of the *Penicillium* genus maintained the

The taxonomy of this clade thus remains unclear, and population structure has not been studied with powerful genetic markers. Differences in phenotypic traits have not been extensively studied either. Humans might have selected strains for specific traits beneficial for cheese production, such as faster growth in cheese, attractive color, and a greater ability to compete against food-spoiling micro-organisms, but without the production of detrimental mycotoxins. Mycotoxins are natural secondary metabolites produced by fungi under certain conditions that can inhibit competitor growth and which can have a toxicological impact on humans [37]. Cyclopiazonic acid is one of these toxic secondary metabolites produced by *P. camemberti* under certain conditions [37, 38].

Our goal here was to investigate the footprints of domestication by assessing genetic and phenotypic differentiation, by sequencing the genomes of strains isolated from various environments. We performed a whole-genome-based analysis of population structure, which revealed that *P. bifforme*, *P. camemberti*, and *P. fuscoglaucum* formed separate and specific genetic clusters, *P. bifforme* and *P. camemberti* being sister clades (Figure 1A). *Penicillium camemberti* and *P. bifforme* were both specific to the cheese environment, whereas *P. fuscoglaucum* was mostly found in natural environments. These findings suggest an ancient domestication event separating *P. bifforme* from *P. fuscoglaucum* and a more recent domestication event separating the *P. camemberti* clonal lineage from *P. bifforme*. Consistent with this scenario, we found evidence of phenotypic adaptation to cheese making in *P. bifforme* compared with *P. fuscoglaucum*, with a whiter color, faster growth on cheese medium under cave conditions, and lower amounts of toxin production. We also reveal the diversification of *P. camemberti* into two varieties, *P. camemberti* var. "camemberti" and var. "caseifulvum," also displaying genetic and phenotypic differentiation and maybe used for the production of different cheese types.



**Figure 2. Population Structure of 61 Strains from the *Penicillium camemberti* Species Complex, Based on Whole-Genome Data**

(A) Principal-component analysis (PCA) based on the 61 strains.

(B) PCA based on only *P. camemberti* strains ( $n = 19$ ).

(legend continued on next page)



## RESULTS

### *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, and blue cheeses), 11 from other food products (e.g., dry sausages and 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* (Table S1). We resequenced the reference genome of *P. camemberti* (LCP06093, also known as FM013, initially sequenced with 454 technology) [26] with PacBio long-read technology and used the ca. 35-Mb assembly obtained for mapping. We identified 392,072 SNPs across all strains (Data S1). 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* ( $n = 14$ ). NGSadmix 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 S2).

A recent study described rapid phenotypic change from green-gray to white in a “wild *P. commune*” strain [39], interpreting these changes as evidence that domestication can occur within 4 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 or 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 3). The *P. commune* neotype strain (LCP05531) and the *P. biforme* type strain (LCP05529) belonged to the same clade (Figures 2C, 2D, and 3), further indicating that *P. commune* and *P. biforme* overlap. In contrast, the *P. caseifulvum*-type strain (LCP05630) and the *P. camemberti*-type strain belonged to different clusters at  $K$  values  $\geq 5$  in the population structure analysis (red stars in Figure 2C), suggesting that they do correspond to different genetic clusters.

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

*Penicillium fuscoglaucum* displayed the highest genetic diversity level, followed by *P. biforme*, as indicated by the number of SNPs,  $\pi$ , and Watterson's  $\theta$  (Table S2), as well as strain scattering in PCA and neighbor-net network (Figures 2A and 2C). The very low genetic diversity detected within *P. camemberti* suggests that it represents a clonal lineage (Table S2; Figures 2A and 2C). The average nucleotide identity between *P. camemberti* and *P. biforme* was very high (99.8%) and similar to the 99.5% identity between the domesticated mold *Aspergillus oryzae* and its wild relative *A. flavus* [40].

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

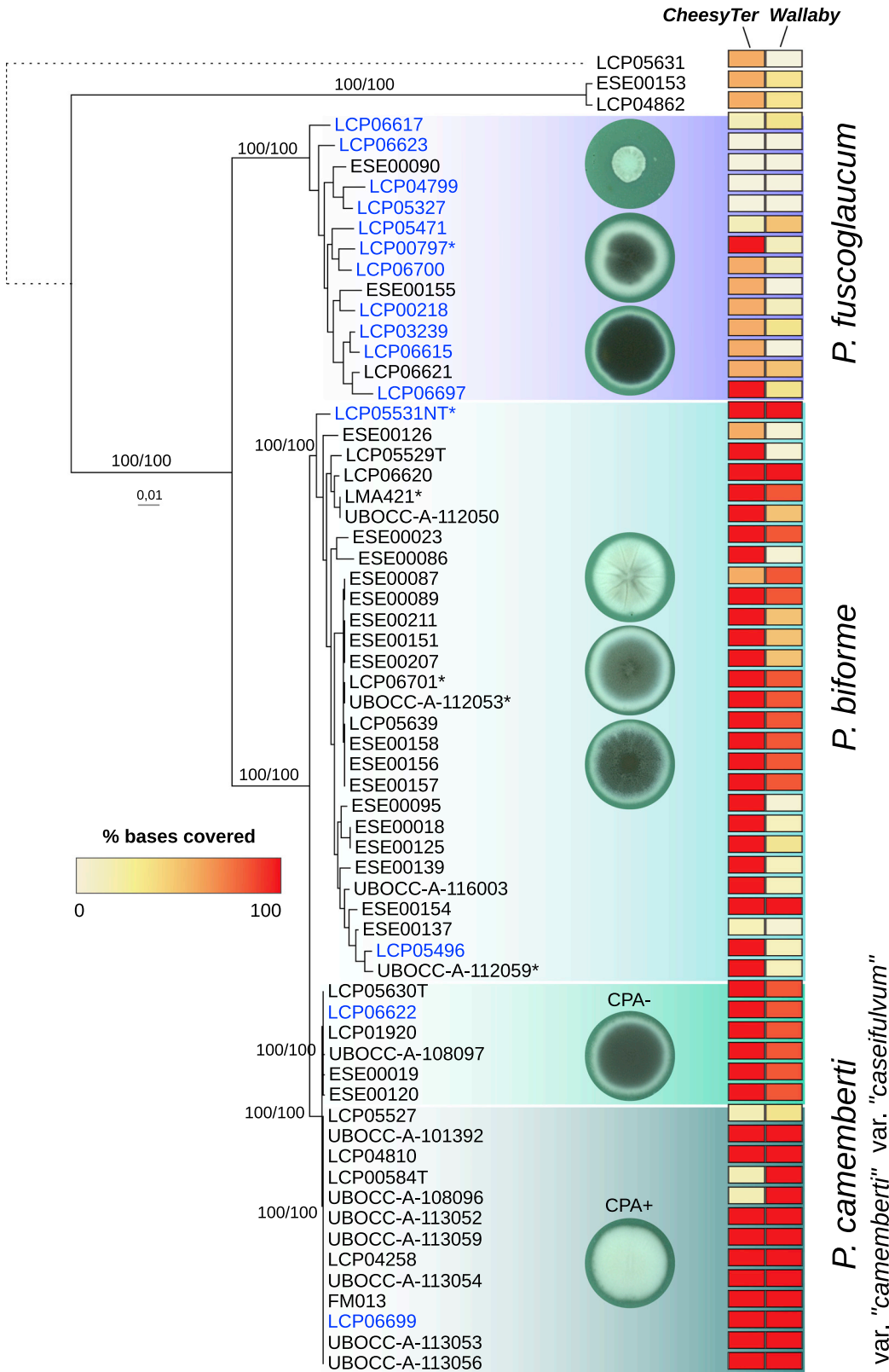
In heterothallic fungi, sexual reproduction occurs only between two haploid individuals carrying different mating-type alleles (MAT1-1 and MAT1-2 in heterothallic ascomycetes). A lack of sexual reproduction leads to relaxed selection on the mating-type genes and a departure from a balanced ratio between mating types. Neither mating-type allele presented any evidence of loss-of-function mutations or SNPs in any of the genomes. We found a significant departure from the 1:1 ratio in *P. biforme* ( $\chi^2 = 14.29$ ;  $df = 1$ ;  $p < 1e^{-3}$ ), with MAT1-2 strains predominating, but not in *P. fuscoglaucum* ( $\chi^2 = 1.14$ ;  $df = 1$ ;  $p = 0.29$ ). All 19 *P. camemberti* strains carried the MAT1-2 allele, providing further evidence of the clonality of *P. camemberti* [41, 42]. The observed clonality in *P. camemberti* might be explained by the recent selection of a white mutant and its use in industrial cheese production [43], whereas *P. biforme* is more widely used in the production of artisanal cheeses and farmers might make use of diverse local strains.

(C) Neighbor-net analysis based on SNP data. Cross-linking indicates the likely occurrence of recombination. Branch lengths are shown, and the scale bar represents 0.07 substitutions per site.

(D) Population subdivision inferred from  $K = 2$  to  $K = 6$ . Each column represents a strain, and colored bars represent their coefficients of membership for the various gene pools.

For all panels, genetic clusters are represented by the same colors: dark purple for *P. fuscoglaucum*; light blue for *P. biforme*; dark green for *P. camemberti* var. “*camemberti*”; and light green for *P. camemberti* var. “*caseifulvum*”. The strains identified as *P. commune* (non-monophyletic group) by collectors are indicated in red. For PCA and neighbor-net analysis, symbols correspond to the environment of collection: circles for cheese; triangles for dried sausages; and squares for other environments. On the structure analysis, letters below the plots indicate the origin of the strains: C for cheese; F for food strains other than cheese; and E for other environments. Red asterisks indicated the type strains of *P. biforme*, *P. camemberti*, and *P. caseifulvum*.

See also Tables S1 and S2, Figures S1 and S2, and Data S3 and S4.



(legend on next page)

### Two Genetic Clusters within *Penicillium camemberti*

We found two genetic clusters within *P. camemberti*, as indicated by the population structure analysis implemented in NGSadmix at  $K = 5$ , the PCA, and the neighbor-net without *P. fuscoglaucum*, as well as the PCA with only *P. camemberti* (Figures 2B, 2D, and S2). The two *P. camemberti* genetic clusters displayed very few shared SNPs (13% of the 19,510 SNPs) (Table S2) but a very high nucleotide identity (99.998%). The two genetic clusters within *P. camemberti* separated strains isolated from soft cheeses, such as Camembert, and strains isolated from other kinds, such as Rigotte de Condrieu and Saint Marcellin. The *P. caseifulvum*-type strain (red star in Figure 2D) clustered with *P. camemberti* strains isolated from non-soft cheeses. The genetic cluster including the *P. caseifulvum*-type strain will be referred to hereafter as *P. camemberti* var. “*caseifulvum*”, and the genetic cluster including the *P. camemberti*-type strain as *P. camemberti* var. “*camemberti*”. The nucleotide diversity was very low within both varieties (Table S2; Figure 2B). We designed two pairs of primers to identify by a single PCR reaction the *P. camemberti* variety, each yielding a specific amplicon size.

The population structure analysis also suggested further genetic subdivision within each of *P. fuscoglaucum* and *P. biforme*; these subdivisions, however, did not correspond to the environment from which the isolate was obtained, type of cheese, or to a strong subdivision in the PCAs or the neighbor-nets (Figures 2 and S1).

### Presence-Absence Polymorphism of the Two Horizontally Transferred Regions Likely to Be Advantageous in Cheese

The PacBio-based genome assembly of the FM013 reference genome showed that *Wallaby* formed a single block (scaffold 18; positions 102,711–538,265). Two genes in *Wallaby* have been highlighted as putatively important for cheese making. One, encoding the *Penicillium* antifungal protein (PAF) [26], was absent from FM013, whereas the other, *Hce2*, encoding a protein with antimicrobial activities, was present. All *P. camemberti* var. “*camemberti*” strains harbored the same *Wallaby* region as FM013 (Figures 3 and S3B), whereas strains from the *P. camemberti* var. “*caseifulvum*” lineage carried fragmented and duplicated segments of the *Wallaby* present in FM013, and *Hce2* was systematically present. We found *Wallaby* fragmented with some duplicated regions in 18 *P. biforme* strains and completely absent from 10 strains (Figures 3 and S3B). *Wallaby* appeared to be even more fragmented in *P. fuscoglaucum* strains, always lacking *Hce2* but in one strain, and completely absent in nine strains. Fragments of *Wallaby* were found in some non-cheese *P. fuscoglaucum* strains. Even in *P. fuscoglaucum*, no genetic variability could be detected within the *Wallaby* fragments, indicating that they have also likely been acquired by horizontal gene transfer in this species. So far,

we found no species in which *Wallaby* showed genetic variability, indicating that we have not identified the donor species.

We found no strict association between *Wallaby* and the 57 kb *CheesyTer* region, the horizontally transferred region thought to be beneficial for growth on cheese [27]. In *P. camemberti*, 16 of 19 strains carried the whole *CheesyTer* region, with an identity of 100% (Figures 3 and S3A) (scaffold 17; positions 690,636–751,978 in FM013). The other three *P. camemberti* strains, belonging to the *P. camemberti* var. “*camemberti*” cluster, completely lacked *CheesyTer* and an 80-kb-downstream fragment. These three strains were isolated from Camembert cheeses around 1905, suggesting that the horizontal transfer of *CheesyTer* to *P. camemberti* occurred after 1905, in a single strain that has been since clonally cultured. In *P. biforme*, 25 of 28 strains carried *CheesyTer*, with an identity >99.8% between strains. The identity between *P. biforme* and *P. camemberti* was also >99.8%, except in the 5-kb region at the start of *CheesyTer*, in which the similarity between the two species dropped to 90%, mainly due to C:G to T:A mutations in *P. camemberti*, corresponding to the repeat-induced point (RIP) mutations specific to fungi occurring in repeated sequences during sexual reproduction [44]. As *P. camemberti* was found to be clonal and no RIP footprints were detected within the rest of *CheesyTer*, these findings suggest that *CheesyTer* was transferred to an ancestor of the two *P. camemberti* varieties with this RIPed sequence already present. The two *CheesyTer* genes involved in lactose metabolism [27], one encoding a lactose permease and the other a  $\beta$ -galactosidase, were present in all strains carrying at least one part of *CheesyTer* (Figures 3 and S3A).

### Looking for Genomic Footprints of Adaptation in *P. camemberti* and *P. biforme*

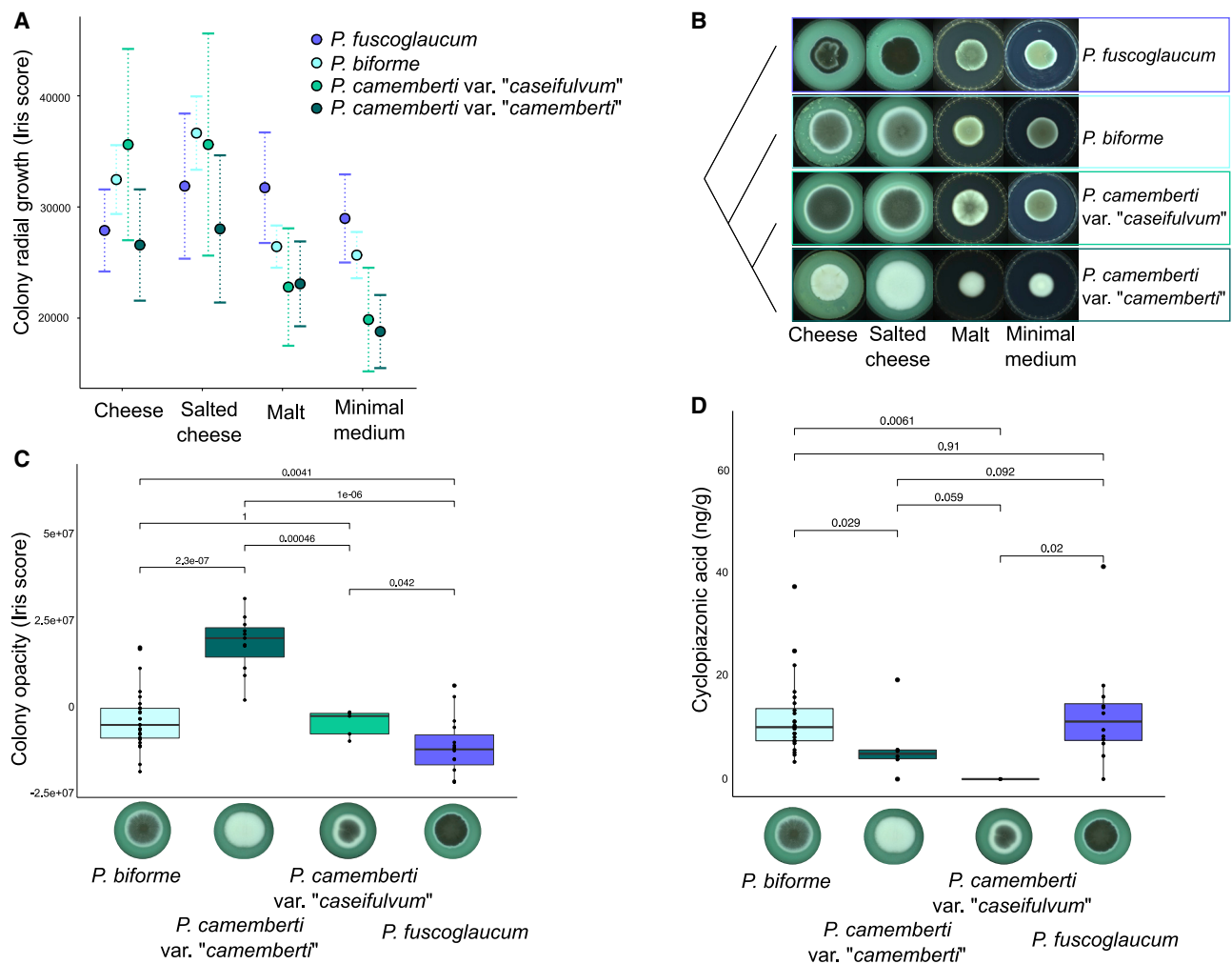
Not many genomic selection scans could be performed given the high level of clonality in *P. camemberti*, as any selective sweep will hitchhike the whole genome. We investigated whether some genomic regions showed greater differentiation or higher diversity than the genomic background when comparing *P. biforme* to either *P. camemberti* or *P. fuscoglaucum*. No particularly striking genomic island of differentiation based on  $d_{xy}$  or diversity based on  $\pi$  and Watterson's  $\theta$  could be identified (Data S3 and S4). The genes within windows in the 1% highest  $d_{xy}$  were not enriched in any particular function but contained significantly less genes with predicted functions than the rest of the reference genome, suggesting that these regions contained rapidly evolving genes.

### Phenotypes Better Suited for Cheese Production in *P. camemberti* and *P. biforme* Than in the Closely Related Wild *P. fuscoglaucum*

We performed a set of laboratory experiments to determine whether cheese strains had evolved traits beneficial for cheese

### Figure 3. Phylogenetic Relationships and Presence/Absence of the *Wallaby* and *CheesyTer* Horizontally Transferred Regions in the 61 Studied Strains of the *Penicillium camemberti* Species Complex and the Three Outgroups

Maximum-likelihood tree for the 64 strains based on 1,124,983 SNPs (586,131 parsimony informative sites). Outgroups are the following: *P. cavernicola* (LCP05631) and *P. palitans* (ESE00153 and LCP04862). Branch supports, represented as SH-aLRT/ultrafast bootstrap in percentage, are only shown for clades but were always >75%. The strain name color indicates the environment of collection: black for food and blue for other environments. Asterisks after strains indicate strains previously named *P. commune*. On the right panel, heatmap representing the percentage of bases covered in *CheesyTer* and *Wallaby* regions is shown. See also Tables S1 and S2, Figures S1–S3, and Data S1, S3, and S4.



**Figure 4. Phenotypic Traits Distinguishing the Four Lineages *Penicillium biforme*, *P. camemberti* var. "caseifulvum", *P. camemberti* var. "camemberti", and *P. fuscoglaucum***

(A) Mean radial growth of the four lineages on unsalted cheese, salted cheese, and malt and minimal media.

(B) Pictures of colonies on unsalted cheese, salted cheese, and malt and minimal media; *P. fuscoglaucum*: LCP05471; *P. camemberti* var. "camemberti": UBOCC-A-113052; *P. biforme*: ESE00211; *P. camemberti* var. "caseifulvum": LCP06622. Vertical dotted lines represent standard deviation bars.

(C) Difference in opacity between the four lineages on salted cheese at 25°C.

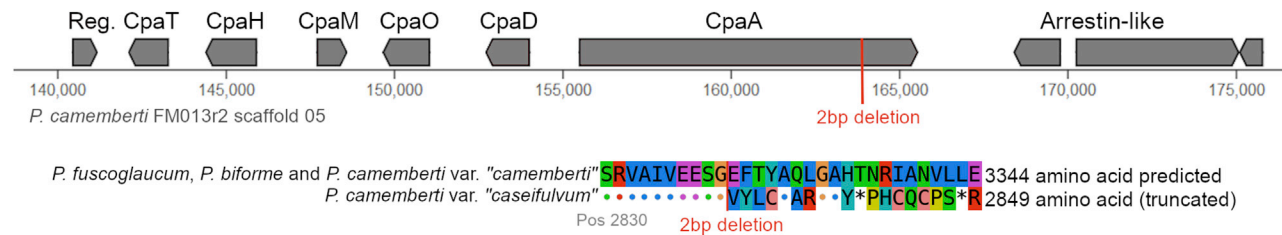
(D) Difference in production of cyclopiazonic acid between the four lineages on yeast extract sucrose.

(C and D) Horizontal lines of the boxplots represent the upper quartile, the median, and the lower quartile. Each dot represents a strain, and p values are given for the tests of differences between the lineages (Wilcoxon test). The color indicates assignments to the four lineages as in other figures. See also [Table S1](#) and [Data S1](#) and [S2](#).

production not present in the closely related wild *P. fuscoglaucum* lineage occurring in other environments. In each of *P. camemberti* and *P. biforme*, we considered as a single group the strains from cheese and other food types (e.g., dry sausage and bread) because they were not genetically differentiated and the same strains are sold by producers for cheese and dry sausage maturation.

We first investigated whether *P. camemberti* and *P. biforme* had acquired traits enabling them to grow more rapidly than *P. fuscoglaucum* in cheese making conditions. Rapidly growing fungi can be beneficial for cheese making, as they are more capable of excluding spoiler bacteria, yeasts, and molds [10] and promote faster cheese maturation. Cheese maturation occurs in dark caves at a low temperature and high humidity. We

therefore grew 61 strains on various media (unsalted cheese, salted cheese, and minimal and malt media), in the dark and at different temperatures and humidity levels (25°C at ambient relative humidity, 10°C at 98% humidity, or 8°C at 85% humidity), and we measured colony radial growth after 10 days. The non-parametric Kruskal-Wallis test and the ANOVA showed no significant effect on growth of the substrate from which the isolate was obtained (i.e., food versus non-food). We found a significant effect on growth of the culture media, species, temperature/humidity conditions, and the *P. camemberti* variety. The ANOVA further showed significant effects on growth of the interactions between culture medium and species on the one hand and between culture medium and temperature/humidity on the other hand ([Data S1A](#) and [S2A](#); [Figure 4A](#)). However, the significance



**Figure 5. Cyclopiazonic Acid (CPA) Biosynthesis Cluster in the *Penicillium camemberti* Species Complex**

(A) Schematic representation of the CPA biosynthesis cluster in the *P. camemberti* var. “*camemberti*” reference genome FM013.

(B) Amino acid alignment of the *CpaA* gene showing a 2-bp deletion in strains belonging to the *P. camemberti* var. “*caseifulvum*” genetic cluster, leading to a truncated protein.

See also [Table S1](#) and [Data S1](#) and [S2](#).

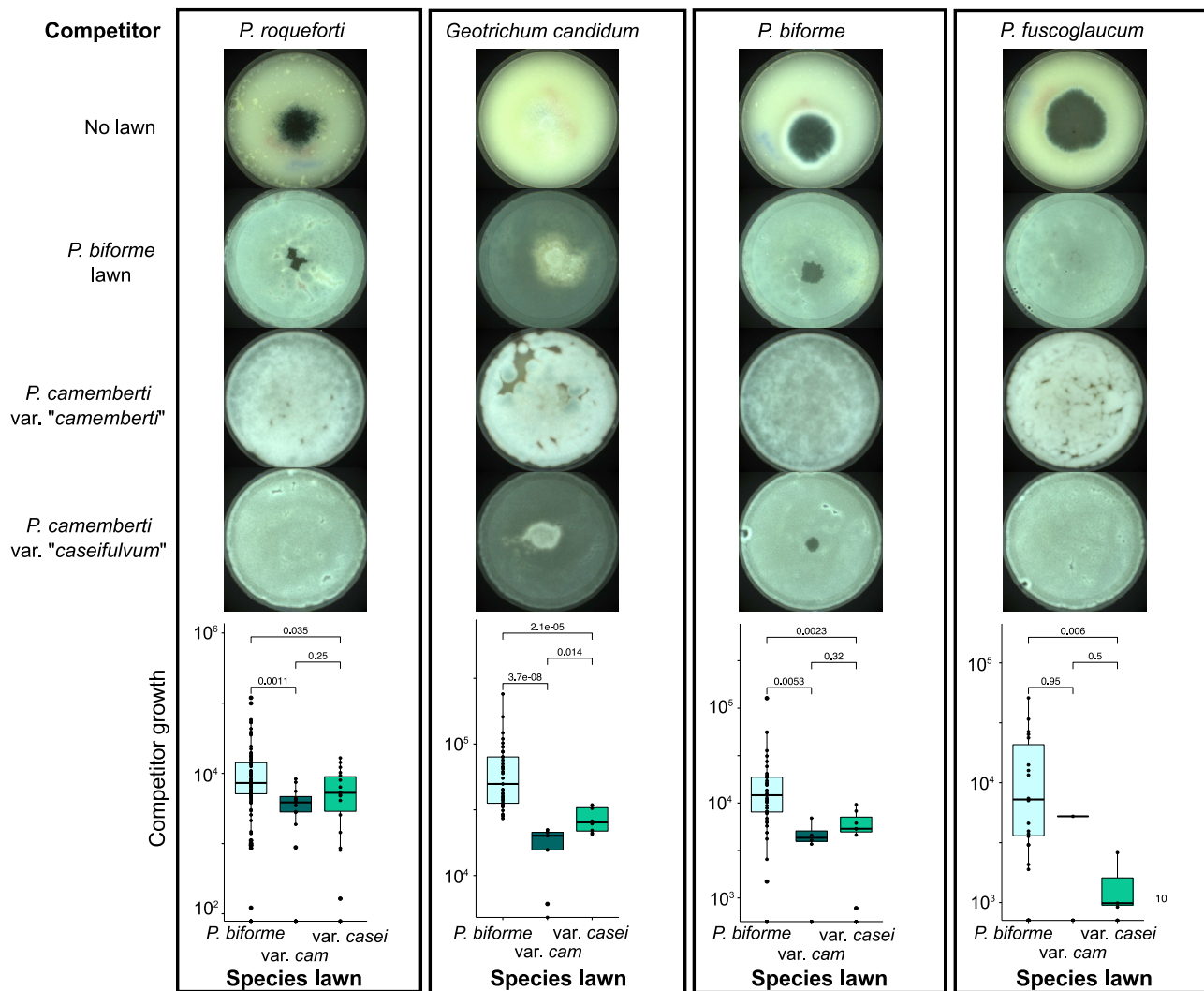
of the interaction between culture medium and temperature/humidity conditions ([Data S1A](#) and [S2A](#)) reflected a very slow growth at low temperatures, making it difficult to detect differences between media. Although the two *P. camemberti* varieties were genetically very closely related, *P. camemberti* var. “*caseifulvum*” grew similarly to *P. biforme* (post hoc Tukey test  $p = 1$ ) and had greater radial growth on unsalted cheese and salted cheese and slower growth on malt and minimal media than *P. fuscoglaucum*. By contrast, *P. camemberti* var. “*camemberti*” displayed weaker radial growth than any other lineage on all media ([Figure 4A](#)) but was fluffier ([Figure 1B](#)). Furthermore, *P. camemberti* and *P. biforme* grew less rapidly than *P. fuscoglaucum* on minimal and malt media, indicating a disadvantage for growth in harsh conditions, as expected for cheese strains because of relaxed selection on functions useful only in wild environments. These findings demonstrated that *P. camemberti* var. “*caseifulvum*” and *P. biforme* have acquired rapid radial growth in cheese that was likely beneficial for fast maturation. *Penicillium camemberti* var. “*camemberti*” grew much less radially on cheese than any other lineage but grew much more vertically (i.e., being fluffier), and the mycelium grew up to the lid of the Petri dishes ([Figure 1B](#)). *Penicillium camemberti* var. “*camemberti*” might have been selected for fast vertical growth to make soft cheeses with a fluffy aspect within a short maturation period. Short maturation is beneficial for production when using cow or goat milk available throughout the year, as is the case for Camembert cheeses. In contrast, ewe milk is only available a few months a year, so that Roquefort cheeses need to be stored for several months to be sold and consumed throughout the year. Such long storage can explain a selection for slow growth in the Roquefort *P. roqueforti* population before fridges were available [25], in contrast to a selection for fast growth in *P. camemberti* and *P. biforme*.

High salt concentrations in cheeses prevent the growth of contaminants, and cheese fungi might have adapted to such conditions. We found an effect of salt on *P. biforme* growth, with strains growing more rapidly on salted than unsalted cheese medium (post hoc Tukey test  $p = 0.006$ ) and more rapidly on salted cheese medium than *P. camemberti* (post hoc Tukey test  $p = 0.0001$ ). Salt had no effect on the growth of *P. camemberti* (post hoc test  $p = 1$ ) or *P. fuscoglaucum* (post hoc test  $p = 1$ ). There might have been stronger selection for salt tolerance in *P. biforme*, which is used to inoculate blue and goat cheeses, both being more salty than soft cheeses, such as Brie or

Camembert [45], for which *P. camemberti* var. “*camemberti*” is used.

We investigated whether cheese lineages had become whiter, which can be more attractive to some consumers than gray-green mold on cheese, by comparing the opacity of lineages on cheese medium, as opacity increases with the brightness and fluffiness of a colony [46]. We found significant effects of species and variety ([Figure 4C](#); [Data S1B](#) and [S2B](#)), with *P. camemberti* being significantly more opaque (i.e., brighter and fluffier) than other lineages and *P. camemberti* var. “*camemberti*” even more so compared with var. “*caseifulvum*” (post hoc Tukey test  $p = 0.03$ ). This is consistent with the white and fluffy aspect of the crust of Camembert and Brie, made with *P. camemberti* var. “*camemberti*”, whereas *P. camemberti* var. “*caseifulvum*” was found here in cheeses with a grayer and less fluffy crust, such as Saint Marcellin or Rigotte de Condrieu ([Figures 1, 4B](#), and [4C](#)). The *P. camemberti* var. “*caseifulvum*” and *P. biforme* lineages did not differ significantly from each other (post hoc Tukey test  $p = 1$ ), and both were brighter than the wild blue-green *P. fuscoglaucum* ([Figure 4B](#)).

We also investigated whether cheese lineages produced smaller amounts of cyclopiazonic acid (CPA), which is cytotoxic to humans [37, 47], than strains isolated from other environments. *Penicillium camemberti* has been reported to produce CPA on yeast extract sucrose (YES) medium and at very low, non-toxic concentrations in cheese [38]. None of the *P. camemberti* var. “*caseifulvum*” strains tested here produced CPA on YES ([Figure 4D](#)), consistent with findings for the type strain [48]. By contrast, *P. camemberti* var. “*camemberti*”, *P. biforme*, and *P. fuscoglaucum* produced CPA, the highest levels being obtained with *P. biforme* and *P. fuscoglaucum* ([Figure 4D](#); [Data S1C](#) and [S2C](#)). The CPA biosynthesis cluster in the genome appeared to be functional, with its six genes present, but not the *cpaR* gene encoding a regulatory protein ([Figure 5](#)). The only exceptions were the six *P. camemberti* var. “*caseifulvum*” strains, in which a 2 bp deletion in the *cpaA* gene led to a frameshift. The *cpaA* gene encodes a polyketide synthase/non-ribosomal peptide synthase responsible for the first step of the CPA biosynthetic pathway so that a non-functional protein probably prevents CPA production on all substrates. Humans have often selected fungal strains unable to produce harmful toxins for use in food; *Aspergillus oryzae* strains used to ferment Asian food products, which do not produce aflatoxins, whereas its wild relative, *A. flavus*, does [49].



**Figure 6. Competitive Abilities of *Penicillium biforme*, *P. camemberti* var. “*caseifulvum*”, and *P. camemberti* var. “*camemberti*” against *Geotrichum candidum*, *P. biforme*, *P. fuscoglaucum*, and *P. roqueforti* Competitors**

Top: pictures of *P. roqueforti* FM164, *Geotrichum candidum* ESE00163, *P. biforme* ESE00154, and *P. fuscoglaucum* ESE00221 on no lawn (first line), *P. biforme* ESE00086 lawn (second line), *P. camemberti* var. “*camemberti*” LCP04810 (third line), and *P. camemberti* var. “*caseifulvum*” ESE00019 (fourth line) lawns on salted cheese. Bottom: boxplots representing the differences of growth abilities of the competitors on different species lawn are shown. Horizontal lines of the boxplots represent the upper quartile, the median, and the lower quartile. Each dot represents a strain, and p values are given for the tests of differences between the lineages (Wilcoxon test). The color indicates assignments to the three lineages as in other figures. var. “*cam*”, *P. camemberti* var. “*camemberti*”; var. “*casei*”, *P. camemberti* var. “*caseifulvum*”. See also [Table S1](#) and [Data S1](#) and [S2](#).

The *P. roqueforti* non-Roquefort population was found unable to produce mycophenolic acid because of a 174 bp deletion in the *mpaC* gene [50].

Cheese is a nutrient-rich environment in which many micro-organisms can thrive, including undesirable food spoilage organisms. We therefore investigated whether the cheese lineages were better able to outcompete challengers. We allowed strains of *P. camemberti* var. “*camemberti*” (n = 3), *P. camemberti* var. “*caseifulvum*” (n = 2), or *P. biforme* (n = 10) to grow as lawns on the surface of the cheese medium, which were inoculated 24 h later with a single spot, in the middle of the Petri dish, by using a challenger: *Geotrichum candidum* (n = 5); *P. biforme* (n = 6); *P. fuscoglaucum* (n = 5); or *P. roqueforti* (n = 12, 3 strains from

each of the 4 known genetic clusters) [25]. The species and the variety used as the lawn had significant effects on the growth of the challenger (ANOVA and non-parametric Kruskal-Wallis tests) ([Data S1D](#) and [S2D](#); [Figure 6](#)). The two *P. camemberti* lineages prevented the growth of all challengers more effectively than *P. biforme*, and *P. camemberti* var. “*camemberti*” was the most effective. We observed significant differences between the species used as challengers, with *G. candidum* having the highest growth differential on the various lawns, growing less well on a *P. camemberti* var. “*camemberti*” lawn than other lawns. *Geotrichum candidum* is a fungus that is also present on the surface of soft cheeses. The exclusion effect might be mediated by the fluffy morphology, resulting in the fungus

occupying more space and using more resources, and/or by biochemical interactions.

## DISCUSSION

Whole-genome sequences revealed that *P. biforme*, *P. camemberti*, and *P. fuscoglaucum* each formed a separate genetic cluster, with *P. biforme* and *P. camemberti* being sister clades. Strains identified as *P. commune* corresponded to either *P. biforme* or *P. fuscoglaucum*, being thus non-monophyletic (Figure 1A), so this name should not be used any longer. Furthermore, considering all *P. commune* strains to be wild, as it is currently common in the literature, leads to wrong inferences about trait evolution [39], as some *P. commune* strains belong to a domesticated genetic cluster. We found that *P. camemberti* and *P. biforme* were sister clades specific to the cheese environment, whereas the more distantly related clade *P. fuscoglaucum* was found mostly in natural environments. These relationships and the lower diversity of *P. biforme* than *P. fuscoglaucum*, with even lower levels of diversity in *P. camemberti*, suggest that an ancient domestication event led to the divergence between *P. biforme* and *P. fuscoglaucum* and that a much more recent domestication event led to the differentiation between *P. camemberti* and *P. biforme*. We used here the *P. biforme*, *P. camemberti*, and *P. fuscoglaucum* Latin species names as they are commonly used and corresponded to differentiated and monophyletic clades on the basis of genome-wide information. We did not address here the question of species status or definition, as they are notoriously difficult to elucidate among closely related fungal groups, especially in clonal ascomycetes [51, 52].

We found evidence of phenotypic adaptation to cheese making in *P. biforme* in relation to the wild *P. fuscoglaucum* clade, and *P. biforme* had a whiter color, faster growth on cheese medium under cave conditions, and lower amounts of toxin production. These signs of domestication were even more marked in *P. camemberti* than in *P. biforme*. We observed a similar evolution of traits as in populations of the domesticated blue-cheese fungus *P. roqueforti*, which grows more rapidly on cheese medium, is more competitive against contaminant micro-organisms, and grows less well under harsh conditions than wild strains [25, 27]. Such convergent evolution under similar selective pressures suggests that evolution may repeat itself, as already reported in natural populations of *Anolis* lizards [53, 54], three-spined sticklebacks [55], or Mexican cavefishes [56]. Phenotypic convergent evolution has been also reported in domesticated organisms, particularly in crop plants, for the loss of seed shattering, minimization of seed dormancy, and increase in seed size and number [57, 58].

We identified two genetically differentiated *P. camemberti* varieties, *P. camemberti* var. “*camemberti*” and *P. camemberti* var. “*caseifulvum*”, with contrasting phenotypic features, used in the production of different kinds of cheese. The *P. camemberti* var. “*camemberti*” strains were white and were isolated from Camembert or Brie. Their radial growth was slower, but their mycelia were fluffier than the other *Penicillium* strains tested, and they produced similar amounts of CPA compared with *P. biforme* strains. They excluded fungal competitors effectively, as previously suggested [59], probably due to their fluffy morphology, taking up the available space and monopolizing resources. *Penicillium*

*camemberti* var. “*caseifulvum*” strains were gray-green, unable to produce CPA, and were isolated from cheeses other than Camembert, such as St. Marcellin or Rigotte de Condrieu. They displayed more rapid radial growth than *P. camemberti* var. “*camemberti*” and grew similarly to *P. biforme* strains on cheese medium in cave conditions. The difference in terms of cheese type usage between the two varieties should be confirmed by using more strains, but it seems to correspond well to the contrasted identified phenotypes and, in particular, color and fluffiness. The existence of two genetically and phenotypically differentiated lineages with possibly different uses in cheese making suggested that these lineages emerged as a result of different human selection pressures, as reported for the domesticated blue-cheese fungus *P. roqueforti* [25], wine-making yeasts [11, 14], maize, rice, tomatoes, dogs, chickens, and horses [1–6].

*Penicillium camemberti* var. “*camemberti*” is the emblematic fungus used to inoculate soft cheeses, such as Camembert and Brie. According to the technical specifications for Brie de Meaux and Brie de Melun PDOs and Camembert, the crust of the cheese must be white and fluffy, and inoculation with *P. candidum* (a synonym of *P. camemberti*) is even specified for Brie de Meaux. These specifications are recent (20<sup>th</sup> century) and seem to have had a negative impact on the diversity of the fungi used for making these kinds of cheeses, as a single clonal lineage now predominates. The genetic diversity of *P. roqueforti* has also been greatly reduced by recent industrialization [25, 28], although PDO specifications that local strains must be used have protected diversity to some extent in Roquefort cheeses. Camembert and Brie cheeses were gray-green until the middle of the 20<sup>th</sup> century [31, 33, 43]. First mentions of Brie cheeses date from Middle Ages [31]. These historical records are consistent with our findings, which suggest a first domestication event leading to the emergence of the gray-green mold *P. biforme*, subsequently followed, around 1900, by the domestication of *P. camemberti*, with the selection of two different varieties displaying more marked signs of domestication than *P. biforme*.

These findings have industrial implications, as they reveal the existence of three closely related but different lineages that have evolved traits beneficial for cheese making, with different phenotypic traits selected according to usage. This study should foster further research, as it would be interesting to assess other important traits, such as volatile compound production and the efficiencies of lipolysis and proteolysis. Our findings raise questions about the use of limited number of clonal strains for cheese making, which tends to lead to degeneration, limiting the possibilities for further improvement, which is currently a major concern in the agrofood sector [60], despite the great geneticist Nikolai Vavilov, known for having identified the centers of origin of cultivated plants long ago, highlighting the importance of genetic diversity in domesticated organisms for variety improvement and diversification [61].

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE

- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - DNA extraction
  - Genome sequencing, assembly, and mapping
  - Analysis of the mating-type genes
  - Identification by a single PCR reaction the *P. camemberti* variety
  - Statistics of population genetics and genomic scans
  - Genetic structure
  - Phylogenetic analyses
  - Laboratory experiments
- QUANTIFICATION AND STATISTICAL ANALYSIS

#### SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cub.2020.08.082>.

#### 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 T.G., J.R., E.C., and M.C. and a Fondation Louis D. grant from the Institut de France to T.G.

#### AUTHOR CONTRIBUTIONS

J.R. and T.G. designed and supervised the study and obtained funding. J.R., E.D., A.S., and S.L.P. generated the data. J.R., R.C.R.d.I.V., and B.B. analyzed the genomes. J.R., E.D., and A.S. performed the experiments. M.C., E.C., and E.P. performed the mycotoxin experiment and contributed to CPA biosynthetic gene cluster analyses. J.R. and E.D. analyzed the data from the laboratory experiments. J.R. and T.G. wrote the manuscript, with contributions from all the authors. J.R., T.G., and R.C.R.d.I.V. prepared the final version of the manuscript.

#### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 12, 2020

Revised: July 2, 2020

Accepted: August 24, 2020

Published: September 24, 2020

#### REFERENCES

1. Warmuth, V., Eriksson, A., Bower, M.A., Cañon, J., Cothran, G., Distl, O., Glowatzki-Mullis, M.-L., Hunt, H., Luis, 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, A., 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., Kostriksyna, 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. *Annu. Rev. Genet.* **38**, 37–59.
7. Wydooghe, E., Berghmans, E., Rijsselaere, T., and Van Soom, A. (2013). International breeder inquiry into the reproduction of the English bulldog. *Vlaams Diergeneesk. Tijdschr.* **82**, 38–43.
8. Perrier, X., De Langhe, E., 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. *Proc. Natl. Acad. Sci. USA* **108**, 11311–11318.
9. Gladieux, P., Ropars, J., Badouin, H., Branca, A., Aguilera, G., de Vienne, D.M., Rodriguez de la Vega, R.C., Branco, S., and Giraud, T. (2014). Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Mol. Ecol.* **23**, 753–773.
10. 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. *Microbiol. Spectrum* **5**, FUNK-0030-2016.
11. Almeida, P., Gonçalves, C., Teixeira, S., Libkind, D., Bontrager, M., Masneuf-Pomarède, I., Albertin, W., Durrrens, P., Sherman, D.J., Marullo, P., et al. (2014). A Gondwanan imprint on global diversity and domestication of wine and cider yeast *Saccharomyces uvarum*. *Nat. Commun.* **5**, 4044.
12. Fay, J.C., and Benavides, J.A. (2005). Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. *PLoS Genet.* **1**, 66–71.
13. Gallone, B., Steensels, J., Prah, 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.
14. Legras, J.-L., Merdinoglu, D., Cornuet, J.-M., and Karst, F. (2007). Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol. Ecol.* **16**, 2091–2102.
15. 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. *Proc. Natl. Acad. Sci. USA* **108**, 14539–14544.
16. 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. *Mol. Biol. Evol.* **32**, 1695–1707.
17. Novo, M., Bigey, F., Beyne, E., Galeote, V., Gavory, F., Mallet, S., Cambon, B., Legras, J.-L., Wincker, P., Casaregola, S., and Dequin, S. (2009). Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *Proc. Natl. Acad. Sci. USA* **106**, 16333–16338.
18. Peter, J., De Chiara, M., 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.
19. 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.
20. Spor, A., Kvittek, 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.
21. 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 Genet.* 9, e1003366.
22. Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.-J., Wortman, J.R., Batzoglu, 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.
  23. 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*. *Curr. Biol.* 22, 1403–1409.
  24. Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., Kusumoto, K., Arima, T., Akita, O., Kashiwagi, Y., et al. (2005). Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 438, 1157–1161.
  25. Dumas, E., Feurtey, A., Rodríguez de la Vega, R.C., Le Prieur, S., Snirc, A., Coton, M., Thierry, A., Coton, E., Le Piver, M., Roueyre, D., et al. (2020). Independent domestication events in the blue-cheese fungus *Penicillium roqueforti*. *Mol. Ecol.* 29, 2639–2660.
  26. 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.
  27. 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., and Giraud, T. (2015). Adaptive horizontal gene transfers between multiple cheese-associated Fungi. *Curr. Biol.* 25, 2562–2569.
  28. Gillot, G., Jany, J.-L., Coton, M., Le Floch, G., Debaets, S., Ropars, J., López-Villavicencio, M., Dupont, J., Branca, A., Giraud, T., and Coton, E. (2015). Insights into *Penicillium roqueforti* morphological and genetic diversity. *PLoS ONE* 10, e0129849.
  29. 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.
  30. Pitt, J.I. (1979). The Genus *Penicillium* and Its Teleomorphic States *Eupenicillium* and *Talaromyces* (Academic).
  31. Delfosse, C. (2008). Histoires de Bries (Illustria Librairie des Musées).
  32. Pitt, J.I., Cruickshank, R.H., and Leistner, L. (1986). *Penicillium commune*, *P. camemberti*, the origin of white cheese moulds, and the production of cyclopiazonic acid. *Food Microbiol.* 3, 363–371.
  33. Pouriau, A.-F. (1895). *La Laiterie, Art de Traiter le Lait, de Fabriquer le Beurre et les Principaux Fromages Français et Étrangers* (Paris: Librairie Audot. Lebroc et cie, successeurs 8, rue Garancière).
  34. Garnier, L., Valence, F., Pawtowski, A., Auhustsinava-Galerie, L., Frotté, N., Baroncelli, R., Deniel, F., Coton, E., and Mounier, J. (2017). Diversity of spoilage fungi associated with various French dairy products. *Int. J. Food Microbiol.* 241, 191–197.
  35. Giraud, F., Giraud, T., Aguilera, 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. *Int. J. Food Microbiol.* 137, 204–213.
  36. 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*. *Stud. Mycol.* 78, 343–371.
  37. 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. *Compr. Rev. Food Sci. Food Saf.* 13, 437–456.
  38. 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.
  39. 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.
  40. Gibbons, J.G., and Rokas, A. (2013). The function and evolution of the *Aspergillus* genome. *Trends Microbiol.* 21, 14–22.
  41. 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.
  42. 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 Genet. Biol.* 39, 151–167.
  43. Boisard, P. (2007). *Le Camembert, Mythe Français* (Odile Jacob).
  44. 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.
  45. 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.
  46. Kritikos, G., Banzhaf, M., Herrera-Dominguez, L., Koumoutsis, A., Wartel, M., Zietek, M., and Typas, A. (2017). A tool named Iris for versatile high-throughput phenotyping in microorganisms. *Nat. Microbiol.* 2, 17014.
  47. 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.
  48. Lund, F., Filtenborg, O., and Frisvad, J.C. (1998). *Penicillium caseifulvum*, a new species found on *Penicillium roqueforti* fermented cheeses. *J. Food Mycol.* 1, 95–101.
  49. Barbesgaard, P., Heldt-Hansen, H.P., and Diderichsen, B. (1992). On the safety of *Aspergillus oryzae*: a review. *Appl. Microbiol. Biotechnol.* 36, 569–572.
  50. 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.
  51. Giraud, T., Refrégier, G., Le Gac, M., de Vienne, D.M., and Hood, M.E. (2008). Speciation in fungi. *Fungal Genet. Biol.* 45, 791–802.
  52. Le Gac, M., and Giraud, T. (2008). Existence of a pattern of reproductive character displacement in Homobasidiomycota but not in Ascomycota. *J. Evol. Biol.* 21, 761–772.
  53. Mahler, D.L., Ingram, T., Revell, L.J., and Losos, J.B. (2013). Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* 341, 292–295.
  54. Thorpe, R.S., Barlow, A., Malhotra, A., and Surget-Groba, Y. (2015). Widespread parallel population adaptation to climate variation across a radiation: implications for adaptation to climate change. *Mol. Ecol.* 24, 1019–1030.
  55. Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A., and Cresko, W.A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6, e1000862.
  56. Duboué, E.R., Keene, A.C., and Borowsky, R.L. (2011). Evolutionary convergence on sleep loss in cavefish populations. *Curr. Biol.* 21, 671–676.
  57. Martínez-Ainsworth, N.E., and Tenaillon, M.I. (2016). Superheroes and masterminds of plant domestication. *C. R. Biol.* 339, 268–273.
  58. Gross, B.L., and Olsen, K.M. (2010). Genetic perspectives on crop domestication. *Trends Plant Sci.* 15, 529–537.
  59. Decker, M., and Nielsen, P.V. (2005). The inhibitory effect of *Penicillium camemberti* and *Geotrichum candidum* on the associated fungi of white mold cheese. *Int. J. Food Microbiol.* 104, 51–60.
  60. Gouyon, P.-H., Leriche, H., Civard, A., Reeves, H., and Hulot, N. (2010). *Aux Origines de l’Environnement* (Fayard).

61. Vavilov, N.I., Vavilov, M.I., Vavilov, N.I., and Dorofeev, V.F. (1992). *Origin and Geography of Cultivated Plants* (Cambridge University).
62. Glass, N.L., and Donaldson, G.C. (1995). Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. *Appl. Environ. Microbiol.* **61**, 1323–1330.
63. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.
64. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303.
65. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **27**, 722–736.
66. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963.
67. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359.
68. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356.
69. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534.
70. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589.
71. Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E., and Lercher, M.J. (2014). PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936.
72. Schliep, K.P. (2011). phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592–593.
73. Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C., and Gascuel, O. (2011). Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* **60**, 685–699.
74. Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188–1195.
75. Hill, T.W., and Kafer, E. (2001). Improved protocols for *Aspergillus* minimal medium: trace element and minimal medium salt stock solutions. *Fungal Genet. Rep.* **48**, 8.
76. Peterson, R.A., and Cavanaugh, J.E. (2019). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *J. Appl. Stat.* Published online June 15, 2019. <https://doi.org/10.1080/02664763.2019.1630372>.
77. Porcher, E., Giraud, T., Goldringer, I., and Lavigne, C. (2004). Experimental demonstration of a causal relationship between heterogeneity of selection and genetic differentiation in quantitative traits. *Evolution* **58**, 1434–1445.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Raw reads (Illumina sequencing)	This paper	GenBank: PRJNA655754
Experimental Models: Organisms/Strains		
Strains of <i>Penicillium camemberti</i> var. “ <i>camemberti</i> ,” var. “ <i>caseifulvum</i> ,” <i>P. biforme</i> , <i>P. fuscoglaucum</i> , <i>P. palitans</i> and <i>P. cavernicola</i>	Table S1	N/A
Oligonucleotides		
Bt2a: GGTAACCAAATCGGTGCTGCTTTC	[62]	N/A
Bt2b: AACCTCAGTGTAGTGACCCTTGCC	[62]	N/A
PC4F: CAAGCTGGCCGATAACCTG	[35]	N/A
PC4R: CCATCCGCTTGATTCTCCT	[35]	N/A
Pcamcasei_F: CAAATCGAGGAATCACAAGTCT	This paper	N/A
Pcam_R: GAAATTCGATCTTGGGCAAA	This paper	N/A
Pcasei_R: GAGTAAGCATTGGCCTCGAC	This paper	N/A
Software and Algorithms		
Trimmomatic	[63]	<a href="http://www.usadellab.org/cms/?page=trimmomatic">http://www.usadellab.org/cms/?page=trimmomatic</a>
GATK	[64]	<a href="https://gatk.broadinstitute.org/hc/en-us">https://gatk.broadinstitute.org/hc/en-us</a>
Canu	[65]	<a href="https://github.com/marbl/canu">https://github.com/marbl/canu</a>
pilon	[66]	<a href="https://github.com/broadinstitute/pilon">https://github.com/broadinstitute/pilon</a>
Bowtie2	[67]	<a href="http://bowtie-bio.sourceforge.net/bowtie2/index.shtml">http://bowtie-bio.sourceforge.net/bowtie2/index.shtml</a>
ANGSD	[68]	<a href="http://www.popgen.dk/angsd/index.php/ANGSD#Overview">http://www.popgen.dk/angsd/index.php/ANGSD#Overview</a>
IQ-Tree2	[69]	<a href="http://www.iqtree.org/">http://www.iqtree.org/</a>
ModelFinder	[70]	<a href="http://www.iqtree.org/ModelFinder/">http://www.iqtree.org/ModelFinder/</a>
Iris	[46]	<a href="http://critichu.github.io/Iris">http://critichu.github.io/Iris</a>
R Project for statistical computing	R core team 2017	<a href="https://www.r-project.org/">https://www.r-project.org/</a>

### RESOURCE AVAILABILITY

#### Lead Contact

Further information and requests for resources and strains with ESE numbers should be directed to and will be fulfilled by the Lead Contact, Jeanne Ropars ([jeanne.ropars@universite-paris-saclay.fr](mailto:jeanne.ropars@universite-paris-saclay.fr)).

#### 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 generated in this study have been deposited to GenBank: PRJNA655754.

The dataset corresponding to phenotype measurements is available in [Data S1](#).

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

We analyzed strains isolated from 37 cheeses from five countries around the world (e.g., Camembert, Saint Marcellin, tomme, Sein de Nounou). We also collected strains from dried sausages (n = 8) and moldy food (e.g., bread, n = 4). Twelve were isolated from other

substrates than food. Spores were sampled from the food sources and spread on Petri dishes containing malt-agar medium, which were then incubated for five days at 25°C. A total of 41 strains were obtained from public strain collections (the Laboratoire de Cryptogamie (LCP strains) at the National Museum of Natural History in Paris (France), the Laboratoire Universitaire de Biodiversité et Ecologie Microbienne (LUBEM, UBOCC strains) in Plouzané (France) and the Westerdijk Fungal Biodiversity Institute (CBS strains) in Utrecht (the Netherlands)). The LMA421 strain was kindly given by Steve Labrie (University of Laval, Canada). We obtained strains from natural environments (e.g., wood or natural cave walls) from eight different countries. Detailed information about the strains used in this study can be found in the [Table S1](#). For each strain, single-spore cultures were generated by a dilution method, to ensure that only a single haploid genotype was cultured for each strain. We checked the species identification of all strains by Sanger sequencing of the  $\beta$ -tubulin gene (primers Bt2a - GGTAACCAAATCGGTGCTGCTTTC and bt2b - AACCTCAGTGAGT GACCCTTGGC [62]) and the PC4 microsatellite flanking regions [35] (primers PC4F - CAAGCTGGCCGATAACCTG and PC4R - CCATCCGCTTGATTCTCTCT), to distinguish between *P. bifforme* and *P. camemberti*. All the strains were added to the ESE (Ecology Systematics Evolution Laboratory) collection, under an ESE accession number, and are available from the corresponding collections (ESE, LUBEM, MNHN).

## METHOD DETAILS

### DNA extraction

For each strain, we used the Nucleospin Soil Kit (Macherey-Nagel, Düren, Germany) to extract DNA from fresh haploid mycelium grown for five days on malt agar.

### Genome sequencing, assembly, and mapping

Sequencing was performed with Illumina HiSeq 2500 paired-end technology (Illumina Inc.), with a mean insert size of 400 bp, at the INRA GenoToul platform, to obtain 10–50X coverage ([Table S1](#)). In addition, the genome of the *P. camemberti* LCP06093 (also known as FM013) reference strain, initially sequenced with 454 technology [26], was resequenced at 34x using the PacBio's long-read technology, assembled with Canu 1.8 [65] and polished with pilon 1.23 [66]. The PacBio assembly and reads for all samples have been deposited in GenBank: PRJNA655754.

Reads were trimmed with Trimmomatic [63] (cut adaptor, options PE, LEADING:3 TRAILING:3 SLIDINGWINDOW:4:25 MINLEN:36) and mapped onto the high-quality PacBio reference genome of LCP06093 with Bowtie2 [67] (options very-sensitive-local, -phred33 PL:ILLUMINA -X 1000). Mapped reads were filtered for PCR duplicates with picard tools (<http://broadinstitute.github.io/picard>) and realigned on the PacBio reference genome with GATK [64]. Single-nucleotide polymorphisms (SNPs) were called with GATK HaplotypeCaller, which provides one gVCF per strain (option -ERC GVCF). GVCFs were combined with GATK CombineGVCFs, genotyped with GATK genotype and hard filtered in accordance with GATK best practice workflow recommendations (<https://software.broadinstitute.org/gatk/best-practices/>).

### Analysis of the mating-type genes

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

### Identification by a single PCR reaction the *P. camemberti* variety

We designed two pairs of primers to identify by a single PCR reaction the *P. camemberti* variety, each yielding a specific amplicon size. The forward primer is common to both reactions (Pcamcasei\_F - CAAATCGAGGAATCACAAAGTCT); Reverse primers sequences are Pcam\_R - GAAATTCGATCTTGGGCAAA and Pcasei\_R - GAGTAAGCATTGGCCCTCGAC. The expected product size for *P. camemberti* var. “*camemberti*” isolates is of 776bp, whereas the expected product size for *P. camemberti* var. “*caseifulvum*” isolates is of 233bp.

PCR was performed in 30  $\mu$ L reactions, using 15  $\mu$ L of template DNA (5ng/ $\mu$ L), 0.15  $\mu$ L of Taq DNA polymerase (5 U/ $\mu$ L, mpbio medical), 3  $\mu$ L of buffer with MgSO<sub>4</sub> (25mM), 1.2  $\mu$ L of dNTPs (10mM), 0.6  $\mu$ L of each reverse primer, 1.2  $\mu$ L of the forward primer and 15  $\mu$ L of water.

Amplifications were performed on a Thermal cycler using a touchdown PCR, with 8 cycles of 30 s at 94°C, an initial annealing temperature of 60°C for 30 s gradually reduced by 2°C every cycle and 1min30 at 72°C. The program is followed by 24 cycles of 30 s at 94°C, 30 s at 53°C and 1min30 at 72°C. The program was ended by a 7 min extension step at 72°C.

PCR products were purified and sequenced by GENEWIZ, in one direction (forward).

Expected DNA sequence for *P. camemberti* var. “*camemberti*” is CCCCGGGAGGTTCCCTAGGCTGGCCGGGATGCTAGCTA TATTAAGAGGCCCGCGGATGCTAAAGCCAATTATCTTTAGTATATCGCATCGCTACCTTAGCATTAGTTACTAATATTTGTTAT ATAAACGGAATTTAACTCCTCGATTCTTTCTACCAAAATGATTGGCTAGTTAAGCTTAACGACAAGGTCGCCGACTTTAGCGCTAAGC AAGTCATCACCTTCGATATTGCTCTCACTTATTACAACCGCCCTAGACTAGCCAATTACACACTAGGCTAAAAAGACCACTATAT CTAAGTATTACGCTCAGATAGTCTAGGCTCCGTTATATATAGTGCCTGTTATAACATTACTACCATCATATTGACAATATCATTATC CTCGCTCCCCGATCGCCCTCGCTATCCGAGTCCATAAACCCAAAGAGAGCAAAGTATCAACACTATCAAGATCCACTACAATTC CACCATGATAGCGAAAAACCAACCATTAGCAATAAAAGGTCTATAGATCAGGGAGCTTCTCCAAGGTATCGCAAGGTAAGTACTGGATAT ACAACATAGACGACAAAATCACTAAGTCTTAACCACCCCTACGCTCCGACCCTAAATATATAGCGACTTCACCATTATGTGTAAG

GAGAAGAATATACTGTATATAGGTATATAGTTTGCCCAA and for *P. camemberti* var. “*caseifulvum*” CTGTTCTCCCCATATATGAAGAGAACAGTTTTGTCTCCACTCCCACTATTTTTCTGCGTCTCTATTTTATAATTTGTGATTACCTGATCCCATCATGGCA GAAAAGATCTCCGGCAATATAGGTGTGAGGCA.

### Statistics of population genetics and genomic scans

Nucleotide diversity indices ( $\pi$  and Watterson’s  $\theta$ ) and genetic differentiation indices ( $F_{ST}$  and  $d_{xy}$ ) were calculated using the R package PopGenome [71] for the four genetic clusters *P. biforme*, *P. fuscoglaucum*, *P. camemberti* var. “*camemberti*” and *P. camemberti* var. “*caseifulvum*.”  $\pi$  and Watterson’s  $\theta$  were calculated per site. We also performed genomic scans along the 20 largest scaffolds of the *P. camemberti* PacBio reference using sliding windows of 75 kb and step sizes of 7.5 kb. Genome scans for  $\pi$  and Watterson’s  $\theta$  were performed within each of the four genetic clusters (*P. biforme*, *P. fuscoglaucum*, *P. camemberti* var. “*camemberti*” and *P. camemberti* var. “*caseifulvum*”). Genome scans based on  $F_{ST}$  and  $d_{xy}$  indices were performed between all possible pairs of the four clusters (*P. biforme*, *P. fuscoglaucum*, *P. camemberti* var. “*camemberti*” and *P. camemberti* var. “*caseifulvum*”). We tested whether the genes belonging to windows with the 1% highest  $d_{xy}$  values were enriched in particular predicted functions (interpro). We performed gene enrichment analyses using Fisher exact tests (‘fisher.test’ function in the R package stats) corrected for multiple testing (‘p.adjust’ function in the R package stats with the false discovery rate, ‘fdr’, method). We also tested whether the number of genes with predicted functions within windows with the 1% highest  $d_{xy}$  values deviated from expectations based on the proportion in the full genome of the PacBio reference using a  $\chi^2$  test (Data S3 and S4).

We calculated the average pairwise identity between species pairs, as well as the number of fixed, private and shared SNPs. When comparing two populations, private SNPs are sites that are polymorphic in only one population (e.g., A/C for individuals in population 1 and T for individuals in population 2), shared SNPs are polymorphic positions within the two populations (e.g., A/C in population 1 and A/T in population 2) and fixed SNPs are positions that are fixed in each population but different between populations (e.g., A for all individuals of population 1 and G for all individuals of population 2).

### Genetic structure

We used the dataset of 392,072 SNPs to infer population structure. We inferred individual ancestry from genotype likelihoods based on realigned reads, by assuming a known number of admixing populations, ranging from  $K = 2$  to  $K = 6$ , using NGSadmix from the ANGSD package [68]. A hundred independent runs were carried out for each number of clusters ( $K$ ). We identified clustering solutions in replicated runs for each  $K$  value and displayed population structure graphically using the R function ‘barplot’. We used the R package phangorn [72] for neighbor-net analyses, and the ‘prcomp’ function of R for principal component analysis (PCA).

### Phylogenetic analyses

We inferred phylogenetic relationships among the 64 isolates using the dataset of 1,124,983 SNPs in a maximum likelihood framework using IQ-TREE2 [69]. We also used *P. cavernicola* (LCP05631 strain) and *P. palitans* (ESE00153 and LCP04862 strains). The tree has been midpoint rooted. The best-fit model chosen using ModelFinder [70] according to BIC was TVM+F+ASC+R3. Branch supports are SH-aLRT [73] and ultrafast bootstrap support [74].

### Laboratory experiments

#### Sampling and calibration

All experiments were performed on the whole collection of *P. camemberti* ( $n = 19$ ), *P. biforme* ( $n = 28$ ) and *P. fuscoglaucum* ( $n = 14$ ) strains, including 35 strains isolated from cheeses, 10 from other food environments (e.g., bread, sausage) and 14 strains isolated from environments other than food. Experiments were initiated with spore suspensions calibrated to  $10^7$  spores/mL with a hemocytometer, under the constraint of the low rate of sporulation in *P. camemberti*.

#### Growth in different conditions and on different media

We investigated whether strains isolated from cheese displayed faster or slower radial growth than strains isolated from other environments when cultured on cheese or other substrates, and under conditions similar to those in maturation caves or other conditions. We prepared four different culture media: a cheese medium without salt, a salted cheese medium (17 g/L, corresponding to a typical cheese), a malt medium and a minimal medium. The cheese media were produced from an unsalted drained cow’s milk cheese from Coubertin Farm in Saint Rémy-les-Chevreuse (France), as previously described [59]; we added five drops of blue food coloring to these media, to make it easier to distinguish white fungal colonies from the medium. The cheese media were rich in lipids and proteins, whereas the malt medium (20 g/L) was rich in carbohydrates. The minimal medium [75] contained only the trace elements necessary for fungal survival (i.e., iron sulfate, zinc sulfate, boric acid, magnesium chloride, copper sulfate, ammonium heptamolybdate, cobalt chloride, EDTA, sodium nitrate, potassium chloride, potassium phosphate, and magnesium sulfate). All media were sterilized in an autoclave (121°C for 20 min for the malt and minimal media, 110°C for 15 min for the cheese media). Each 90 mm-diameter Petri dish was filled with 20 mL of the appropriate medium.

We allowed the fungi to grow in the dark, under three different conditions for each strain and each medium: 10°C with 98% humidity (Camembert cave conditions), 8°C with 85% humidity (cave conditions for other cheeses) and 25°C with ambient relative humidity (ambient conditions). We had 244 Petri dishes in total for each set of conditions, and we left the colonies to grow for 10 days. Images of the Petri dishes were obtained with a Scan 1200 from Interscience and analyzed with IRIS [46] for growth and color-opacity measurements. The colony size was estimated as the colony area in pixels. The opacity score was defined as the sum of the brightness

values for all the pixels within the colony bounds; this score thus captures both the color and the three-dimensional colony growth (as opacity increases with fluffiness) and can be used as a proxy to assess the fluffiness of the fungal colony (Data S1 and S2).

### Mycotoxin production

We selected at least six strains from each genetic cluster for this experiment (Data S1). For measurements of mycotoxin production, we used 1  $\mu$ L of calibrated spore suspension ( $10^7$  spores/mL) for each of the 61 strains to inoculate YES (yeast extract sucrose) agar medium buffered at pH 4.5 with phosphate-citrate buffer and characterized by a high C/N ratio to favor mycotoxin production as already described in [50]. Each culture was performed in triplicate for mycotoxin analyses as well as for fungal dry weight measurements. The plates were incubated at 25°C in the dark for 10 days and were then stored at –20°C until mycotoxin analysis.

For mycotoxin extractions [50], we homogenized the thawed samples with an Ultraturrax T25 (IKA, Heidelberg, Germany) before recuperating 4g aliquots. Then, 25 mL of acetonitrile (ACN) supplemented with 0.1% formic acid (v/v) was added, samples were vortexed for 30 s followed by 15 min sonication. Extracts were then centrifuged for 10 min at 5000 g and the recuperated supernatants were directly filtered through 0.2  $\mu$ m PTFE membrane filters (GE Healthcare Life Sciences, UK) into amber vials. All samples were stored at –20°C until analyses.

Mycotoxin detection and quantification were performed using an Agilent 6530 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) LC/MS system equipped with a Binary pump 1260 and degasser, well plate autosampler set to 10°C and a thermostated column compartment. Filtered samples (2 $\mu$ L) were injected into a ZORBAX Extend C-18 column (2.1x50mm and 1.8  $\mu$ m, 600 bar) maintained at 35°C with a flow rate of 0.3 ml/min using mobile phase A (milli-Q water + 0.1% formic acid (v/v) and 0.1% ammonium formate (v/v) and mobile phase B (ACN + 0.1% formic acid). Mobile phase B was maintained at 10% for 4 min followed by a gradient from 10 to 100% for 16 min. Then, mobile phase B was maintained at 100% for 2 min before a 5 min post-time. Cyclopiazonic acid (CPA) was ionized in electrospray ionization mode ESI+ in the mass spectrometer with the following parameters: capillary voltage 4 kV, source temperature 325°C, nebulizer pressure 50 psig, drying gas 12 l/min, ion range 100-1000 m/z. CPA had a retention time of 15.3 min and was quantified using the [M+H]<sup>+</sup> 337.1545 m/z ion and [M+Na]<sup>+</sup> 359.1366 qualifier ion (ESI+). Other extrolites included in our internal database that were targeted (in ESI+ and ESI- modes) were andrastin A, citreoviridin, citrinin, ermefortins A & B, (iso)-fumigaclavin A, griseofulvin, meleagrins, mycophenolic acid, ochratoxin A, patulin, penitrem A, PR toxin, roquefortin C, sterigmatocystin.

We used a matrix matched calibration curve ( $R^2 > 0.99$ ) for reliable mycotoxin quantification with final concentrations ranging from 10 to 10000 ng/ml. Method performance and mycotoxin determination was carried out as previously described in [50]. Specific mycotoxin production was expressed as ng per g fungal dry weight (Data S1 and S2).

### Competition

We inoculated salted cheese with 150  $\mu$ L of a calibrated spore solution ( $10^7$  spores/mL), which we allowed to grow as a lawn. We used two *P. camemberti* var. “*caseifulvum*” strains (LCP05630 and ESE00019), three *P. camemberti* var. “*camemberti*” strains (FM013, LCP04810 and LCP05527) and 10 *P. bifforme* strains (ESE00018, ESE00086, ESE00089, ESE00126, ESE00139, LCP05531, LCP06701, ESE00154, LCP06620 and LCP05496). After 24 h of growth, we deposited a single 20  $\mu$ L drop of calibrated spore solution from a challenger species ( $10^7$  spores/mL) in the middle of the Petri dish. The challenger species used were *G. candidum* (n = 4, ESE00163 and ESE00165, CBS11628 and VTTC4559), *P. bifforme* (n = 6, ESE00126, ESE00089, ESE00018, ESE00154, LCP005531, LCP06701), *P. fuscoglaucum* (n = 5, ESE00090, LCP06621, LCP000797, LCP04799, LCP03239), and *P. roqueforti* (n = 12, three from each of the four genetic clusters described in [25]: LCP06036, LCP06037, LCP06128, LCP06043, LCP06098, LCP06064, FM164, LCP06137, LCP06131, LCP06173). The colony size of the challenger was calculated as the colony area in pixels (Data S1 and S2).

## QUANTIFICATION AND STATISTICAL ANALYSIS

We performed analyses of variance (ANOVAs) followed by post-ANOVA Tukey’s HSD (honest significant difference) tests, with R. The data normality was improved with the bestNormalize R package [76]. For improving distribution normality, we used an ordered quantile (ORQ) transformation (a one-to-one transformation for vectors with an arbitrary distribution that generates normally distributed vectors) for diameter measurements and competition data and a Yeo-Johnson transformation for CPA (mycotoxin) data. Some residuals were still not normally distributed, but ANOVA is known to be robust to deviations from normality [77]. When residues deviated from normality (diameter measurements for the growth and competition), we 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.

### III. Article : La domestication des champignons *Penicillium* du fromage



INSTITUT DE FRANCE  
Académie des sciences

# *Comptes Rendus*

---

## *Biologies*


Jeanne Ropars, Thibault Caron, Ying-Chu Lo, Bastien Bennetot  
et Tatiana Giraud

**La domestication des champignons *Penicillium* du fromage**

Volume 343, issue 2 (2020), p. 155-176.

<https://doi.org/10.5802/crbiol.15>

© Académie des sciences, Paris and the authors, 2020.  
*Some rights reserved.*

 This article is licensed under the  
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.  
<http://creativecommons.org/licenses/by/4.0/>



*Les Comptes Rendus. Biologies* sont membres du  
Centre Mersenne pour l'édition scientifique ouverte  
[www.centre-mersenne.org](http://www.centre-mersenne.org)





Vie de l'Académie / *Life of the Academy*

# La domestication des champignons *Penicillium* du fromage

## *The domestication of Penicillium cheese fungi*

Jeanne Ropars<sup>a</sup>, Thibault Caron<sup>a</sup>, Ying-Chu Lo<sup>a</sup>, Bastien Bennetot<sup>a</sup> et Tatiana Giraud<sup>\*, a</sup>

<sup>a</sup> Ecologie Systématique et Evolution, CNRS, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France

*Courriels* : [jeanne.ropars@universite-paris-saclay.fr](mailto:jeanne.ropars@universite-paris-saclay.fr) (J. Ropars),  
[thibaultcaron864@gmail.com](mailto:thibaultcaron864@gmail.com) (T. Caron), [yingchu.lo@gmail.com](mailto:yingchu.lo@gmail.com) (Y.-C. Lo),  
[bastien.bennetot@universite-paris-saclay.fr](mailto:bastien.bennetot@universite-paris-saclay.fr) (B. Bennetot),  
[tatiana.giraud@universite-paris-saclay.fr](mailto:tatiana.giraud@universite-paris-saclay.fr) (T. Giraud)

**Résumé.** La domestication est le processus d'évolution d'un organisme en cours de sélection par l'homme et, à ce titre, elle a servi de modèle pour l'étude de l'adaptation depuis Charles Darwin. Nous passons ici en revue les études récentes sur la génomique de l'adaptation et le syndrome de domestication dans deux lignées de champignons fromagers, *Penicillium roqueforti* utilisé pour l'affinage des fromages bleus, et le complexe d'espèces *Penicillium camemberti* utilisé pour la fabrication de fromages à pâte molle comme le camembert et le brie. La génomique comparative a révélé des transferts de gènes horizontaux impliqués dans l'adaptation convergente du fromage. La génomique des populations a identifié des populations différenciées aux caractères contrastés, plusieurs populations ayant été indépendamment domestiquées pour la fabrication de fromage à la fois chez *P. roqueforti* et le complexe d'espèces *Penicillium camemberti*, et ayant subi de forts goulots d'étranglement. Les différentes populations fromagères ont acquis des caractères avantageux pour la fabrication du fromage par rapport aux populations non fromagères, en ce qui concerne la couleur, la production de spores, les taux de croissance sur le fromage, la tolérance au sel, la lipolyse, la protéolyse, la production de composés volatils ou de toxines et/ou la capacité d'exclusion compétitive. Les populations fromagères présentent également une dégénérescence de certaines fonctions non utilisées, telles que la capacité réduite de reproduction sexuée ou de croissance dans des conditions difficiles. Ces récentes découvertes ont une importance fondamentale pour notre compréhension de l'adaptation et ont un intérêt appliqué pour l'amélioration des souches.

**Abstract.** Domestication is the process of organism evolution under selection by humans, and as such has been a model for studying adaptation since Charles Darwin. Here we review recent studies on the genomics of adaptation and domestication syndrome in two cheese-making fungal lineages, *Penicillium roqueforti* used for maturing blue cheeses, and the *Penicillium camemberti* species complex used for making soft cheeses such as Camembert and Brie. Comparative genomics have revealed horizontal gene transfers involved in convergent adaptation to cheese. Population genomics have identified differentiated populations with contrasted traits, several populations having independently been domesticated for cheese making in both *P. roqueforti* and the *Penicillium camemberti* species complex, and

\* Auteur correspondant.

having undergone bottlenecks. The different cheese populations have acquired traits beneficial for cheese making in comparison to non-cheese populations, regarding color, spore production, growth rates on cheese, salt tolerance, lipolysis, proteolysis, volatile compound or toxin production and/or competitive ability. The cheese populations also show degeneration for some unused functions such as decreased ability of sexual reproduction or of growth under harsh conditions. These recent findings have fundamental importance for our understanding of adaptation and have applied interest for strain improvement.

**Mots-clés.** Domestication, Fromage, Champignon, *Penicillium*, Adaptation, Génomique.

**Keywords.** Domestication, Cheese, Fungi, *Penicillium*, Adaptation, Genomics.

*Manuscrit reçu et accepté le 29 juillet 2020.*

## 1. Introduction

La divergence adaptative (différenciation des populations associée à l'adaptation à une nouvelle niche écologique) est un processus clé de la biologie évolutive. L'étude de ce processus est essentielle pour comprendre l'adaptation des organismes à leur environnement et la façon dont la biodiversité se développe. La domestication est un cas de divergence adaptative sous une forte pression humaine et à l'échelle temporelle humaine. Charles Darwin a utilisé des pigeons pour illustrer le processus et le pouvoir de sélection dans des formes de rendement radicalement différentes de la forme sauvage ancestrale en comparant les pigeons sauvages et domestiqués [1]. Depuis lors, la domestication est utilisée comme un bon modèle pour comprendre l'adaptation, car elle implique une sélection forte et récente pour des caractères qui peuvent être facilement identifiés. Cette sélection forte et récente par l'être humain laisse souvent des traces d'adaptation dans les génomes qui sont plus faciles à détecter que celles laissées par la sélection naturelle. La comparaison des populations domestiquées et non domestiquées aux niveaux phénotypique et génomique peut permettre d'identifier l'architecture génétique des caractères adaptatifs et les mécanismes génomiques permettant une adaptation et une différenciation rapides, comme cela a été démontré pour le maïs et les chiens par exemple [2, 3].

Dans les cultures domestiquées, de multiples événements convergents d'évolution vers des caractères phénotypiques similaires ont été documentés, tels que la perte de la dispersion des graines, la minimisation de la dormance des graines et l'augmentation de la taille et du nombre des graines, principalement due à des changements génomiques différents chez différentes espèces [4, 5]; la perte de la dispersion des graines est une exception, avec la même base

génétique dans le sorgho, le riz, le maïs et le millet sétaire [6, 7]. Les organismes domestiqués sont donc des modèles utiles pour les études de convergence au niveau phénotypique et génomique.

Les espèces domestiquées présentent également souvent une dégénérescence pour des caractères qui sont utiles dans l'environnement sauvage mais qui ne présentent pas d'avantages dans les environnements humains, ce qui conduit à une sélection relâchée. Les espèces domestiquées peuvent également dégénérer en raison d'une accumulation délétère d'allèles due à de forts goulets d'étranglement et/ou à un auto-stop avec des allèles sélectionnés positivement. L'accumulation d'allèles délétères a été démontrée par exemple dans le riz, les porcs et les chevaux [8–11]. Les cas de dégénérescence de la fertilité comprennent une baisse de la fertilité des épillets latéraux chez l'orge domestiquée [12] et des très petites graines non fonctionnelles dans les cultivars de bananiers triploïdes [13, 14].

La domestication microbienne a été beaucoup moins étudiée que la domestication végétale ou animale, bien que les micro-organismes aient été utilisés pour la fabrication d'aliments et aient développé de nouveaux caractères dans le cadre de la sélection dirigée par l'homme. Les études sur les champignons domestiqués sont rares au-delà de l'espèce modèle de levure *Saccharomyces cerevisiae* utilisée pour la production de vin et de bière. Pourtant, les champignons domestiqués constituent des modèles eucaryotes très précieux pour l'étude de la divergence adaptative, avec leurs petits génomes, leur accès facile à la phase haploïde et leur traçabilité expérimentale pour les expériences *in vitro* [15–19]. De nombreux champignons ont été domestiqués, pour la fermentation et la maturation des aliments (par exemple pour la bière, le pain, le vin, la saucisse sèche ou le fromage), pour la production de métabolites

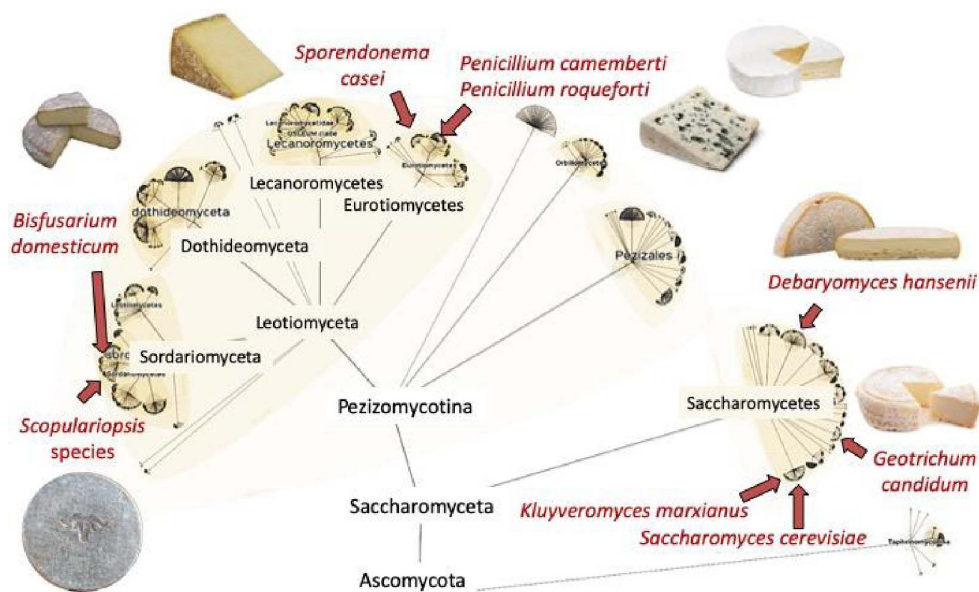
secondaires utilisés en pharmacie (par exemple la pénicilline), ou pour leurs valeurs nutritionnelles et gustatives (par exemple les champignons de Paris et shiitake) [20]. La fermentation est utilisée par les populations humaines depuis des millénaires [21, 22] comme moyen de conserver les aliments périssables, et de conférer au produit de nouvelles caractéristiques organoleptiques et nutritionnelles. Les produits fermentés qui en résultent ont souvent une forte valeur économique et culturelle. La fermentation repose sur les capacités métaboliques des microorganismes à transformer un substrat en un produit présentant un mélange de métabolites responsables d'aspects, de textures, de saveurs et de caractéristiques nutritionnelles particuliers.

Les microorganismes de fermentation, qui résultent à l'origine d'une contamination environnementale, ont été améliorés par l'homme, grâce à un processus de sélection d'abord involontaire, puis plus conscient, sur de nombreuses générations microbiennes, pour les souches donnant les meilleurs produits en termes de caractéristiques sensorielles, de durée de conservation et de sécurité. Cette sélection par l'homme a eu lieu au cours des siècles et a conduit à la domestication microbienne, c'est-à-dire à la différenciation génétique des populations sauvages et alimentaires [16, 17, 23] et à la sélection de phénotypes spécifiques présentant un intérêt pour l'être humain [24]. L'étude des champignons utilisés et sélectionnés au cours des siècles pour la production alimentaire peut ainsi fournir des indications sur les processus d'adaptation. Pourtant, malgré leur importance économique et industrielle, les champignons utilisés par l'être humain ont été peu étudiés, à l'exception de la levure de boulanger *Saccharomyces cerevisiae* utilisée dans la production de bière, de vin et de pain [23–33], et le champignon filamenteux *Aspergillus oryzae*, qui est utilisé pour la fermentation des produits à base de soja et de riz en Asie [34, 35].

Les champignons sont encore d'excellents modèles pour étudier l'évolution et l'adaptation des eucaryotes, compte tenu de leurs nombreux atouts expérimentaux [15]. De nombreux champignons ont de petits génomes, ils sont faciles à cultiver dans des conditions de laboratoire, ont des temps de génération courts, peuvent survivre de longues périodes au congélateur, sont faciles à transformer génétiquement, et les champignons présentent une riche diversité de lignées et de niches écologiques. Les cham-

pignons peuvent être utilisés pour répondre à des questions de biologie de l'évolution en combinant des approches complémentaires, telles que des méthodes génomiques et expérimentales. Les études sur les champignons domestiqués ont en fait déjà permis de faire la lumière sur les processus génomiques impliqués dans l'adaptation aux environnements anthropiques, notamment l'expansion des familles de gènes pour des voies métaboliques spécifiques, la perte de gènes qui ne sont plus nécessaires dans le nouvel environnement, l'acquisition de gènes par transfert vertical, l'hybridation et l'introgession [19, 23, 24, 28, 29, 32, 36]. L'existence de lignées multiples et distantes génétiquement de champignons utilisés pour la production de produits alimentaires similaires crée une opportunité idéale pour étudier la répétabilité de l'évolution, en fournissant plusieurs répliques indépendantes récentes d'adaptation à la même niche écologique sous sélection forte.

La domestication des champignons pour la fabrication du fromage constitue un système modèle particulièrement attrayant. En effet, le fromage est l'un des aliments fermentés les plus populaires dans le monde entier. On le trouve en Europe, en Amérique, en Asie et au Moyen-Orient avec différentes variétés de fromage, comme l'Anthotyros et la feta en Grèce, le Comté, le Reblochon et le Roquefort en France et le Chhurpi au Népal. Les champignons sont utilisés pour la maturation du fromage et jouent un rôle essentiel dans la fermentation, la texture, la couleur et la saveur du fromage. En outre, les multiples espèces fongiques éloignées sur le plan phylogénétique qui prospèrent dans la niche écologique du fromage offrent une "expérience" fascinante de sélection similaire récente et des possibilités uniques d'évaluer la répétabilité de l'évolution, d'autant plus que plusieurs lignées éloignées ont été utilisées pour des produits similaires (Figure 1). Nous passons ici en revue les études récentes sur l'histoire et le processus de domestication de deux espèces fromagères emblématiques, *Penicillium roqueforti* et *P. camemberti*, utilisées respectivement pour la fabrication de fromages bleus et de fromages à pâte molle (Figure 1). Après un bref aperçu des champignons *Penicillium*, nous résumons les connaissances récentes acquises sur leur histoire démographique, leur diversité génétique et leur structure de population, ainsi que sur leurs changements phénotypiques et génétiques dans le cadre de la domestication.



**FIGURE 1.** Les champignons du fromage ne sont pas tous étroitement liés, mais plutôt dispersés dans l'arbre généalogique des champignons. On trouve les champignons *Scopulariopsis* sur Ossau-Iraty et Tommes, *Bisfusarium domesticum* sur Saint-Nectaire et Reblochon, *Sporendonema casei* sur Cantal et Salers, *Penicillium camemberti* sur Camembert et Brie, *P. roqueforti* dans tous les fromages bleus du monde, *Debaromyces hansenii*, *Geotrichum candidum*, *Kluyveromyces marxianus* et *Saccharomyces cerevisiae* dans de nombreuses sortes de fromages. Phylogénie dessinée sur la base de LifeMap [37]. Images de fromages : *Scopulariopsis* : Ossau Iraty, *B. domesticum* : Saint-Nectaire, *S. casei* : Cantal, *P. camemberti* : Camembert, *P. roqueforti* : Roquefort, *D. hansenii* : Reblochon, *G. candidum* : fromage de chèvre. Les photos de Saint-Nectaire, Reblochon et Cantal sont tirées de Wikipédia.

## 2. Les champignons *Penicillium*

*Penicillium* est un genre de champignons des Ascomycota, décrit pour la première fois en 1809 par Link. Le nom *Penicillium* vient du mot latin qui signifie pinceau, reflétant la forme conidiophore des champignons *Penicillium* [38]. Le genre *Penicillium* contient 354 espèces acceptées taxonomiquement, isolées dans le monde entier dans divers environnements, par exemple le sol, l'air et la nourriture [38–40]. La plupart des espèces de *Penicillium* peuvent se reproduire par voie sexuée et asexuée. Pour la reproduction asexuée, les champignons *Penicillium* génèrent des conidies en chaîne (spores asexuées) via des structures spécifiques appelées conidiophores. Les conidiophores de *Penicillium* sont plus ou moins ramifiés, certaines espèces de *Penicillium* ayant des conidiophores uniquement composés de phialides d'où sortent les conidies (*Penicillia monoverticillate*), tandis que d'autres espèces ont des conidiophores

avec une (biverticillate) ou plusieurs (*Penicillia verticillate*) branches (appelées metula); ces schémas de ramification des conidiophores ont été utilisés par Pitt pour classer les espèces de *Penicillium* avant l'avènement de l'identification moléculaire. Les conidies de *Penicillium* sont généralement vertes ou bleues/vertes, la couleur changeant toutefois en fonction des conditions de croissance, telles que le milieu de culture, la lumière, l'humidité et la température [41]. La dispersion asexuée des spores peut se produire par dispersion aérienne, aquatique ou animale lorsque les spores germent en hyphes qui vont croître et se développer en un nouveau mycélium. Les champignons *Penicillium* peuvent également se reproduire par voie sexuée, lorsque les hyphes de deux individus haploïdes portant des types d'accouplement opposés, appelés MAT1-1 et MAT1-2, se rencontrent dans des conditions adéquates. Ensuite, les noyaux fusionnent (caryogamie) et une méiose se produit, donnant naissance à huit ascospores. Ces

ascospores sont enveloppées dans des asques, eux-mêmes étant à l'intérieur du cléistothecium, le corps fructifère des Ascomycètes.

Les champignons *Penicillium* sont bien connus en tant que champignons saprophytes décomposant les fruits et les cultures, en tant que contaminants alimentaires, et certaines espèces sont des agents pathogènes des cultures humaines ou alimentaires. Par exemple, *P. expansum* et *P. digitatum* sont des agents de dégradation post-récolte des agrumes et des poires [42, 43], *P. marneffeii* est un agent pathogène pour l'homme [44]; *P. oxalicum*, *P. viridicatum* et *P. cyclopium* sont couramment utilisés pour endommager les cultures [45]; *P. glabrum*, *P. allii*, *P. sclerotigenum*, *P. roqueforti* et *P. paneum* sont des agents d'altération des aliments [39, 40]. Les champignons *Penicillium* ont également d'énormes impacts positifs sur la société humaine : Alexander Fleming a accidentellement découvert en 1928 la capacité de production d'antibiotiques de *P. rubens*, nommé à l'époque *P. notatum*, puis *P. chrysogenum* et récemment rebaptisé *P. rubens* [46], ce qui lui a valu le prix Nobel de médecine en 1945. Comme beaucoup de champignons, les champignons *Penicillium* produisent des toxines dans certaines conditions pour inhiber la croissance et la reproduction de leurs concurrents [47, 48], qui peuvent avoir un impact toxicologique sur l'homme selon les conditions d'exposition. En outre, les champignons *Penicillium* sont utilisés pour la fermentation des aliments, étant par exemple inoculés pour la production de fromage et de viande séchée [20, 49–51]. En effet, les champignons *Penicillium* utilisés dans les aliments fermentés permettent de conserver le lait cru ou la viande périssable, en transformant les aliments crus grâce à leur activité métabolique telle que la lipolyse et la protéolyse, en générant des barrières physiques à la surface, ainsi que des goûts et arômes spécifiques.

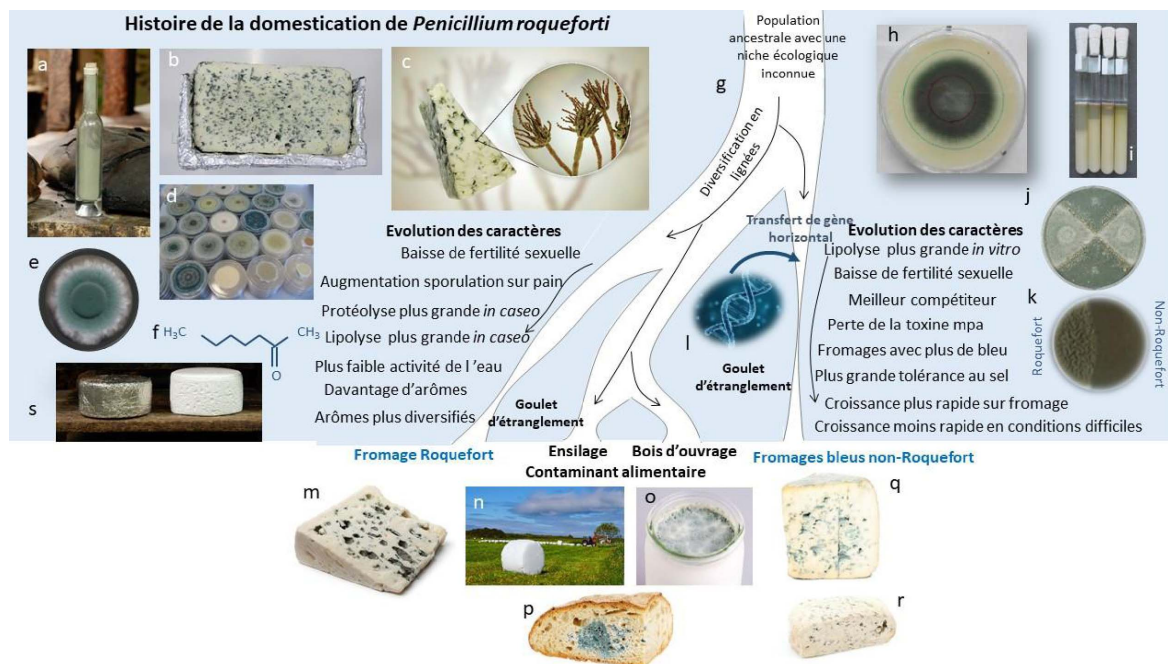
### 3. Domestication du *Penicillium roqueforti*

*Penicillium roqueforti* est l'espèce de moisissure utilisée comme ferment secondaire pour la production de Roquefort et de tous les fromages bleus [41, 48, 52], avec plus de 16 synonymes tels que *P. glaucum*, *P. gorgonzolae* ou *P. stilton*. *Penicillium roqueforti* est un champignon saprophyte parfois trouvé dans le sol ou la matière organique en décomposition, bien que son réservoir naturel soit encore inconnu. Il est

fréquemment isolé à partir de matières végétales mal stockées dans des conditions micro-aérophiles telles que l'ensilage, où il peut produire des toxines ayant des effets néfastes sur le bétail ou les humains [48, 53–57]. Sa tolérance au froid et aux acides organiques, en particulier lorsque l'activité de l'eau est supérieure à 95%, fait de cette espèce un contaminant alimentaire fréquent du pain, de la bière, de l'orge, du riz, des olives, des fruits, des fromages à pâte dure, de la farine ou de la viande séchée [58]. L'espèce a été décrite en 1906 par Charles Thom [59]. Il s'agissait en fait d'un complexe de plusieurs espèces, aujourd'hui appelé section Roquefortorum [60], comprenant également *P. carneum*, *P. paneum* et *P. psychrosexualis*, que l'on pouvait distinguer sur la base d'analyses moléculaires, morphologiques et biochimiques [61–67].

Au sein de l'espèce *P. roqueforti*, plusieurs populations différenciées ont été identifiées sur la base de marqueurs microsattellites [68, 69]. Sur la base de génomes entiers, quatre populations génétiquement et phénotypiquement différenciées ont été identifiées [70] (Figure 2). Deux populations sont utilisées pour la fabrication du fromage et les deux autres populations vivent dans d'autres environnements [70]. Parmi les deux populations de fromage, l'une comprend des souches liées à la production française de Roquefort en appellation d'origine protégée (AOP), tandis que la population non-Roquefort comprend des souches isolées dans le monde entier de tous les autres fromages bleus, comme le Stilton en Angleterre, le Cabrales en Espagne, le Gorgonzola en Italie, la Fourme d'Ambert en France. Les deux autres populations se trouvent sur le bois, l'ensilage et comme contaminants dans divers produits alimentaires (Figure 2).

Les deux populations de fromage présentent une diversité génétique beaucoup moins importante que les deux autres populations [70]. Toutefois, la population de Roquefort a conservé une diversité génétique légèrement plus riche que la population non-Roquefort, probablement en raison de l'obligation d'utiliser des souches locales dans le cadre du cahier des charges de l'AOP [70]. La perte drastique de diversité génétique dans la population non-Roquefort reflète une forte sélection pour la production à l'échelle industrielle et le commerce mondial qui ont conduit à l'utilisation d'une seule lignée clonale dans tous les fromages bleus analysés dans le monde [70].



**FIGURE 2.** Histoire de la domestication du *Penicillium roqueforti* pour la fabrication des fromages bleus. D'une population ancestrale avec une niche écologique jusqu'alors inconnue, les quatre populations identifiées de *P. roqueforti* se sont diversifiées en s'adaptant aux niches écologiques créées par l'homme, avec deux populations domestiquées indépendamment pour la fabrication de fromage, et deux populations contaminant la nourriture et l'ensilage. L'une des deux populations ne se trouve que dans les fromages appartenant à l'appellation d'origine protégée (AOP) Roquefort et abrite une certaine diversité génétique, tandis que l'autre population est une lignée clonale unique utilisée pour tous les autres fromages bleus du monde. Les goulets d'étranglement se sont produits dans toutes les populations, étant les plus forts dans les populations de fromages. Les populations fromagères ont développé des caractères spécifiques bénéfiques pour la fabrication du fromage, tels que les arômes, la coloration bleue des cavités du fromage, la lipolyse, la protéolyse et la tolérance au sel, avec des spécificités différentes dans les deux populations. La population de Roquefort produit une plus grande quantité de spores lorsqu'elle est cultivée sur du pain, le milieu de culture ancestral de Roquefort-sur-Soulzon. **a** : spores de *P. roqueforti* prélevées sur du pain inoculé, cuit auparavant en produisant des croûtes brûlées, comme on l'a longtemps fait pour l'inoculation du fromage bleu. **b** : tranche de fromage expérimental ressemblant au Roquefort. **c** : Morceau de fromage bleu avec des spores de *P. roqueforti* grossies en incrustation, portées dans les "pincesaux" typiques des champignons *Penicillium*. **d** : Boîtes de Pétri telles qu'elles sont utilisées aujourd'hui pour la culture stérile de *P. roqueforti*. **e** : Colonie de *P. roqueforti* sur une boîte de Pétri avec un milieu malté. **f** : Molécule de 2-heptatone responsable de l'un des arômes typiques du fromage bleu de Roquefort. **g** : Illustration de l'histoire de la divergence déduite des quatre populations de *P. roqueforti* identifiées, deux étant utilisées pour la fabrication du fromage et ayant souffert de forts goulets d'étranglement, une contaminant les aliments et le bois d'œuvre et une l'ensilage et également les aliments. **h** : Colonie de *P. roqueforti* sur une boîte de Pétri avec un milieu de fromage, avec des caractéristiques de mesures de taux de croissance. **i** : Tubes avec tributyrine utilisés pour mesurer le taux de lipolyse. **j** : Boîte de Pétri avec deux souches de *P. roqueforti* et structures sexuées (appelées cléistothèces, en marron) induites dans la zone de confrontation. **k** : Expérience de compétition sur une boîte de Pétri avec deux souches de *P. roqueforti*, montrant une plus grande capacité de compétition et une couleur plus foncée de la souche non-Roquefort. **l** : Acquisition horizontale de gènes dans la lignée non-Roquefort. **m** : Morceau de fromage roquefort. **n** : Balle d'ensilage dans un champ. **o** : Yaourt contaminé par *P. roqueforti*. **p** : Pain contaminé par *P. roqueforti*. **q** et **r** : Fromages bleus non Roquefort (Fourme d'Ambert et Bleu d'Auvergne). **s** : Fromages de Roquefort sur leurs étagères en bois dans les caves, l'une étant enveloppée de la typique feuille de plomb.

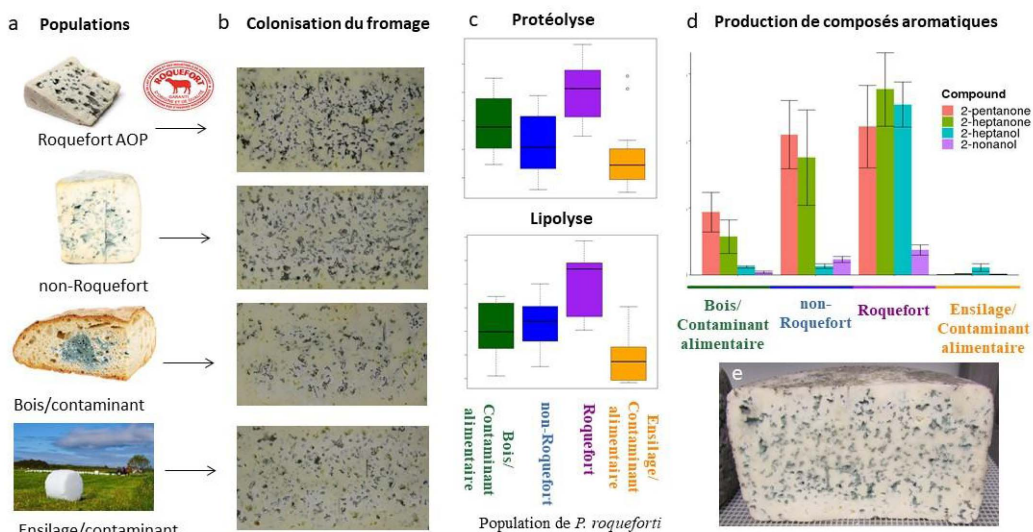
La lignée non-Roquefort a acquis de grands fragments génomiques par des transferts horizontaux de gènes [68, 71, 72]. Deux des régions transférées horizontalement semblent présenter un intérêt particulier en raison de la fonction de certains gènes qu'elles portent. Ces deux régions à transfert horizontal, appelées *Wallaby* et *CheesyTer*, ont une longueur respective de 600 et 80 kb et portent des gènes ayant des fonctions liées au métabolisme du lactose/lactate et à la compétition avec d'autres microorganismes [71, 72].

La comparaison des scénarios de divergence des populations dans le cadre du calcul approximatif bayésien a révélé que les populations d'ensilage, de bois et de contaminants alimentaires étaient les plus proches génétiquement les unes des autres, ayant divergé le plus récemment, tandis que les deux populations de fromage ont divergé plus tôt et étaient le résultat de deux événements de domestication indépendants [70] (Figure 2g). En effet, malgré un échantillonnage plus important des populations de fromage, en particulier pour la population non-Roquefort, de forts goulets d'étranglement ont été détectés dans les deux populations de fromage, étant un peu plus faibles dans la population roquefortaine [70] (Figure 2g). Deux goulets d'étranglement successifs ont même pu se produire dans la population de Roquefort [70]. Ces goulets d'étranglement sont le résultat de la domestication et plus probablement aussi de la sous-traitance récente de la production de ferments à un petit nombre de sociétés spécialisées, et peut-être de la sélection récente d'une lignée clonale performante après l'acquisition de gènes conférant des caractères avantageux par transferts horizontaux pour la population non-Roquefort [71, 72]. Les tests comparant les modèles démographiques ont également détecté un flux de gènes entre les deux populations non fromagères, mais pas avec les populations fromagères [70]. Les dates de domestication estimées à partir des derniers événements de flux de gènes entre les populations ont suggéré que la population de Roquefort aurait divergé il y a environ sept fois plus longtemps que la population non fromagère, il y a environ 760 et 140 générations, respectivement [70]. Cependant, l'estimation du temps absolu à partir du nombre de générations est difficile dans le cas des champignons, en particulier dans les lignées clonales.

Les souches de fromage ont des caractéristiques spécifiques, notamment en termes de couleur, de fertilité, de taux de croissance et de capacité à exclure d'autres micro-organismes en concurrence sur un milieu de culture [68–70, 73] (Figure 2i). Les souches non-Roquefort se développent plus rapidement sur milieu fromager que toutes les autres souches et excluent mieux les concurrents que toutes les autres souches [72] (Figure 2k). Ces capacités sont probablement liées, au moins en partie, aux régions transférées horizontalement *Wallaby* et *CheesyTer* (Figure 2), car elles portent des gènes ayant des fonctions liées au métabolisme du lactose/lactate et à la concurrence avec d'autres micro-organismes [71, 72]. La population non-Roquefort a également montré une lipolyse plus rapide, une plus grande tolérance au sel et une colonisation de la cavité du fromage [70], et la perte de la capacité à produire de l'acide mycophénolique en raison d'une délétion de 174 pb dans le gène *mpaC* [74] (Figure 2). La population de Roquefort a plus grande capacité de sporulation sur milieu à base de pain [70].

La lente croissance de la population de Roquefort sur milieu fromager et son activité lipolytique plus faible observée *in vitro* par rapport à la population non Roquefort devraient ralentir la maturation du fromage. Ce ralentissement de la maturation peut être lié à la nécessité de conserver les fromages de Roquefort, fabriqués à partir de lait de brebis disponible uniquement entre février et juillet, sur de longues périodes et historiquement sans réfrigération [76], et à l'exigence d'un affinage d'au moins 90 jours dans le cahier des charges de l'AOP Roquefort. La meilleure sporulation sur pain de la population de Roquefort fait également probablement partie du syndrome de domestication, en raison de la méthode de conservation et de production des spores de moisissure sur le pain de seigle [76–80] (Figure 2a). Les performances de la population non-Roquefort en termes de croissance sur milieu fromager [72], de colonisation des cavités fromagères, de tolérance au sel, de lipolyse *in vitro* [70], et de perte de capacité de production de toxines [74], résulteraient donc d'une sélection très récente de caractéristiques bénéfiques pour la production moderne et rapide de tous types de fromages bleus.

Une analyse de 52 composés volatils produits par les deux populations de fromages dans des modèles fromagers très simplifiés a permis d'identifier des



**FIGURE 3.** Principaux résultats d'une étude récente [75], dans laquelle des analyses de fromages expérimentaux fabriqués selon le mode de production du Roquefort (e) et par inoculation de souches provenant des quatre populations de *Penicillium roqueforti* (a) ont montré des caractères bénéfiques dans les deux populations de fromage, produisant une surface bleue plus élevée sur les tranches de fromage (b), une lipolyse et une protéolyse plus fortes *in caseo*, et en particulier dans la population de Roquefort (alors que la population non-Roquefort s'est avérée avoir le taux de lipolyse le plus élevé *in vitro* [70]), (c) et des quantités plus importantes d'arômes positifs (comme les méthylcétones, avec par exemple, dans d, la 2-pentanone et la 2-heptanone, et les alcools secondaires, avec par exemple, dans d, le 2-heptanol et le 2-nonanol, les quantités d'alcool multipliées par quatre par rapport aux méthylcétones).

profils de composés volatils significativement différents entre les deux populations de fromages, notamment en distinguant les composés typiques des fromages bleus au goût prononcé et une plus grande diversité de composés volatils produits par la population de Roquefort [70]. Des empreintes de sélection positive ont également été identifiées dans plusieurs gènes potentiellement impliqués dans la production de composés volatils [70].

Une étude récente a examiné les différences entre les populations de *P. roqueforti* dans le contexte de la domestication pour la fabrication de fromage en comparant les fromages bleus fabriqués avec les quatre populations fongiques suivant des conditions aussi proches que possible des spécifications de production commerciale de type Roquefort [75] (Figure 3). Les populations de *P. roqueforti* ont eu un impact mineur sur la diversité bactérienne des fromages et aucun sur l'abondance des principaux microorganismes. Les populations de *P. roqueforti* ont produit des fromages avec des pourcentages plus élevés de surface bleue en raison d'une croissance plus

rapide et/ou d'une meilleure sporulation, peut-être liée à une meilleure tolérance aux faibles niveaux d'oxygène. Les populations de fromages ont également produit des fromages avec différentes séries et des quantités plus élevées de composés volatils souhaités. La population Roquefort de *P. roqueforti* en particulier a produit des composés aromatiques plus variés et en plus grande quantité, ce qui a donné des notes plus fruitées et plus fromagères aux fromages; la production de ces composés étaient principalement à sa lipolyse et sa protéolyse les plus efficaces *in caseo*. Il est frappant de constater que la lipolyse et la protéolyse étaient effectivement les plus efficaces dans les fromages pour la population de *P. roqueforti* de Roquefort [75], tandis que la lipolyse était en revanche la plus efficace *in vitro* pour la population de *P. roqueforti* non Roquefort [70], ce qui montre l'importance des études *in caseo*. La population de Roquefort produit également des fromages à faible activité de l'eau, ce qui peut prévenir l'apparition de micro-organismes contaminants. Il y avait donc une forte influence des populations de *P. roqueforti* sur



plusieurs aspects importants de la sécurité, de l'apparence et de la saveur du fromage [75].

Des empreintes de dégénérescence ont été trouvées chez les populations de *P. roqueforti* domestiquées. La culture clonale de souches de *P. roqueforti* utilisées pour la production de fromage, et leur maintien isolé des populations dans d'autres environnements depuis le 19<sup>ème</sup> siècle, a entraîné une dégénérescence des fonctions sexuelles, avec une perte de fertilité à différents stades dans les deux populations de fromage [73]. La population de fromage de Roquefort a pu effectuer les premières étapes de la reproduction sexuée mais a montré une baisse de la fertilité post-zygotique, alors que la population non-Roquefort n'initie même souvent pas la reproduction sexuée [73]. Les souches non-Roquefort se développent plus lentement sur un milieu pauvre en nutriments, ce qui peut être dû à une dégénérescence des fonctions non utilisées [72]. Les importants goulets d'étranglement dans les populations de fromage [70] peuvent également avoir contribué à la dégénérescence.

#### 4. Domestication du *Penicillium camemberti*

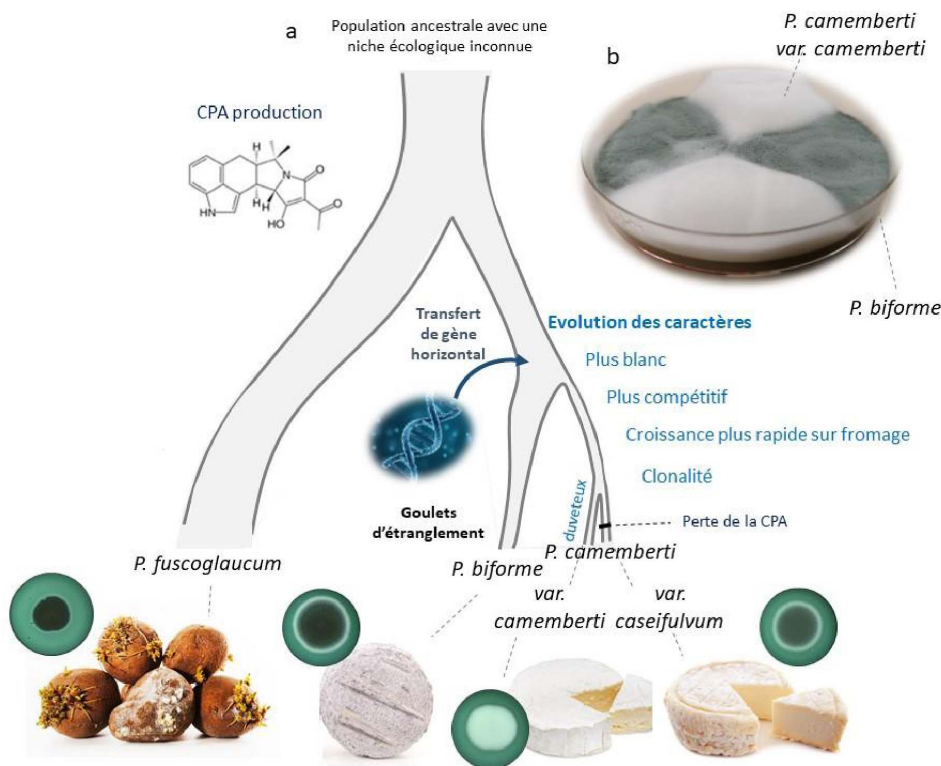
Le *Penicillium camemberti* est un autre champignon fromager emblématique, utilisé pour l'affinage des fromages à pâte molle comme le Camembert et le Brie, tout en étant phylogénétiquement éloigné de *P. roqueforti* [71]. Le complexe d'espèces *Penicillium camemberti* comprend trois espèces, *P. biforme*, *P. camemberti* et *P. fuscoglaucum*, toutes présentes dans l'environnement du fromage, *P. biforme* et *P. camemberti* étant spécifiques à l'environnement du fromage tandis que *P. fuscoglaucum* se trouve principalement dans les environnements naturels [81]. Au début du 20<sup>ème</sup> siècle, un mutant blanc a été généré dans l'industrie fromagère, par mutagenèse artificielle à partir d'un champignon gris-vert, à l'époque nommé *P. commune*, aujourd'hui divisé en deux espèces étroitement apparentées, *P. biforme* et *P. fuscoglaucum* [81, 82]. Le brie était en effet d'aspect gris-vert jusqu'au 20<sup>ème</sup> siècle. On peut le voir sur le tableau "Symphonie des fromages en brie majeur; Nature morte au fromage" de Marie Julesx Justin (1847; <https://www.pop.culture.gouv.fr/notice/joconde/M0794015842>).

Une récente étude de génomique des populations a déduit un événement de domestication ancien

séparant le champignon fromager *P. biforme* de *P. fuscoglaucum*, que l'on trouve principalement dans d'autres environnements que le fromage, par exemple le caoutchouc, les racines, le bois, les noix, l'eau ou les aliments contaminés. Un événement de domestication beaucoup plus récent a été détecté, séparant la lignée clonale de *P. camemberti* de celle de *P. biforme* [81] (Figure 4). La lignée clonale de *P. camemberti* a été trouvée dans des fromages à pâte molle de type Camembert, tandis que *P. biforme* a été trouvé dans des fromages à pâte dure comme la Tomme ou l'Ossau Iraty (Figure 4). Une adaptation phénotypique à la fabrication du fromage a été trouvée chez *P. biforme* par rapport à *P. fuscoglaucum*, avec une couleur plus blanche, une croissance plus rapide sur le milieu fromager dans des conditions de caves d'affinages (basse température et forte humidité), une tolérance au sel et des niveaux plus faibles de production de toxines de l'acide cyclopiazonique (Figure 4).

Deux lignées différenciées de *P. camemberti* ont été identifiées, avec des caractéristiques phénotypiques contrastées, qui peuvent donc être considérées comme deux variétés domestiquées différentes. L'une des variétés de *P. camemberti* contient les mêmes régions transférées que la population non Roquefort de *P. roqueforti*, *CheesyTer* et *Wallaby* [71], à l'exception de trois souches isolées du fromage camembert avant 1905 qui ne contiennent pas *CheesyTer* [81]. Cette variété de *P. camemberti*, appelée *var. camemberti*, présente une croissance radiale plus faible, mais était plus duveteuse que l'autre variété de *P. camemberti* et que *P. biforme* et *P. fuscoglaucum* (Figure 4), ce qui permet de former la croûte blanche et duveteuse typique des fromages Camembert et Brie. La variété gris-vert *P. camemberti*, appelée variété *caseifulvum*, est présente dans les fromages autres que le Camembert, comme le St Marcellin ou la Rigotte de Condrieu, porte *CheesyTer* mais pas toute la région de *Wallaby*, et est incapable de produire l'acide cyclopiazonique en raison d'une mutation de décalage du cadre de lecture dans un des gènes impliqués dans sa biosynthèse [81]. Les deux variétés de *P. camemberti* excluent les concurrents fongiques plus efficacement que *P. biforme*, la lignée blanche duveteuse de *P. camemberti* étant plus efficace à l'exclusion que la lignée gris-vert de *P. camemberti*.

Ainsi, le processus de domestication a conduit à la diversification de *P. camemberti* en deux variétés, *P.*



**FIGURE 4.** Historique de la domestication du *Penicillium camemberti* et du *P. bifforme* pour la fabrication de fromages à pâte molle ou dure (a). Issu d'une population ancestrale dont la niche écologique était jusqu'alors inconnue, *P. fuscoglaucum* a divergé d'abord, se trouvant surtout dans des environnements non fromagers, puis *P. bifforme* a été domestiqué pour la production de fromage, et plus tard les deux variétés de *P. camemberti*, var. *camemberti* et var. *caseifulvum*, avec de forts goulets d'étranglement. Le processus de domestication a conduit à une adaptation concernant la couleur (qui devient plus blanche), la capacité d'exclusion compétitive, la production de toxines (perte de la production d'acide cyclopiazonique CPA), la croissance radiale et/ou verticale (duveteuse). L'environnement d'origine des lignées est montré (de gauche à droite : racines pourries, fromage de chèvre, camembert et Saint-Marcellin), ainsi que des photos de la colonie de mycélium sur des boîtes de Pétri avec du fromage salé en milieu (coloré en vert avec un colorant alimentaire) en cercles. (b). Mycélium du *P. camemberti* var. *camemberti* blanc et du *P. bifforme* gris-vert cultivés en secteurs dans la même boîte de Pétri.

*camemberti* var. *camemberti* et *P. camemberti* var. *caseifulvum*, présentant des différenciations génétique et phénotypique (en termes de taux de croissance, de couleur, de capacités concurrentielles et de production de toxines). En outre, les deux variétés de *P. camemberti* présentaient des empreintes de goulets d'étranglement importants, représentant des lignées clonales (Figure 4a). Les deux variétés de *P. camemberti* poussent plus lentement dans des conditions difficiles que *P. bifforme* ou *P. fuscoglaucum*, ce qui suggère que l'adaptation à la fabrication de fromage

a entraîné une dégénérescence des fonctions utiles dans les environnements sauvages [81]. Une autre empreinte de la dégénérescence est le niveau général très faible de production de spores chez *P. camemberti*, encore plus chez la variété *camemberti*, ce qui pose problème pour la production de fromage.

## 5. Conclusion

Les études récentes sur *P. roqueforti* et *P. camemberti* ont montré que l'aspect et le goût typiques des

fromages bleus et des fromages à pâte molle sont le résultat d'une sélection humaine sur *P. roqueforti* et *P. camemberti*, constituant ainsi une domestication, et que différentes variétés au sein de chaque espèce ont acquis des spécificités pour la fabrication du fromage. Ceci a des implications importantes pour notre compréhension des processus d'adaptation et de domestication. Plusieurs caractéristiques d'adaptation convergente ont par exemple été trouvées dans les champignons fromagers domestiqués, les mêmes régions génomiques conférant des avantages pour la croissance du fromage ayant été acquises horizontalement à la fois chez *P. roqueforti* et *P. camemberti*, une même tolérance au sel, une plus grande capacité concurrentielle et la perte de production de toxines. Des goulots d'étranglement importants et une clonalité ont également été documentés dans les populations de fromage des deux espèces, ainsi qu'une capacité réduite à croître dans des conditions difficiles.

Les études récentes sur les champignons du fromage sont également importantes pour améliorer la production de fromage. La découverte de deux lignées clonales différentes utilisées pour la fabrication du fromage chez *P. roqueforti*, avec des phénotypes contrastés et des types sexuels opposés [70], ainsi que l'élaboration d'un protocole pour induire la reproduction sexuée [68], ouvrent la voie à l'amélioration des souches par la génération de descendants. Ces descendants peuvent en outre permettre d'associer des génotypes et des phénotypes pour

élucider la base génétique de caractères importants pour la fabrication du fromage et l'adaptation. *Penicillium camemberti* présente des traces de dégénérescence sous la forme d'un niveau de production de spores dramatiquement bas, ce qui rend difficile la réplication et l'inoculation de cette moisissure pour les producteurs de fromage. Alors que *P. camemberti* représente une seule lignée clonale avec un seul type sexuel [81], les deux types sexuels ont été trouvés chez *P. bifforme* [68], ouvrant la voie à la génération de recombinants pour l'amélioration des souches.

Les produits de viande séchée, comme le salami, le jambon et d'autres charcuteries comme les produits de viande séchée de canard ou de dinde, utilisent également des champignons *Penicillium* pour la fermentation, *P. nalgiovense* et *P. salamii* servant de ferment, ainsi que *P. camemberti* [49–51]. Il serait donc intéressant d'étudier les traces de domestication de ces champignons *Penicillium*. De plus, *P. nalgiovense* et *P. salamii* sont distants génétiquement, ce qui suggère que ces champignons *Penicillium* peuvent être des modèles idéaux pour étudier l'adaptation parallèle.

## Remerciements

Ce travail a été financé par l'ERC starting grant GenomeFun 309403, un financement de la fondation Louis D. grant (Institut de France), l'ANRT, le LIP SAS, et l'ANR FUNGADAPT (ANR-19-CE20-0002).

## English version

### 1. Introduction

Adaptive divergence (population differentiation associated with adaptation to a new ecological niche) is a key process in evolutionary biology. Studying this process is essential for understanding the adaptation of organisms to their environment and how biodiversity arises. Domestication is a case of adaptive divergence under strong human pressure and at human timescale. Charles Darwin used pigeons to illustrate the process and power of selection in yielding forms drastically different from the ancestral wild form by comparing wild and domesticated pigeons [1]. Ever since, domestication has been a good model for understanding adaptation, as it in-

volves strong and recent selection for traits that can easily be identified. This strong and recent selection by humans often leaves footprints of adaptation in genomes that are easier to detect than those left by natural selection. Comparing domesticated and non-domesticated populations at phenotypic and genomic levels can allow elucidating the genetic architecture of adaptive traits and the genomic mechanisms allowing rapid adaptation and differentiation, as shown in maize and dogs for instance [2, 3].

In domesticated crops, multiple convergent events of evolution towards similar phenotypic traits have been documented, such as the loss of seed shattering, the minimization of seed dormancy and the increase in seed size and number, mostly

arising through different genomic changes in different species [4, 5]; the loss of seed shattering is an exception, with the same genetic basis in sorghum, rice, maize and foxtail millet [6, 7]. Domesticated organisms are thus useful models for studies of convergence at both the phenotypic and genomic levels.

Domesticated species also often show degeneration for traits that are useful in the wild environment but do not provide benefit in human-made environments, which leads to relaxed selection. Domesticated species can also degenerate because of deleterious allele accumulation due to strong bottlenecks and/or to hitchhiking with positively selected alleles. Accumulation of deleterious alleles has been shown for example in rice, pigs and horses [8–11]. Cases of degeneration in fertility include decreased fertility of lateral spikelets in domesticated barley [12] and non-functional very small seeds in triploid banana cultivars [13, 14].

Microbial domestication has been much less studied than plant or animal domestication, although microorganisms have been used for making food, and have evolved new traits under human-directed selection. Studies on domesticated fungi are rare beyond the model yeast *Saccharomyces cerevisiae* used for wine and beer production. Yet, domesticated fungi constitute highly valuable eukaryotic models for the study of adaptive divergence, with their small genomes, easy access to the haploid phase and experimental tractability for *in vitro* experiments [15–19]. Many fungi have been domesticated, for food fermentation and maturation (e.g. for beer, bread, wine, dried sausage or cheese), for producing secondary metabolites used in pharmaceuticals (e.g. penicillin), or for their nutritional and gustatory values (e.g. button and shiitake mushrooms) [20]. Fermentation has been used by human populations for millennia [21, 22] as a way to preserve perishable food, and to confer new organoleptic and nutritional traits to the product. The resulting fermented products often have strong economic and cultural values. Fermentation relies on the metabolic abilities of microorganisms to transform a substrate into a product exhibiting a mixture of metabolites responsible for particular aspects, textures, flavors and nutritional traits.

Fermenting microorganisms, originally resulting from environmental contamination, have been im-

proved by humans, via an initially unintentional, and then more conscious, selection process along microbial generations for strains yielding the best products in terms of sensorial traits, shelf life and safety. This selection by humans has occurred over centuries and led to microbial domestication, i.e., genetic differentiation from wild populations [16, 17, 23] and the selection of specific phenotypes of interest for humans [24]. The study of fungi used and selected over centuries for food production can thus provide insights into adaptation processes. Yet, despite their economic and industrial importance, the fungi used by humans have been little studied, with the exception of the budding yeast *Saccharomyces cerevisiae* used in the production of beer, wine and bread [23–33], and the filamentous fungus *Aspergillus oryzae*, which is used to ferment soy and rice products in Asia [34, 35].

Fungi are yet excellent models for studying evolution and adaptation in eukaryotes, given their many experimental assets [15]. Some fungi have small genomes, are easy to culture in laboratory conditions, have short generation times, can survive long periods in the freezer, are easy to transform genetically, and fungi present a rich diversity of lineages and ecological niches. Fungi can be used to address evolutionary biology questions through combinations of complementary approaches, such as genomic and experimental methods. Studies on domesticated fungi have as a matter of fact already shed light on the genomic processes involved in adaptation to anthropic environments, including gene family expansion for specific metabolism pathways, the loss of genes no longer required in the new environment, gene acquisition by horizontal gene transfers, hybridisation and introgression [19, 23, 24, 28, 29, 32, 36]. The existence of multiple, distantly related lineages of fungi used for the production of similar food products creates an ideal opportunity for studying the repeatability of evolution, by providing several recent independent replicates of adaptation to the same ecological niche under strong selection.

The domestication of fungi for cheese-making constitute a particularly attractive model system. Indeed, cheese is one of the most popular fermented food around the world, being found in Europe, America, Asia and Middle East with different cheese varieties, such as the Anthotyros and feta in Greece,

Comté, Reblochon and Roquefort in France and Chhurpi in Nepal. Fungi are used for cheese maturation, playing a critical role in cheese fermentation, texture, color and flavor. Furthermore, the multiple phylogenetically distant fungal species thriving in the cheese ecological niche provide an exciting “natural experiment” of recent similar selection and unique opportunities to assess the repeatability of evolution, especially as several distant lineages have been used for similar products (Figure 1). Here, we review recent studies on the history and process of domestication in two emblematic cheese-making species, *Penicillium roqueforti* and *P. camemberti*, used for making blue and soft cheeses, respectively (Figure 1). After a brief overview on *Penicillium* fungi, we review the recent insights gained on their demographic history, their genetic diversity and population structure, and their phenotypic and genetic changes under domestication.

## 2. *Penicillium* fungi

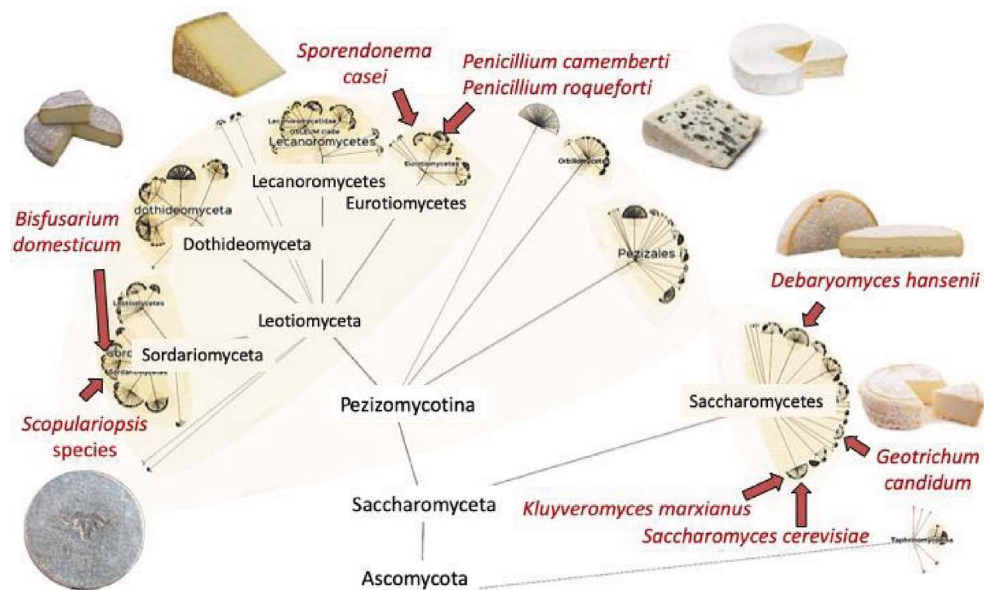
*Penicillium* is a genus of Ascomycota fungi, first described in 1809 by Link. The name *Penicillium* comes from the Latin word meaning brush, reflecting the conidiophore shape of *Penicillium* fungi [38]. The *Penicillium* genus contains 354 accepted species, isolated worldwide from various environments, e.g. soil, air and food [38–40]. Most *Penicillium* species can perform sexual and asexual reproduction. For asexual reproduction, *Penicillium* fungi generate conidia in chains (asexual spores) via specific structures called conidiophores. Conidiophores of *Penicillium* are more or less branched, with some *Penicillium* species having conidiophores only composed of phialides from which conidia go out (monoverticillate *Penicillia*), while some other species have conidiophores with one (biverticillate) or more (terverticillate *Penicillia*) branches (called metula); these conidiophore branching patterns were used by Pitt to classify *Penicillium* species before the advent of molecular identification. *Penicillium* conidia are typically mostly green or blue/green, the color however changing depending on growth conditions, such as culture medium, light, humidity and temperature [41]. Asexual spore dispersion can occur by aerial, aquatic or animal dispersion when spores germinate into hyphae that will grow and develop into a new

mycelium. *Penicillium* fungi can also undergo sexual reproduction, when hyphae from two haploid individuals carrying opposite mating types, named MAT1-1 and MAT1-2, meet under appropriate conditions. Then nuclei fuse (karyogamy) and meiosis occurs, giving rise to eight ascospores. These ascospores are enveloped into asci, themselves being inside the cleistothecium, the fruiting body of Ascomycetes.

*Penicillium* fungi are well known as saprophytic fungi decaying fruits and crops, as food spoilers, and some species are pathogens of human or food crops. For instance, *P. expansum* and *P. digitatum* are post-harvest spoiling agents of citrus and pears [42, 43], *P. marneffeii* is a human pathogen [44] *P. oxalicum*, *P. viridicatum* and *P. cyclopium* are commonly found damaging crops [45] *P. glabrum*, *P. allii*, *P. sclerotigenum*, *P. roqueforti* and *P. paneum* are common food spoilers [39, 40]. *Penicillium* fungi also have huge positive impacts on human society: Alexander Fleming accidentally discovered in 1928 the antibiotic production ability of *P. rubens* (named at that time *P. notatum*, then *P. chrysogenum* and recently renamed *P. rubens* [46]), which made him win the 1945 medicine Nobel Prize. As many fungi, *Penicillium* fungi produce toxins under certain conditions for inhibiting competitor growth and reproduction [47, 48], which can have a toxicological impact on humans depending on exposure conditions. Furthermore, *Penicillium* fungi are used for fermenting food, being for example inoculated for cheese and dry-cured meat production [20, 49–51]. Indeed, *Penicillium* fungi used in fermented food enable preserving perishable raw milk or meat, by transforming the raw food thanks to their metabolic activity such as lipolysis and proteolysis and by generating physical barriers on the surface, and also for the generation of special taste and aromas.

## 3. Domestication of *Penicillium roqueforti*

*Penicillium roqueforti* is the species of mould used as a secondary ferment for the production of Roquefort and all blue cheeses [41, 48, 52], with more than 16 synonyms such as *P. glaucum*, *P. gorgonzolae* or *P. stilton*. *Penicillium roqueforti* is a saprophytic fungus commonly found in soil or decaying organic matter, although its natural reservoir is still unknown. It



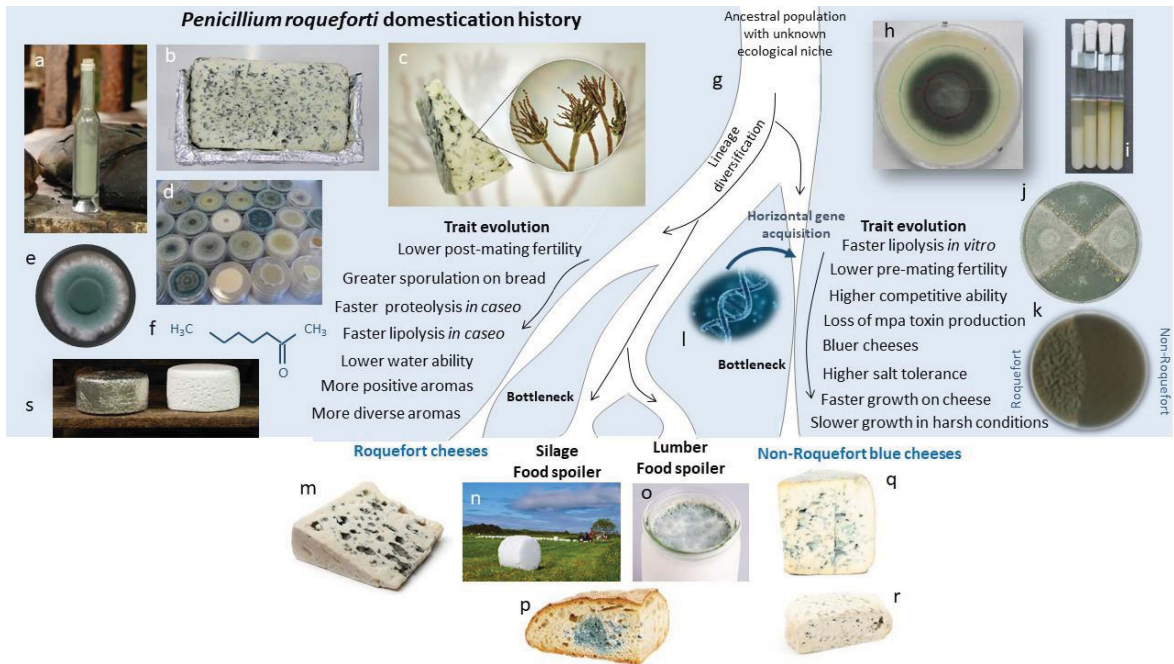
**Figure 1.** Cheese fungi are not all closely related, being instead scattered across the fungal tree of life. *Scopulariopsis* fungi can be found on Ossau-Iraty and Tommes, *Bisfusarium domesticum* on Saint-Nectaire and Reblochon, *Sporendonema casei* on Cantal and Salers, *Penicillium camemberti* on Camembert and Brie, *P. roqueforti* in all blue cheeses worldwide, *Debaryomyces hansenii*, *Geotrichum candidum*, *Kluyveromyces marxianus* and *Saccharomyces cerevisiae* in many different kinds of cheeses. Phylogeny drawn based on LifeMap [37]. Cheese pictures: *Scopulariopsis*: Ossau Iraty, *B. domesticum*: Saint-Nectaire, *S. casei*: Cantal, *P. camemberti*: Camembert, *P. roqueforti*: Roquefort, *D. hansenii*: Reblochon, *G. candidum*: goat cheese. Pictures of Saint-Nectaire, Reblochon and Cantal are from Wikipedia.

is frequently isolated from poorly stored plant material in micro-aerophilic conditions such as silage, where it can produce toxins with adverse effects on livestock or humans [48, 53–57]. Its tolerance to cold and organic acids, particularly when water activity is greater than 95%, makes this species a frequent food contaminant of bread, beer, barley, rice, olives, fruit, hard cheeses, flour or dried meat [58]. The species was described in 1906 by Charles Thom [59], and was in fact a complex of several species, now referred to as the Roquefortorum section [60], also including *P. carneum*, *P. paneum* and *P. psychrosexualis*, which could be distinguished on the basis of molecular, morphological and biochemical analyses [61–67].

Within the species *P. roqueforti*, several differentiated populations have been identified based on microsatellite markers [68, 69]. Based on whole genomes, four genetically and phenotypically differentiated populations have been identified [70] (Figure 2). Two populations are used for cheese making and the two other populations are found in

other environments [70]. Of the two cheese populations, one encompasses strains related to the French Roquefort protected designation of origin (PDO) production, while the non-Roquefort population encompasses strains isolated worldwide from all other blue cheeses, such as Stilton in England, Cabrales in Spain, Gorgonzola in Italy, Fourme d’Ambert in France. The other two populations are found on lumber, silage and as contaminants in various food products (Figure 2).

The two cheese populations display much less genetic diversity than the two other populations [70]. However, the Roquefort population has retained a slightly richer genetic diversity than the non-Roquefort population, probably due to the requirement of using local strains as part of the PDO specifications [70]. The drastic loss of genetic diversity in the non-Roquefort population reflects a strong selection for industrial scale production and worldwide trade that have led to the use of a single clonal line in all analysed blue cheeses in the world [70].



**Figure 2.** Domestication history of *Penicillium roqueforti* for making blue cheeses. From an ancestral population with a so far unknown ecological niche, the four identified *P. roqueforti* populations have diversified by adapting to human-made ecological niches, with two populations independently domesticated for cheese-making, and two populations spoiling food and silage. One of the two cheese populations is only found in cheeses belonging to the Roquefort protected designation of origin and harbors some genetic diversity while the other population is a single clonal lineage used for all other blue cheeses worldwide. Bottlenecks occurred in all populations, being strongest in the cheese populations. Cheese populations evolved specific traits beneficial for cheese-making, such as aromas, blue coloration of cheese cavities, lipolysis, proteolysis and salt tolerance, with different specificities in the two populations. The Roquefort population produces greater spore quantity when grown on bread, the ancestral medium of cultivation in Roquefort-sur-Soulzon. a: spores from *P. roqueforti* collected from inoculated bread with burned crusts, as was done for long for blue cheese inoculation. b: Slice of experimental Roquefort-like cheese. c: Piece of blue cheese with *P. roqueforti* spores magnified in inset, borne in the typical “brushes” of *Penicillium* fungi. d: Petri dishes as nowadays used for sterile *P. roqueforti* cultivation. e: *P. roqueforti* colony on a Petri dish with malt medium. f: 2-heptatone molecule responsible for one of the typical Roquefort blue cheese aromas. g: Illustration of the inferred divergence history of the four identified *P. roqueforti* populations, two being used for cheese-making and having suffered from strong bottlenecks, one thriving in spoiled food and lumber and one in silage and also contaminated food. h: *P. roqueforti* colony on a Petri dish with cheese medium, with traits of growth rate measures. i: Tubes with tributyrin used to measure lipolysis rate. j: Petri dish with two *P. roqueforti* strains and sexual structures (named cleistothecia, in brown) induced at the confrontation zone. k: Competition experiment on a Petri dish with two *P. roqueforti* strains, showing greater competitive ability and darker color of the non-Roquefort strain. l: Horizontal gene acquisition in the non-Roquefort lineage. m: Roquefort cheese piece. n: Silage ball in a field. o: Yoghurt spoiled by *P. roqueforti*. p: Bread spoiled by *P. roqueforti*. q and r: Non-Roquefort blue cheeses (Fourme d’Ambert and Bleu d’Auvergne). s: Roquefort cheeses on their wood shelves in caves, one being enveloped in the typical lead sheet.

The non-Roquefort lineage has acquired large genomic fragments by horizontal gene transfers [68, 71, 72]. Two of the horizontally-transferred regions appeared of particular interest because of the putative function of some genes that they carried. These two horizontally-transferred regions, called *Wallaby* and *CheesyTer*, are 600 and 80 kb long, respectively, and carry genes with functions related to lactose/lactate metabolism and competition against other microorganisms [71, 72].

Comparisons of population divergence scenarios under the approximate Bayesian computation framework revealed that the silage, lumber and food contaminant populations were the genetically closest to each other, having diverged most recently, while the two cheese populations diverged earlier and were the result of two independent domestication events [70] (Figure 2g). Indeed, despite a larger sampling of the cheese populations, especially for the non-Roquefort population, strong bottlenecks were detected in the two cheese populations, being a bit weaker in the Roquefort population [70] (Figure 2g). Two successive bottlenecks may even have occurred in the Roquefort population [70]. These bottlenecks are the result of domestication and also probably of the recent subcontracting of ferment production to a small number of specialist companies, and possibly to the recent selection of a high-performance clonal lineage after the acquisition of genes conferring advantageous traits by horizontal transfers for the non-Roquefort population [71, 72]. Tests comparing demographic models also detected gene flow between the two non-cheese populations, but not with the cheese populations [70]. Domestication dates estimated from the last gene flow events between the populations suggested that the Roquefort population would have diverged about seven times longer than the non-Roquefort population, approximately 760 and 140 generations ago, respectively [70]. However, estimating absolute time from generation numbers is challenging in fungi, especially in clonal lineages.

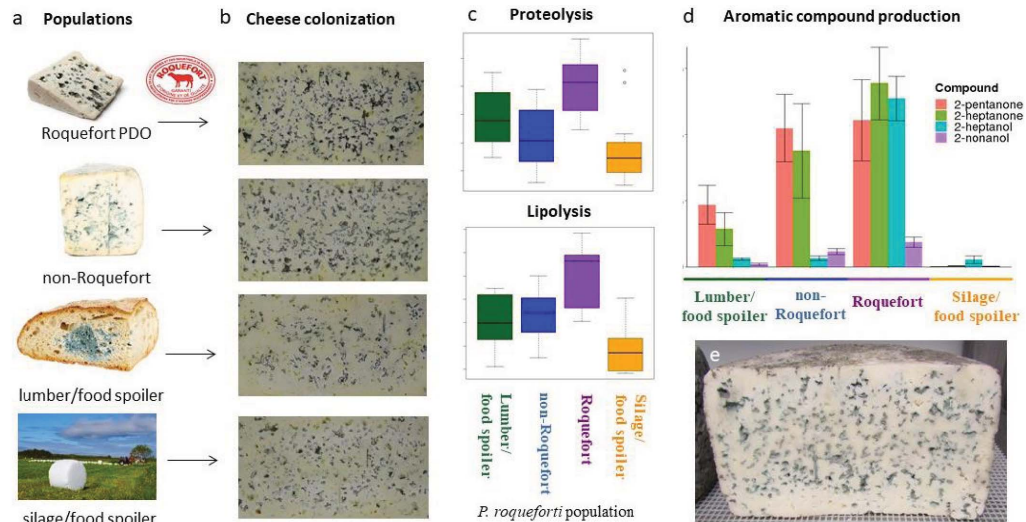
Cheese strains have specific characteristics, in particular in terms of colour, fertility, growth rate and ability to exclude other microorganisms competing on a culture medium [68–70, 73] (Figure 2). Non-Roquefort strains grow faster on cheese medium than all other strains and exclude competitors better than all other strains [72] (Figure 2k). These capabilities are likely at least partly related to the horizontally-

transferred regions *Wallaby* and *CheesyTer* (Figure 2l) as they carry genes with functions related to lactose/lactate metabolism and competition against other microorganisms [71, 72]. The non-Roquefort population also showed faster lipolysis, greater salt tolerance and cheese cavity colonization [70], and the loss of ability to produce mycophenolic acid due to a 174 bp deletion in the *mpaC* gene [74] (Figure 2). The Roquefort population showed greater sporulation on bread medium [70]. The slow growth of the Roquefort population on cheese medium and its lower lipolytic activity observed *in vitro* compared to the non-Roquefort population are expected to slow down cheese maturation. Such slow ripening may be linked to the need to preserve Roquefort cheeses, made from sheep's milk available only between February and July, over long periods and historically without refrigeration [76], and to the requirement of at least 90 days of maturation in the Roquefort PDO specifications. The better sporulation on bread of the Roquefort population is also likely part of the domestication syndrome, due to the method of conservation and production of mould spores on rye bread [76–80] (Figure 2a). The performance of the non-Roquefort population in terms of growth on cheese medium [72], colonisation of cheese cavities, salt tolerance, lipolysis *in vitro* [70], and loss of toxin production ability [74], would thus result from a very recent selection of beneficial characteristics for the modern and rapid production of all types of blue cheeses.

An analysis of 52 volatile compounds produced by the two cheese populations in highly simplified cheese models identified significantly different volatile compound profiles between the two cheese populations, notably distinguished by compounds typical of blue cheeses with a pronounced taste and a greater diversity of volatile compounds produced by the Roquefort population [70]. Footprints for positive selection have also been identified in several genes potentially involved in volatile compound production [70].

A recent study investigated the differences among *P. roqueforti* populations in the context of domestication for cheesemaking by comparing blue cheeses made with the four fungal populations following conditions as close as possible to commercial production Roquefort-type production specifications [75] (Figure 3). The *P. roqueforti* populations had a minor impact on the cheese bacterial diversity and none on





**Figure 3.** Main findings from a recent study [75], in which analyses of experimental cheeses made following the Roquefort mode of production (e) and by inoculating strains from the four *Penicillium roqueforti* populations (a) showed beneficial traits in the two cheese populations, producing a higher blue surface on cheese slices (b), stronger lipolysis and proteolysis *in caseo*, and especially the Roquefort population (while the non-Roquefort population was found to have the highest lipolysis rate *in vitro* [70]), (c) and higher quantities of positive aromas (such as methyl ketones, with for example shown in d the 2-pentanone and the 2-heptanone, and secondary alcohols, with for example shown in d the 2-heptanol and the 2-nonanol, with alcohol quantities multiplied by four compared to methyl ketones to be visible on the same axis).

the main microorganism abundance. The cheese *P. roqueforti* populations produced cheeses with higher percentages of blue area due to faster growth and/or better sporulation, possibly related to a better tolerance to low oxygen levels. The cheese populations also produced cheeses with different sets and higher quantities of desired volatile compounds. The Roquefort *P. roqueforti* population in particular produced higher diversities and quantities of aromatic compounds resulting in more fruity and cheesy notes in cheeses, which was mostly due to its most efficient lipolysis and proteolysis *in caseo*. Strikingly, lipolysis and proteolysis were indeed found highest in cheeses for the Roquefort *P. roqueforti* population [75], while lipolysis was in contrast found highest *in vitro* for the non-Roquefort *P. roqueforti* population [70], showing the importance of studies *in caseo*. The Roquefort population also produced cheeses with lower water activity, which can prevent the occurrence of spoiler microorganisms. There was thus a strong influence of *P. roqueforti* populations on several important aspects of cheese safety, appearance

and flavour [75]. Footprints of degeneration have been found in *P. roqueforti* domesticated. The clonal culture of *P. roqueforti* strains used for cheese production, and their maintenance isolated from populations in other environments since the 19th century, has led to degeneration of sexual functions, with loss of fertility at different stages in the two cheese populations [73]. The Roquefort cheese population could perform the first steps of sexual reproduction but showed decreased post-zygotic fertility, while the non-Roquefort population even often not even initiate sexual reproduction [73]. Non-Roquefort strains grow more slowly on nutrient-poor medium, which may be due to degeneration on unused functions [72]. The strong bottlenecks in cheese populations [70] may have also contributed to degeneration.

#### 4. Domestication of *Penicillium camemberti*

*Penicillium camemberti* is another emblematic cheese-making fungus, used for maturing soft cheeses such as Camembert and Brie, while being

phylogenetically distantly related to *P. roqueforti* [71]. The *Penicillium camemberti* species complex includes three species, *P. biforme*, *P. camemberti* and *P. fuscoglaucum*, all present in the cheese environment, *P. biforme* and *P. camemberti* being specific to the cheese environment while *P. fuscoglaucum* is mostly found in natural environments [81]. At the beginning of the 20th century, a white mutant was generated in the cheese industry, by artificial mutagenesis from a grey-green fungus at that time named *P. commune*, now split into two closely related species, *P. biforme* and *P. fuscoglaucum* [81, 82]. The Brie cheeses indeed used to be of grey-green aspect until the 20th century. This can be seen on the painting “Symphonie des fromages en brie majeur; Nature morte au fromage” by Marie Julesx Justin (1847; <https://www.pop.culture.gouv.fr/notice/joconde/M0794015842>).

A recent population genomic study inferred an ancient domestication event separating the cheese-making *P. biforme* fungus from *P. fuscoglaucum*, mostly found in other environments than cheese, e.g., rubber, root, wood, walnut, water or contaminated food. A much more recent domestication event was detected, separating the *P. camemberti* clonal lineage from *P. biforme* [81] (Figure 4). The *P. camemberti* clonal lineage was found in Camembert-like soft cheeses, while *P. biforme* was found in hard cheeses like Tomme or Ossau Iraty (Figure 4). Phenotypic adaptation to cheese-making was found in *P. biforme* as compared to *P. fuscoglaucum*, with a whiter color, faster growth on cheese medium under cave conditions (low temperature, and high humidity), salt tolerance and lower levels of the cyclopiazonic acid toxin production (Figure 4).

Two differentiated *P. camemberti* lineages were identified, with contrasting phenotypic features, that can therefore be considered as two different domesticated varieties. One of the *P. camemberti* varieties contained the same transferred regions as the non-Roquefort population *P. roqueforti*, *CheesyTer* and *Wallaby* [71], except three strains isolated from Camembert cheese before 1905 which do not carry *CheesyTer* [81]. This *P. camemberti* variety, called var. *camemberti*, displays weaker radial growth, but was fluffier than the *P. camemberti* other variety and than *P. biforme* and *P. fuscoglaucum* (Figure 4), which allows forming the white fluffy crust typical of Camembert and Brie cheeses. The grey-green *P. camem-*

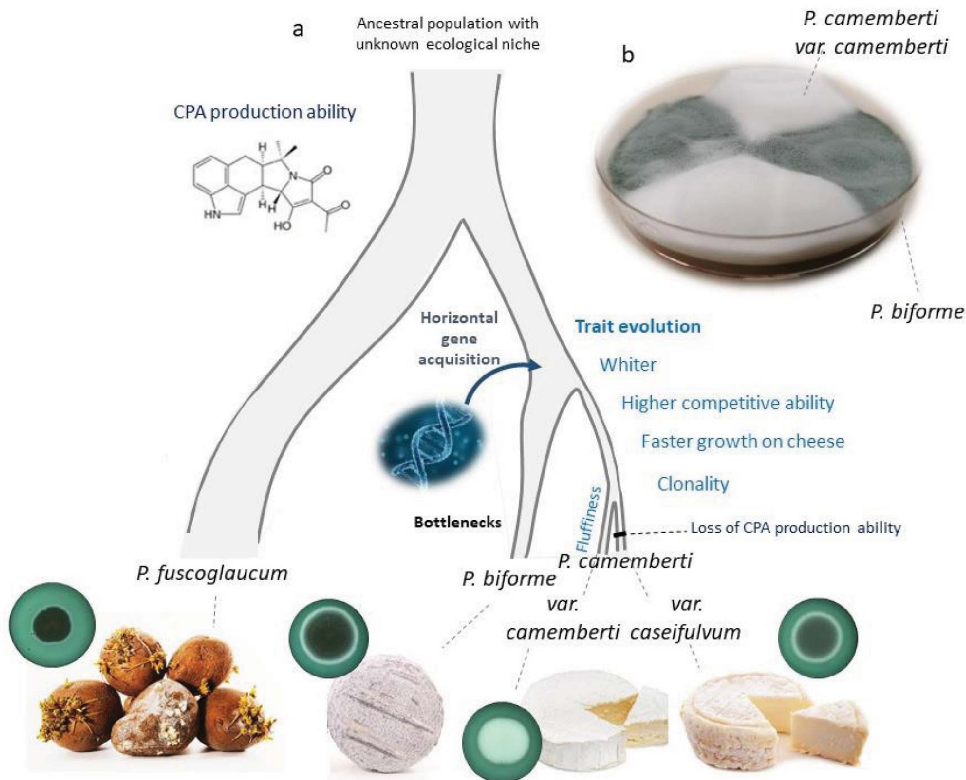
*berti* variety, called the *caseifulvum* variety, is found in cheeses other than Camembert, such as St Marcellin or Rigotte de Condrieu, carries *CheesyTer* but not the entire *Wallaby* region, and is unable to produce the cyclopiazonic acid toxin due to a frameshift mutation in one of the genes involved in its biosynthesis [81]. Both *P. camemberti* varieties exclude fungal competitors more effectively than *P. biforme*, the fluffy white Camembert lineage being more effective at exclusion than the grey-green *P. camemberti* lineage.

Footprints of domestication are thus found in *P. camemberti* and *P. biforme*, with genetic differentiation between cheese and strains from other environments, and the identification of specific phenotypic traits beneficial for cheese-making. Furthermore, the domestication process led to diversification of *P. camemberti* into two varieties, *P. camemberti* var. *camemberti* and *P. camemberti* var. *caseifulvum*, also displaying genetic and phenotypic differentiation (in terms of growth rate, color, competitive abilities and toxin production).

Furthermore, both *P. camemberti* varieties showed footprints of strong bottlenecks, representing clonal lineages (Figure 4a). Both *P. camemberti* varieties grew more slowly in harsh conditions than *P. biforme* or *P. fuscoglaucum*, suggesting that adaptation to cheese making has led to a degeneration of functions useful in wild environments [81]. Another degeneration footprint is the general very low spore production level in *P. camemberti*, even more in the *camemberti* variety, which raises issues for cheese production.

## 5. Conclusion

The recent studies on *P. roqueforti* and *P. camemberti* have shown that the typical appearance and flavour of blue cheeses and soft cheeses are the result of human selection both on *P. roqueforti* and *P. camemberti*, thus constituting domestication, and different varieties within each species have acquired specificities for cheese-making. This has important implications for our understanding of adaptation and domestication processes. Several footprints of convergent adaptation have for example been found in the domesticated cheese fungi, with the same genomic regions conferring benefits for growth on cheese having been horizontally acquired both in *P. roqueforti*



**Figure 4.** Domestication history of *Penicillium camemberti* and *P. bifforme* for making soft or hard cheeses (a). From an ancestral population with a so far unknown ecological niche, *P. fuscoglaucum* first diverged, thriving mostly in non-cheese environments, then *P. bifforme* was domesticated for cheese production, and later the two *P. camemberti* varieties, var. *camemberti* and var. *caseifulvum*, with strong bottlenecks. The domestication process led to adaptation regarding color (becoming whiter), competitive ability, toxin production (loss of CPA—cyclopiazonic acid—production), radial and/or vertical (fluffiness) growth. Environment of origin of the lineages are shown (from left to right: rotten roots, goat cheese, Camembert cheese and Saint-Marcellin cheese), as well as pictures of mycelium colony on Petri dishes with salted cheese medium (colored in green with a food colorant) in circles. (b). Mycelia of the white fluffy *P. camemberti* var. *camemberti* and the grey-green *P. bifforme* grown as sectors in the same Petri dish.

and *P. camemberti*, salt tolerance, higher competitive ability and the loss of toxin production. Strong bottlenecks and clonality were also documented in the cheese populations of the two species, as well as the reduced ability to grow in harsh conditions.

The recent studies on cheese fungi are also important for improving cheese production. The finding of two different clonal lineages used for cheese-making in *P. roqueforti*, with contrasted phenotypes and opposite mating types [70], together with the development of a protocol for inducing sexual reproduction [68], pave the way for strain improvement by generat-

ing progenies. Such progenies can further allow associating genotypes and phenotypes for elucidating the genetic basis for important traits for cheese-making and adaptation. *Penicillium camemberti* shows footprints of degeneration in the form of a dramatically low level of spore production, which renders difficult the replication and inoculation of this mold for cheese producers. While *P. camemberti* represents a single clonal lineage harboring a single mating type [81], both mating types were found in *P. bifforme* [68], paving the way to generate new progenies for strain improvement.

Dry-cured meat products, such as salami, ham and other charcuteries like duck or turkey dry-cured meat products, also use *Penicillium* fungi for fermentation, *P. nalgiovense* and *P. salamii* serving as a starter, and also *P. camemberti* [49–51]. It would thus be interesting to study the footprints of domestication in these *Penicillium* fungi. Furthermore, *P. nalgiovense* and *P. salamii* are distantly related, which suggests that these *Penicillium* fungi can be ideal models to study parallel adaptation.

## Références / References

- [1] C. Darwin, *The Origin of Species*, John Murray, London, 1859.
- [2] H. Wang, A. J. Studer, Q. Zhao, R. Meeley, J. F. Doebley, « Evidence that the origin of naked kernels during maize domestication was caused by a single amino acid substitution in *tg1* », *Genetics* **200** (2015), n° 3, p. 965-974.
- [3] A. H. Freedman, K. E. Lohmueller, R. K. Wayne, « Evolutionary history, selective sweeps, and deleterious variation in the dog », *Annu. Rev. Ecol. Evol. System.* **47** (2016), n° 1, p. 73-96.
- [4] B. Gross, K. Olsen, « Genetic perspectives on crop domestication », *Trends Plant Sci.* **15** (2010), p. 529-537.
- [5] N. Martínez-Ainsworth, M. Tenaillon, « Superheroes and masterminds of plant domestication », *C. R. Biol.* **339** (2016), p. 268-273.
- [6] Z. W. Lin, X. R. Li, L. M. Shannon, C. T. Yeh, M. L. Wang, G. H. Bai, Z. Peng, J. R. Li, H. N. Trick, T. E. Clemente *et al.*, « Parallel domestication of the *Shattering1* genes in cereals », *Nat. Genet.* **44** (2012), n° 6, p. 720-724.
- [7] G. Q. Jia, X. H. Huang, H. Zhi, Y. Zhao, Q. Zhao, W. J. Li, Y. Chai, L. F. Yang, K. Y. Liu, H. Y. Lu *et al.*, « A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*) », *Nat. Genet.* **45** (2013), n° 8, p. 957-961.
- [8] J. Lu, T. Tang, H. Tang, J. Z. Huang, S. H. Shi, C. I. Wu, « The accumulation of deleterious mutations in rice genomes : a hypothesis on the cost of domestication », *Trends Genet.* **22** (2006), n° 3, p. 126-131.
- [9] G. Larson, D. Q. Fuller, « The evolution of animal domestication », *Annu. Rev. Ecol. Evol. Syst.* **45** (2014), p. 115-136.
- [10] M. Schubert, H. Jonsson, D. Chang, C. D. Sarkissian, L. Ermini, A. Ginolhac, A. Albrechtsen, I. Dupanloup, A. Foucal, B. Petersen *et al.*, « Prehistoric genomes reveal the genetic foundation and cost of horse domestication », *Proc. Natl Acad. Sci. USA* **111** (2014), n° 52, p. E5661-E5669.
- [11] C. D. Marsden, V. D. Ortega-Del, D. P. O'Brien, J. F. Taylor, O. Ramirez, C. Vila, T. Marques-Bonet, R. D. Schnabel, R. K. Wayne, K. E. Lohmueller, « Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs », *Proc. Natl Acad. Sci. USA* **113** (2016), n° 1, p. 152-157.
- [12] J. H. Peng, D. Sun, E. Nevo, « Domestication evolution, genetics and genomics in wheat », *Mol. Breed.* **28** (2011), p. 281.
- [13] J. Heslop-Harrison, T. Schwarzacher, « Domestication, genomics and the future for banana », *Ann. Bot.* **100** (2007), p. 1073-1084.
- [14] X. Perrier, E. De Langhe, M. Donohue, C. Lentfer, L. Vrydaghs, F. Bakry, F. Carreel, I. Hippolyte, J.-P. Horry, C. Jenny *et al.*, « Multidisciplinary perspectives on banana *Musa domestica* », *Proc. Natl Acad. Sci. USA* **108** (2011), n° 28, p. 11311-11318.
- [15] P. Gladieux, J. Ropars, H. Badouin, A. Branca, G. Aguilera, D. M. De Vienne, R. C. R. De La Vega, S. Branco, T. Giraud, « Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes », *Mol. Ecol.* **23** (2014), n° 4, p. 753-773.
- [16] M. Gonçalves, A. Pontes, P. Almeida, R. Barbosa, M. Serra, D. Libkind, M. Hutzler, P. Gonçalves, J. P. Sampaio, « Distinct domestication trajectories in top-fermenting beer yeasts and wine yeasts », *Curr. Biol.* **26** (2016), p. 2750-2761.
- [17] P. Almeida, R. Barbosa, D. Bensasson, P. Gonçalves, J. Sampaio, « Adaptive divergence in wine yeasts and their wild relatives suggests a prominent role for introgressions and rapid evolution at noncoding sites », *Mol. Ecol.* **26** (2017), p. 2167-2182.
- [18] B. Carbonetto, J. Ramsayer, T. Nidelet, J. Legrand, D. Sicard, « Bakery yeasts, a new model for studies in ecology and evolution », *Yeast* **35** (2018), p. 591-603.
- [19] J.-L. Legras, V. Galeote, F. Bigey, C. Camarasa, S. Marsit, T. Nidelet, I. Sanchez, A. Couloux, J. Guy, R. Franco-Duarte *et al.*, « Adaptation of *S. cerevisiae* to fermented food environments reveals remarkable genome plasticity and the footprints of domestication », *Mol. Biol. Evol.* **35** (2018), n° 7, p. 1712-1727.
- [20] J. Dupont, S. Dequin, T. Giraud, F. Le Tacon, S. Marsit, J. Ropars, F. Richard, M. Selosse, « Fungi as a source of food », *Microbiol. Spectr.* **5** (2017), n° 3.
- [21] P. McGovern, M. Jalabadze, S. Batiuk, M. P. Callahan, K. E. Smith, G. R. Hall, E. Kvavadze, D. Maghradze, N. Rusishvili, L. Bouby *et al.*, « Early Neolithic wine of Georgia in the South Caucasus », *Proc. Natl Acad. Sci. USA* **114** (2017), n° 48, p. E10309-E10318.
- [22] E. Greco, O. El-Aguizy, M. F. Ali, S. Foti, V. Cunsolo, R. Salletti, E. Ciliberto, « Proteomic analyses on an ancient Egyptian Cheese and Biomolecular Evidence of Brucellosis », *Anal. Chem.* **90** (2018), n° 16, p. 9673-9676.
- [23] P. Almeida, C. Gonçalves, S. Teixeira, D. Libkind, M. Bontrager, I. Masneuf-Pomarède, W. Albertin, P. Durrens, S. D. James, P. Marullo *et al.*, « A Gondwanan imprint on global diversity and domestication of wine and cider yeast *Saccharomyces uvarum* », *Nat. Commun.* **5** (2014), p. 4044.
- [24] S. Marsit, A. Mena, F. Bigey, F.-X. Sauvage, A. Couloux, J. Guy, J.-L. Legras, E. Barrio, S. Dequin, V. Galeote, « Evolutionary advantage conferred by an eukaryote-to-eukaryote gene transfer event in wine yeasts », *Mol. Biol. Evol.* **32** (2015), n° 7, p. 1695-1707.
- [25] J. Fay, J. Benavides, « Evidence for domesticated and wild populations of *Saccharomyces cerevisiae* », *PLoS Genet.* **1** (2005), p. e5.
- [26] J.-L. Legras, D. Merdinoglu, J.-M. Cornuet, F. Karst, « Bread, beer and wine : *Saccharomyces cerevisiae* diversity reflects human history », *Mol. Ecol.* **16** (2007), p. 2091-2102.
- [27] G. Liti, D. M. Carter, A. M. Moses, J. Warringer, L. Parts, S. A. James, R. P. Davey, I. N. Roberts, A. Burt, V. Koufopanou *et al.*, « Population genomics of domestic and wild yeasts », *Nature* **458** (2009), n° 7236, p. 337-341.

- [28] M. Novo, F. Bigey, E. Beyne, V. Galeote, F. Gavory, S. Mallet, B. Cambon, J.-L. Legras, P. Wincker, S. Casaregola *et al.*, «Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118», *Proc. Natl Acad. Sci. USA* **106** (2009), p. 16333-16338.
- [29] D. Libkind, C. Hittinger, E. Valério, C. Gonçalves, J. Dover, M. Johnston, P. Gonçalves, J. Sampaio, «Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast», *Proc. Natl Acad. Sci. USA* **108** (2011), p. 14539-14544.
- [30] B. Dunn, T. Paulish, A. Stanbery, J. Piotrowski, G. Koniges, E. Kroll, E. J. Louis, G. Liti, G. Sherlock, F. Rosenzweig, «Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression», *PLoS Genet.* **9** (2013), n° 3, article ID e1003366.
- [31] A. Spor, D. J. Kvittek, T. Nidelet, J. Martin, J. Legrand, C. Dillmann, A. Bourgeois, D. de Vienne, G. Sherlock, D. Sicard, «Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast», *Evolution* **68** (2014), n° 3, p. 772-790.
- [32] B. Gallone, J. Steensels, T. Prah, L. Soriaga, V. Saels, B. Herrera-Malaver, A. Merlevede, M. Roncoroni, K. Voordeckers, L. Miraglia *et al.*, «Domestication and divergence of *lib* beer yeasts», *Cell* **166** (2016), p. 1397-1410.
- [33] J. Peter, J. Schacherer, «Population genomics of yeasts : towards a comprehensive view across a broad evolutionary scale», *Yeast* **33** (2016), n° 3, p. 73-81.
- [34] J. E. Galagan, S. E. Calvo, C. Cuomo, L. J. Ma, J. R. Wortman, S. Batzoglu, S. I. Lee, M. Basturkmen, C. C. Spevak, J. Clutterbuck *et al.*, «Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*», *Nature* **438** (2005), n° 7071, p. 1105-1115.
- [35] J. G. Gibbons, L. Salichos, J. C. Slot, D. C. Rinker, K. L. McGary, J. G. King, M. A. Klich, D. L. Tabb, W. H. McDonald, A. Rokas, «The evolutionary imprint of domestication on genome variation and function of the filamentous fungus *Aspergillus oryzae*», *Curr. Biol.* **22** (2012), n° 15, p. 1403-1409.
- [36] A. R. Borneman, A. H. Forgan, R. Kolouchova, J. A. Fraser, S. A. Schmidt, «Whole genome comparison reveals high levels of inbreeding and strain redundancy across the spectrum of commercial wine strains of *Saccharomyces cerevisiae*», *G3-Genes Genomes Genet.* **6** (2016), n° 4, p. 957-971.
- [37] D. M. de Vienne, «Lifemap : Exploring the entire tree of life», *PLoS Biol.* **14** (2016), n° 12, article ID e2001624.
- [38] C. M. Visagie, J. Houbbraken, J. C. Frisvad, S.-B. Hong, C. H. W. Klaassen, G. Perrone, K. A. Seifert, J. Varga, T. Yaguchi, R. A. Samson, «Identification and nomenclature of the genus *Penicillium*», *Stud. Mycol.* **78** (2014), p. 343-371.
- [39] O. Filtenborg, J. C. Frisvad, U. Thrane, «Moulds in food spoilage», *Specific Spoilage Organisms* **33** (1996), p. 85-102.
- [40] J. Dijksterhuis, «The fungal spore and food spoilage», *Curr. Opin. Food Sci.* **17** (2017), p. 68-74.
- [41] R. A. Samson, J. Houbbraken, U. Thrane, J. C. Frisvad, B. Andersen, *Food and Indoor Fungi*, CBS-KNAW Fungal Biodiversity Centre, 2010.
- [42] D. R. L. Caccioni, M. Guizzardi, D. M. Biondi, A. Renda, G. Ruberto, «Relationship between volatile components of citrus fruit essential oils and antimicrobial action on *Penicillium digitatum* and *Penicillium italicum*», *Int. J. Food Microbiol.* **43** (1998), p. 73-79.
- [43] M. Marcet-Houben, A.-R. Ballester, B. de la Fuente, E. Harries, J. F. Marcos, L. González-Candelas, T. Gabaldón, «Genome sequence of the necrotrophic fungus *Penicillium digitatum*, the main postharvest pathogen of citrus», *BMC Genom.* **13** (2012), p. 646.
- [44] K. F. LoBuglio, J. W. Taylor, «Phylogeny and PCR identification of the human pathogenic fungus *Penicillium marneffei*», *J. Clin. Microbiol.* **33** (1995), p. 85.
- [45] P. Mislivec, «Toxic species of *Penicillium* common in food», *J. Food Prot.* **44** (1981), p. 723-726.
- [46] J. Houbbraken, J. C. Frisvad, R. A. Samson, «Fleming's penicillin producing strain is not *Penicillium chrysogenum* but *P. rubens*», *IMA Fungus* **2** (2011), n° 1, p. 87-95.
- [47] N. Hymery, V. Vasseur, M. Coton, J. Mounier, J.-L. Jany, G. Barbier, E. Coton, «Filamentous Fungi and Mycotoxins in Cheese : A Review», *Compr. Rev. Food Sci. Food Saf.* **13** (2014), n° 4, p. 437-456.
- [48] E. Coton, M. Coton, N. Hymery, J. Mounier, J.-L. Jany, «*Penicillium roqueforti* : an overview of its genetics, physiology, metabolism and biotechnological applications», *Fungal Biol. Rev.* **34** (2020), p. 59-73.
- [49] L. O. Sunesen, L. H. Stahnke, «Mould starter cultures for dry sausages-selection, application and effects», *Meat. Sci.* **65** (2003), n° 3, p. 935-948.
- [50] D. Magistà, M. Ferrara, M. Del Nobile, D. Gammariello, A. Conte, G. Perrone, «*Penicillium salamii* strain ITEM 15302 : A new promising fungal starter for salami production», *Int. J. Food Microbiol.* **16** (2016), p. 33-41.
- [51] R. Chávez, F. Fierro, R. O. García-Rico, F. Laich, 2019. «Mould-fermented foods : *Penicillium* spp. as ripening agents in the elaboration of cheese and meat products». *Mycofactories*, 73-98.
- [52] M. D. Cantor, T. van den Tempel, T. K. Hansen, Y. Ardö, «Blue Cheese», in *Cheese : Chemistry, Physics and Microbiology* (P. F. Fox, P. L. H. McSweeney, T. M. Cogan, T. P. Guinee, eds.), Elsevier Academic Press, London, 2004, p. 175-198.
- [53] S. Ohmomo, H. K. Kitamoto, T. Nakajima, «Detection of roquefortines in *Penicillium roqueforti* isolated from moulded maize silage», *J. Sci. Food Agric.* **64** (1994), p. 211-215.
- [54] M. O'Brien, K. F. Nielsen, P. O'Kiely, P. D. Forristal, H. T. Fuller, J. C. Frisvad, «Mycotoxins and other secondary metabolites produced in vitro by *Penicillium paneum* Frisvad and *Penicillium roqueforti* Thom isolated from baled grass silage in Ireland», *J. Agric. Food Chem.* **54** (2006), p. 9268-9276.
- [55] M. O'Brien, D. Egan, P. O'kiely, P. D. Forristal, F. M. Doohan, H. T. Fuller, «Morphological and molecular characterisation of *Penicillium roqueforti* and *P. paneum* isolated from baled grass silage», *Mycol. Res.* **112** (2008), p. 921-932.
- [56] C. W. Basse, «Mitochondrial inheritance in fungi», *Curr. Opin. Microbiol.* **13** (2010), n° 6, p. 712-719.
- [57] V. A. Alonso, C. M. Pereyra, L. A. M. Keller, A. M. Dalcerro, C. A. R. Rosa, S. M. Chiacchiera, L. R. Cavaglieri, «Fungi and mycotoxins in silage : an overview», *J. Appl. Microbiol.* **115** (2013), p. 637-643.
- [58] J. I. Pitt, A. D. Hocking, *Fungi and Food Spoilage*, 2nd éd., Blackie Academic & Professional, London, 1997.

- [59] C. Thom, « Fungi in cheese ripening : Camembert and Roquefort », *US Dep. Agric. Bur. Anim. Ind.* **82** (1906), p. 1-39.
- [60] J. Houbraken, L. Wang, H. B. Lee, J. C. Frisvad, « New sections in *Penicillium* containing novel species producing patulin, pyripyropens or other bioactive compounds », *Persoonia - Mol. Phylog. Evol. Fungi* **36** (2016), n° 1, p. 299-314.
- [61] J. C. Frisvad, O. Filtenborg, « Terverticillate *Penicillia* : chemotaxonomy and mycotoxin production », *Mycologia* **81** (1989), p. 837-861.
- [62] M. Boysen, P. Skouboe, J. Frisvad, L. Rossen, « Reclassification of the *Penicillium roqueforti* group into three species on the basis of molecular genetic and biochemical profiles », *Microbiol. Spect.* **142** (1996), p. 541-549.
- [63] J. C. Frisvad, R. A. Samson, « Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*- A guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins », *Stud. Mycol.* **49** (2004), p. 1-174.
- [64] R. A. Samson, K. A. Seifert, A. F. A. Kuijpers, J. Houbraken, J. C. Frisvad, « Phylogenetic analysis of *Penicillium* subgenus *Penicillium* using partial  $\beta$ -tubulin sequences », *Stud. Mycol.* **49** (2004), p. 175-200.
- [65] K. Karlshøj, T. O. Larsen, « Differentiation of species from the *Penicillium roqueforti* group by volatile metabolite profiling », *J. Agric. Food Chem.* **53** (2005), p. 708-715.
- [66] J. Houbraken, J. C. Frisvad, R. A. Samson, « Sex in *Penicillium* series *Roqueforti* », *IMA Fungus* **1** (2010), p. 171-180.
- [67] J. Houbraken, R. A. Samson, « Phylogeny of *Penicillium* and the segregation of *Trichocomaceae* into three families », *Stud. Mycol.* **70** (2011), p. 1-51.
- [68] J. Ropars, M. Lopez-Villavicencio, J. Dupont, A. Snirc, G. Gillot, M. Coton, J. L. Jany, E. Coton, T. Giraud, « Induction of sexual reproduction and genetic diversity in the cheese fungus *Penicillium roqueforti* », *Evol. Appl.* **7** (2014), n° 4, p. 433-441.
- [69] G. Gillot, J. L. Jany, M. Coton, G. Le Floch, S. Debaets, J. Ropars, M. Lopez-Villavicencio, J. Dupont, A. Branca, T. Giraud *et al.*, « Insights into *Penicillium roqueforti* morphological and genetic diversity », *PLoS One* **10** (2015), n° 6, article ID e0129849.
- [70] E. Dumas, A. Feurtey, R. C. Rodríguez de la Vega, S. Le Prieur, A. Snirc, M. Coton, A. Thierry, E. Coton, M. Le Piver, D. Roueyre *et al.*, « Independent domestication events in the blue-cheese fungus *Penicillium roqueforti* », *Mol. Ecol.* **29** (2020), p. 2639-2660.
- [71] K. Cheeseman, J. Ropars, P. Renault, J. Dupont, J. Gouzy, A. Branca, A. L. Abraham, M. Ceppi, E. Conseiller, R. Debuchy *et al.*, « Multiple recent horizontal transfers of a large genomic region in cheese making fungi », *Nat. Commun.* **5** (2014), article ID 2876.
- [72] J. Ropars, R. C. R. de la Vega, M. Lopez-Villavicencio, J. Gouzy, E. Sallet, E. Dumas, S. Lacoste, R. Debuchy, J. Dupont, A. Branca *et al.*, « Adaptive horizontal gene transfers between multiple cheese-associated fungi », *Curr. Biol.* **25** (2015), n° 19, p. 2562-2569.
- [73] J. Ropars, Y. C. Lo, E. Dumas, A. Snirc, D. Begerow, T. Rollnik, S. Lacoste, J. Dupont, T. Giraud, M. Lopez-Villavicencio, « Fertility depression among cheese-making *Penicillium roqueforti* strains suggests degeneration during domestication », *Evolution* **70** (2016), n° 9, p. 2099-2109.
- [74] G. Gillot, J. Jany, E. Poirier, M. Maillard, S. Debaets, A. Thierry, E. Coton, M. Coton, « Functional diversity within the *Penicillium roqueforti* species », *Int. J. Food Microbiol.* **241** (2017), p. 141-150.
- [75] T. Caron, M. Le Piver, A.-C. Péron, P. Lieben, R. Lavigne, S. Brunel, D. Roueyre, M. Place, P. Bonnarne, T. Giraud *et al.*, 2020. « Strong effect of *Penicillium roqueforti* populations on volatile and metabolic compounds responsible for aromas, flavour and texture in blue cheeses ». *BioRxiv* <https://doi.org/10.1101/2020.03.02.974352>.
- [76] E. Marre, *Le roquefort*, E. Carrère, Rodez, 1906.
- [77] R. Aussibal, *Les caves de roquefort*, Cahors, France, 1983.
- [78] M. Labbe, J. P. Serres, *Chroniques du Roquefort-De la préhistoire à l'aube industrielle*, Graphi Imprimeur, La Primaube, France, 2004.
- [79] M. Labbe, J. P. Serres, *Chroniques du Roquefort-Des hommes, des entreprises, des marques, période moderne*, Graphi Imprimeur, La Primaube, France, 2009.
- [80] S. Vabre, *Le sacre du Roquefort. Lémergence d'une industrie agroalimentaire (fin XVIIIe-1925)*. 588 pp, ISBN : 9782753535572. Collection Table des Hommes, 2015.
- [81] J. Ropars, E. Didiot, R. C. Rodríguez de la Vega, B. Bennetot, M. Coton, E. Poirier, E. Coton, A. Snirc, S. Le Prieur, T. Giraud, « Domestication of the emblematic white cheese-making fungus *Penicillium camemberti* and its diversification into two varieties », *Curr. Biol.* (2020), in press.
- [82] F. Giraud, T. Giraud, G. Aguilera, E. Fournier, R. Samson, C. Cruaud, S. Lacoste, J. Ropars, A. Tellier, J. Dupont, « Microsatellite loci to recognize species for the cheese starter and contaminating strains associated with cheese manufacturing », *Int. J. Food Microbiol.* **137** (2010), p. 204-213.

#### IV. Synthèse de la domestication de *Penicillium camemberti*

Comme on peut le voir sur le schéma ci-dessus (Figure 4 de l'article du III) qui résume la domestication de *P. camemberti*, celle-ci s'est réalisée en deux étapes successives. Un premier évènement de domestication a mené à la différenciation de *P. biforme* (présent dans l'environnement agro-alimentaire) et *P. fuscoglaucum*. Un second évènement de domestication a mené à la différenciation de *P. camemberti* (uniquement présent sur fromage) aux alentours des années 1900. Les données historiques corroborent l'absence de diversité génétique de *P. camemberti*. Bien que la diversité de *P. camemberti* soit extrêmement faible, nous avons mis en évidence deux lignées distinctes, une lignée où les souches sont de couleur blanche et très duveteuses, *P. camemberti* var. *camemberti* (retrouvée sur bries et camemberts), et une autre lignée, *P. camemberti* var. *caseifulvum*, moins duveteuse et de couleur grise (retrouvée sur rigottes de Condrieu et St Marcellin).

La majeure partie des études phénotypiques se sont concentrées sur l'espèce *P. camemberti*. L'espèce *Penicillium biforme* est aussi présente sur fromage. Un meilleur échantillonnage de *P. biforme* ainsi qu'une caractérisation phénotypique du complexe d'espèces de *P. camemberti* (lipolyse, protéolyse et composés volatils) sont donc de bonnes pistes de poursuite de ces travaux. Le cas de *P. camemberti* est un cas particulier où les données historiques corroborent un effet fondateur que l'on retrouve au niveau génétique : *P. camemberti* n'a quasi-aucune diversité génétique. Or de nombreuses analyses de génomique quantitative reposent sur la présence de variations génétiques le long du génome. Il est donc particulièrement difficile de déterminer quels ont été les gènes qui ont subi la sélection. De plus, lors de la comparaison phénotypique de population clonale domestiquée et de

populations naturelles très diversifiées génétiquement, on se retrouve dans un cadre où assez souvent l'homoscédasticité n'est pas respectée ce qui peut compliquer l'analyse statistique.

*Geotrichum candidum* est une levure poussant sur le même type de fromage que *P. camemberti*. Cette levure est même parfois utilisée en synergie avec *P. camemberti*. Elle est donc soumise à des contraintes similaires. Cela en fait, un excellent second modèle de domestication de champignons dans l'environnement fromager afin de généraliser ou nuancer les résultats obtenus chez *P. camemberti*.



Chapitre 2 Domestication de *Geotrichum*  
*candidum*

## I. Résumé

*Geotrichum candidum* est un champignon que l'on retrouve dans des fromages tel que le St Nectaire, Reblochon, Ossau Iraty, Tilsit, brie et fromages de chèvre mais aussi dans des environnements plus naturels tel que le sol, le sable et en forêt tropicale. Certaines souches sont inoculées tandis que d'autres sont naturellement présentes dans le lait cru. *Geotrichum candidum* est un ascomycète dimorphique pouvant prendre une forme de levure ou de moisissure (comme montré ci-dessous; figure 15).

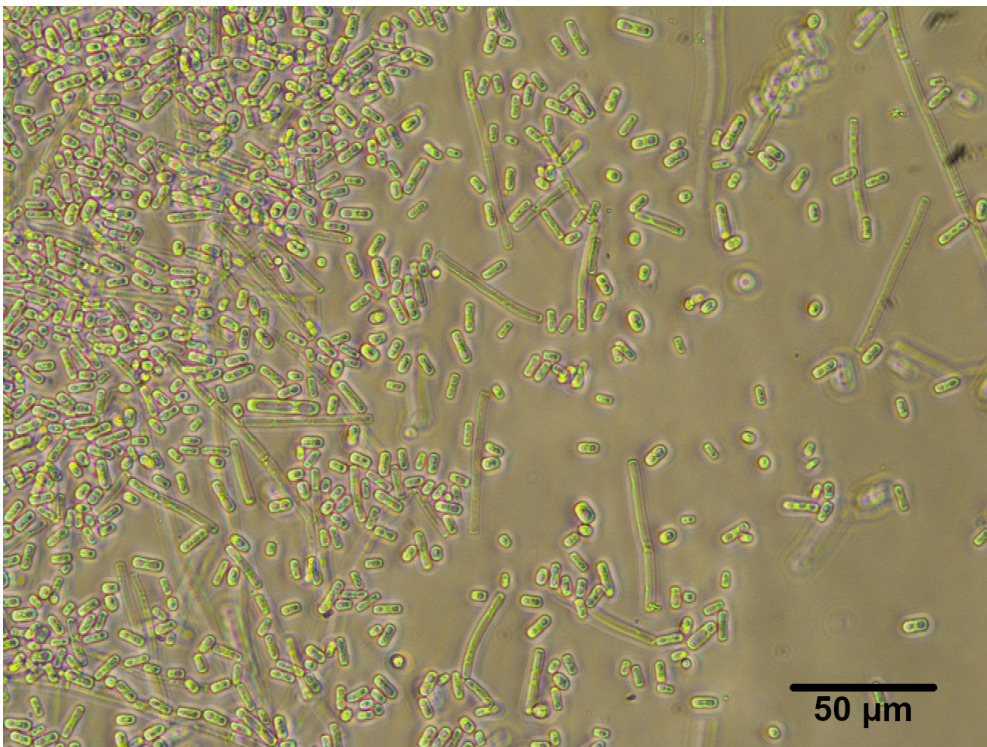
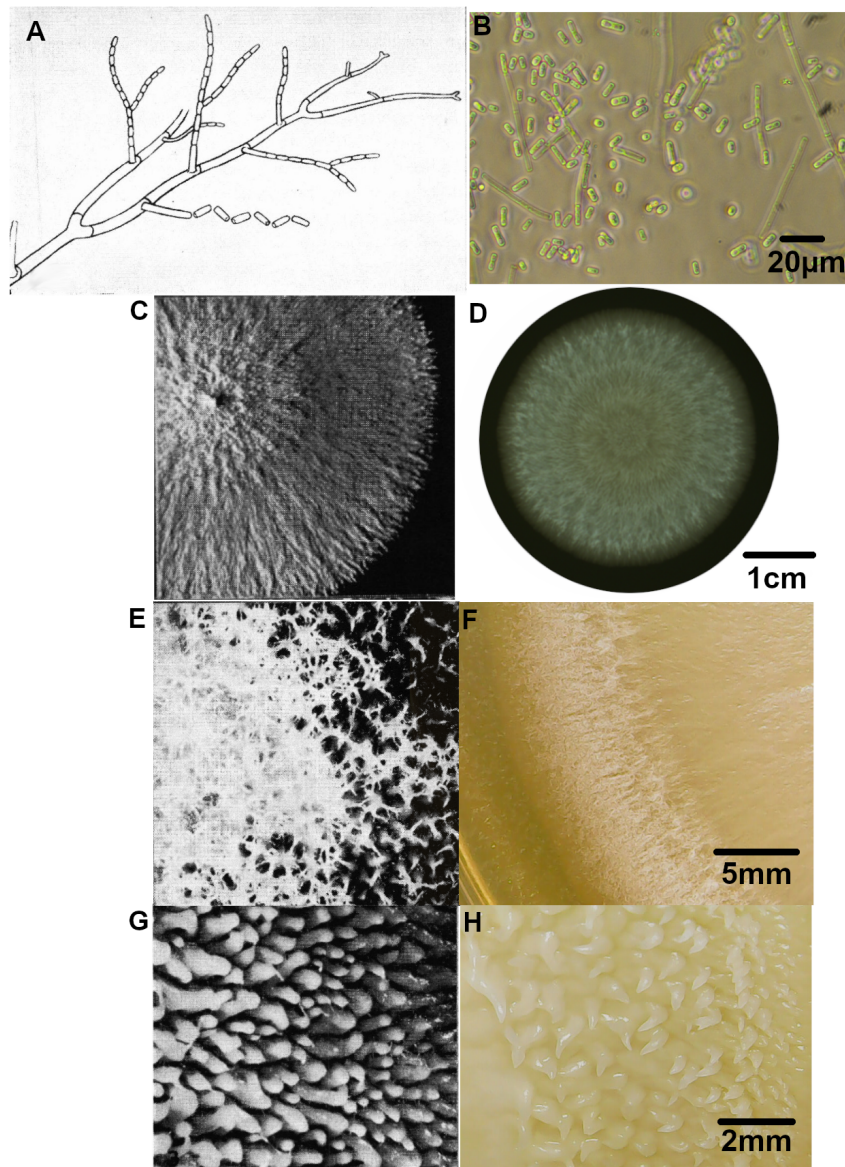


Figure 15 : *Geotrichum candidum* en microscopie optique montrant sa forme levure et sa forme mycélienne (x200). Source : Bastien Bennetot

Les premières citations scientifiques de *G. candidum* se font sous le nom d'*Oidium lactis* et remonte à 1905 dans les comptes rendus hebdomadaires des séances de l'académie des sciences (Tome 140, p 1475). On y parle de l'isolation de *G.*

*candidum* dans le camembert et de son rôle essentiel dans la production d'arômes et la désacidification du fromage. La première trace iconographique de *G. candidum* accompagnée d'une description morphologique date de 1957 (Carmichael 1957).



**Figure 16 : Illustration de Carmichael 1957 (côté gauche) et prise actuelle (2022) de *G. candidum* (côté droit).** Dessin d'observation (A) et microscopie optique (x200) (B). Photographie de l'aspect morphologique d'une colonie de *G. candidum* (C,D). Aspect de l'extérieur (E,F) et de l'intérieur (G,H) d'une colonie à la loupe binoculaire.

De nombreuses études ont porté sur la diversité phénotypique de *G. candidum* (Dieuleveux et al. 1998; Maldonado 2017; Jollivet et al. 1994; Amrane et Prigent 1997; Molimard et al. 1994; Boutrou et Gueguen 2005) sans pour autant prendre en

compte la structure de populations de *G. candidum*. Jusqu'à récemment, la structure génétique de *G. candidum* était principalement établie à partir d'étude de typage de séquence multilocus (MLST) (Tinsley et al. 2022; Jacques et al. 2017). Cela a permis de montrer l'existence de plusieurs clusters génétiques dont certains étaient spécifiques de l'environnement fromager. En 2020, un génome par cluster génétique a été séquencé afin de rendre compte de la relation phylogénétique au sein de *G. candidum* et ses espèces proches (Perkins et al. 2020). Ma thèse représente donc le premier projet de génomique à une échelle populationnelle chez *G. candidum*. Nous avons séquencé des souches de collections publiques, d'industriels, mais aussi des souches isolées au laboratoire. Afin que notre jeu de données soit le plus complet possible, toutes les souches disponibles provenant du milieu naturel (sol, eaux usées, sable) ont été utilisées et les souches de l'environnement fromager ont été choisis de manière à maximiser la diversité géographique des souches et la diversité de fromage sur lesquelles ces souches ont été isolées. Seules les souches haploïdes ont été conservées pour les analyses, d'une part car les cas de polyploïdie sont relativement rares (8 sur 106 souches séquencés) et d'autre part car très peu de choses sont connues du cycle de vie de *G. candidum*. On ne sait pas si *G. candidum* passe à un état polyploïde au cours de son cycle de vie, sous certaines conditions ou s'il s'agit d'hybrides. A partir des données génomiques, j'ai pu identifier 5 populations : une population inféodée au milieu naturel, une d'origine mixte (produits alimentaires, sol et sable) et un clade contenant la quasi-majorité des souches du fromages subdivisé en trois populations. Ces résultats sont présentés dans l'article de la partie II.

II. Article : Domestication of different varieties in the  
cheese-making fungus *Geotrichum candidum*

# Domestication of different varieties in the cheese-making fungus *Geotrichum candidum*

BASTIEN BENNETOT<sup>1</sup>, JEAN-PHILIPPE VERNADET<sup>1</sup>, VINCENT PERKINS<sup>3</sup>, SOPHIE HAUTEFEUILLE<sup>1</sup>, RICARDO C. RODRÍGUEZ DE LA VEGA<sup>1</sup>, SAMUEL O'DONNELL<sup>1</sup>, ALODIE SNIRC<sup>1</sup>, CÉCILE GRONDIN<sup>2</sup>, MARIE-HÉLÈNE LESSARD<sup>3</sup>, ANNE-CLAIRE PERON<sup>4</sup>, STEVE LABRIE<sup>3</sup>, SOPHIE LANDAUD<sup>4</sup>, TATIANA GIRAUD<sup>1</sup>, AND JEANNE ROPARS<sup>1</sup>

<sup>1</sup>Laboratoire Ecologie Systématique et Evolution, Université Paris-Saclay, CNRS, AgroParisTech, site IDEEV, Gif-sur-Yvette, France

<sup>2</sup>SPO, Univ Montpellier, INRAE, Institut Agro, Montpellier, France

<sup>3</sup>Département de sciences des aliments et de nutrition, Université Laval, Québec, Canada

<sup>4</sup>UMR SayFood, Université Paris-Saclay, INRAE, AgroParisTech, Thiverval Grignon, France

Keywords: domestication, adaptive divergence, lipolysis, proteolysis, genomics, volatiles, laboratory experiments, *Galactomyces candidus*

Domestication is an excellent model for studying adaptation processes, involving recent adaptation and diversification, convergence following adaptation to similar conditions, as well as degeneration of unused functions. *Geotrichum candidum* is a fungus used for cheese making and is also found in other environments such as soil and plants. By analyzing whole-genome data from 98 strains, we found that all strains isolated from cheese formed a monophyletic clade. Within the cheese clade, we identified three genetically differentiated populations and we detected footprints of recombination and admixture. The genetic diversity in the cheese clade was similar as that in the wild clade, suggesting the lack of strong bottlenecks. Commercial starter strains were scattered across the cheese clade, thus not constituting a single clonal lineage. The cheese populations were phenotypically differentiated from other populations, with a slower growth on all media, even cheese, a prominent production of typical cheese volatiles and a lower proteolytic activity. One of the cheese clusters encompassed all soft goat cheese strains, suggesting an effect of cheese-making practices on differentiation. Another of the cheese populations seemed to represent a more advanced stage of domestication, with stronger phenotypic differentiation from the wild clade, harboring much lower genetic diversity, and phenotypes more typical of cheese fungi, with denser and fluffier colonies and a greater ability of excluding cheese spoiler fungi. Cheese populations lacked two beta lactamase-like genes present in the wild clade, involved in xenobiotic clearance, and displayed higher contents of transposable elements, likely due to relaxed selection. Our findings suggest the existence of genuine domestication in *G. candidum*, which led to diversification into different varieties with contrasted phenotypes. Some of the traits acquired by cheese strains indicate convergence with other, distantly related fungi used for cheese maturation.

<https://doi.org/10.1101/2022.05.17.492043>

## 1. INTRODUCTION

Understanding how populations adapt to their environment is a key question in evolutionary biology. Domestication, the change in the genetic and phenotypic make-up of populations under human artificial selection, is an excellent model for studying adaptation processes, as it in-

volves recent adaptation events under strong selection on known traits, rapid diversification and reduced gene flow between wild and domesticated populations. Numerous studies have documented the specific traits acquired in domesticated animals (dog, horse, pig, cow) (Diamond, 2002; Frantz et al., 2015; Petersen et al., 2013; Warmuth

et al., 2011) and plants (cabbage, corn, wheat) (Hufford et al., 2012; Mabry et al., 2021; Peng et al., 2011), as well as their genetic differentiation from wild populations and their adaptive genomic changes. For example, domesticated animals (dog, horse and cattle) have been selected for coat color, size, rapidity and docility (Liu et al., 2022; Plassais et al., 2022; Qanbari et al., 2014). In plants too, similar traits have been selected in different lineages, such as bigger grains with more nutrients and lack of dormancy (Cornille et al., 2014; Hufford et al., 2012; Peng et al., 2011; Purugganan, 2019). On the other hand, functions essential in wild environments but unused in anthropic environments have often degenerated due to relaxed selection, for example, reductions in defense mechanisms in plants (Cornille et al., 2014; Hufford et al., 2012). Domestication also often leads to strong reduction in genetic diversity due to bottlenecks in animals (e.g. dog) (Marsden et al., 2016) and annual plants (e.g. rice) (Zhu et al., 2007).

Humans have domesticated several fungi for the fermentation of foods (e.g. beer, bread, wine, coffee, dry-cured meat and cheese), to produce secondary metabolites used in pharmaceuticals (e.g. penicillin), or for their nutritional and gustatory values (e.g. button and shiitake mushrooms) (Steensels et al., 2021). Fungi are excellent models for studying evolution and adaptation in eukaryotes, given their many experimental assets (Gladieux et al., 2014): fungi have relatively small genomes, many are easy to culture in laboratory conditions, and spores can survive long periods in the freezer. However, despite their economic and industrial importance, and their utility as biological models for studying adaptive divergence, the domestication in fungi has yet been less studied than in plants or animals. An exception is the budding yeast *Saccharomyces cerevisiae* used in the production of beer, wine, sake, bread, coffee and cacao (Bai et al., 2022; Barbosa et al., 2018; Bigey et al., 2021a; Duan et al., 2018; Fay and Benavides, 2005; Gallone et al., 2016; Gonçalves et al., 2016; Lahue et al., 2020; Legras et al., 2018; Libkind et al., 2011; Liti et al., 2009; Ludlow et al., 2016; Peter et al., 2018), and to a lesser extent the filamentous fungus *Aspergillus oryzae* used to ferment soy and rice products in Asia (Galagan et al., 2005; Gibbons et al., 2012; Machida et al., 2005) and the *Penicillium* species used for cheese ripening, e.g. *P. camemberti* for soft cheeses (Ropars et al., 2020), and *P. roqueforti* for blue cheeses (Cheeseman et al., 2014; Dumas et al., 2020; Ropars et al., 2015). Phenotypic traits beneficial for food production have been acquired in fungal domesticated populations, being different from wild populations. The domestication having led to *P. camemberti* occurred in several steps, with the successive differentiation of several lineages displaying decreasing diversity and increasing beneficial traits for cheese maturation, from the wild *P. fuscoglaucum*, to *P. bifforme* and then the two clonal *P. camemberti* varieties, *caseifulvum* and *camemberti* (Ropars et al., 2020b). Domesticated populations of fermented food microorganisms can for example better assimilate the carbon sources present

in the anthropic environment, e.g. lactose for *Penicillium* cheese fungi (Ropars et al., 2015) and maltose for *S. cerevisiae* sourdough strains (Bigey et al., 2021b). Furthermore, volatile organic compounds crucial for cheese flavor are more appealing in cheese populations compared to wild populations in *P. roqueforti* (Caron et al., 2021).

The genomic processes involved in adaptation to human-made environments in domesticated fungi include gene family expansion for specific metabolism pathways, gene gain, inter-specific hybridization, introgression and horizontal gene transfer (Almeida et al., 2014; Barros Lopes et al., 2002; Borneman et al., 2016; Cheeseman et al., 2014; Gallone et al., 2016; Libkind et al., 2011; Machida et al., 2005; Naumova et al., 2005; Novo et al., 2009; Ropars et al., 2015). Domesticated fungi also have lost parts of genes no longer useful in the food environment; for example a cheese *P. roqueforti* population and *P. camemberti* var. *caseifulvum* are no longer able to produce some of their toxins due to deletions in the corresponding genes (Gillot et al., 2017; Ropars et al., 2020b). Such losses are likely due to relaxed selection in terms of competition ability in cheese, in which desired fungi are often inoculated in large quantities compared to possible competitors. Bottlenecks (leading to genetic drift) and degeneration have also been documented in domesticated fungi, with reduced fertility and genetic diversity in the cheese fungi *P. roqueforti* and *P. camemberti* (Dumas et al., 2020; Ropars et al., 2020b), likely due to an accumulation of deleterious mutations because of drift.

While several cheese-making fungi have been studied recently, it is important to add study cases in additional lineages, as it allows addressing the question of whether adaptation to a similar medium leads to convergent traits. In the case of cheese-making fungi, one can expect phenotypic convergence, for example for more or less rapid growth on cheese, higher proteolysis and lipolysis abilities, higher competitive abilities and greater production of positive volatile compounds (Ropars et al., 2020a, 2020b), as these traits may have been selected similarly in the different fungi used for cheese making. *Geotrichum candidum* (teleomorph *Galactomyces candidus*) is a dimorphic fungus (i.e., able to grow as a yeast or a mycelial form), commonly used for cheese-making, but also thriving in other environments such as soil, plants and fruits. *Geotrichum candidum* is naturally present in raw milk and is also often added as a starter culture for the production of semi-hard, mold-ripened, smeared soft cheeses, fresh goat and ewe cheeses. Analyses based on genetic markers have revealed genetic differentiation between cheese and wild strains (Alper et al., 2013; Jacques et al., 2017; Perkins et al., 2020; Tinsley et al., 2022). Phenotypic diversity within *G. candidum* has been reported in terms of carbon and nitrogen assimilation, lipolysis and proteolysis (Boutroun and Gueguen, 2005; Perkins et al., 2020). However, it has not been tested whether *G. candidum* cheese populations have evolved specific traits that could be beneficial for cheese making.

127 By analyzing the genomes of 98 strains isolated from  
128 different kinds of cheeses and other environments, we  
129 confirmed the genetic differentiation between cheese and  
130 wild strains and the occurrence of recombination. Within  
131 the cheese clade, we reveal the existence of three varieties,  
132 i.e., three genetic clusters with contrasted traits and levels  
133 of diversity. One of the cheese clusters encompassed all  
134 goat soft cheese strains, suggesting an effect of cheese-  
135 making practices on differentiation. Commercial strains  
136 did not belong to a single clonal lineage, but were in-  
137 stead present in all cheese clusters, some corresponding to  
138 admixed strains. We found phenotypic differentiation be-  
139 tween cheese and wild populations, and between cheese  
140 populations, in terms of growth, proteolysis and volatile  
141 compounds. The cheese clade lacked two tandem beta  
142 lactamase-like genes, present in wild strains, and involved  
143 in xenobiotic clearance, and contained more repetitive  
144 elements than the wild clade. Altogether our findings  
145 suggest the existence of genuine domestication in *G. can-*  
146 *didum*, with both genetic and phenotypic differentiation of  
147 cheese strains from their wild counterparts, and a stronger  
148 domestication syndrome in one of the cheese clusters. We  
149 also show convergence in some traits with domesticated  
150 *Penicillium* cheese fungi, such as a denser and fluffier  
151 mycelium, at the expense of radial growth, and the pro-  
152 duction of typical cheese volatiles.

## 153 2. RESULTS

### 154 A. Genetic differentiation between wild and cheese 155 strains in *Geotrichum candidum*

156 We collected and sequenced the genomes of 88 *G. can-*  
157 *didum* strains with Illumina technology and included in  
158 our analyses ten available genomes (Illumina and PacBio)  
159 (Perkins et al., 2020). Our dataset included 61 strains  
160 isolated from different kinds of cheeses (semi-hard, mold-  
161 ripened, smeared soft and fresh soft goat cheeses), 16  
162 industrial strains used for cheese-making, seven strains  
163 from dairy products, four strains from other food sub-  
164 strates (e.g., sausage or vegetables) and all the 10 wild  
165 strains available in public collections worldwide (isolated  
166 for example from soil or plant; Table S1). We identified  
167 699,755 SNPs across the 98 strains by mapping against the  
168 CLIB 918 reference genome (cheese strain, NCBI accession:  
169 PRJEB5752). All the 98 *G. candidum* strains analysed were  
170 haploid.

171 The maximum likelihood tree, principal component  
172 analysis (PCA) and neighbor-net (SplitsTree) analyses all  
173 identified the same three clades (Figure 1A), with one con-  
174 taining mostly wild strains (corresponding to the GeoC  
175 group identified previously based on genetic markers)  
176 (Perkins et al., 2020), one composed of strains of varying  
177 origins (i.e. dairy products and other environments, corre-  
178 sponding to the group previously named GeoB) and one  
179 containing mostly cheese and dairy strains (previously  
180 named GeoA). The larger sampling and the genome se-  
181 quencing of the present study further revealed genetic

182 subdivision in the cheese clade, with three clearly differ-  
183 entiated populations and several admixed strains (Figure  
184 1B).

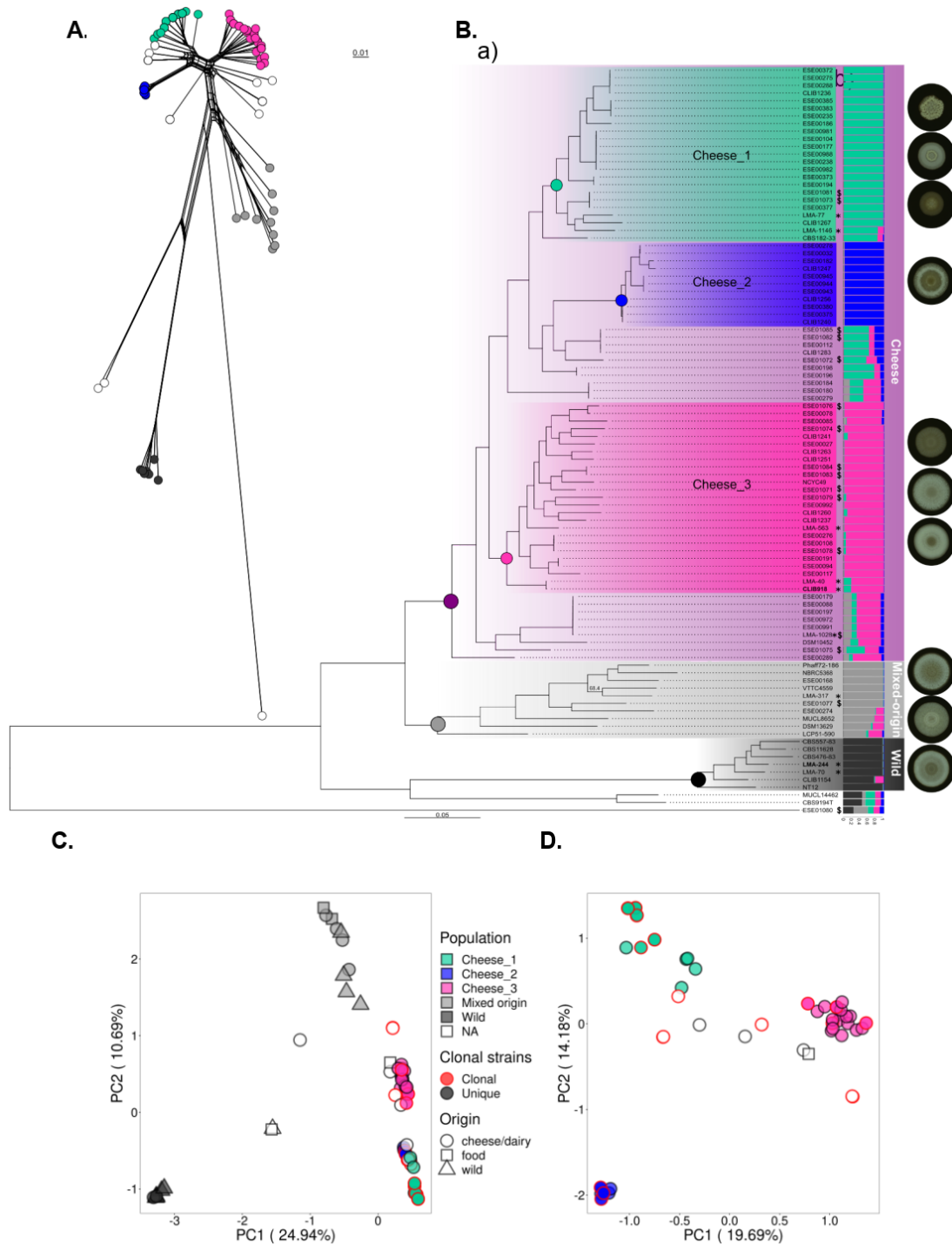
185 We performed an admixture analysis, assigning strains  
186 to K populations, with K ranging from two to ten. At  
187 K=3, the cheese, mixed-origin and wild clades were sep-  
188 arated (Figure S1). At K=5, the cheese clade was di-  
189 vided into three genetic clusters, corresponding to mono-  
190 phyletic groups in the maximum likelihood tree and well-  
191 separated genetic clusters in the PCA and the neighbor-  
192 net (Figure 1 B, D). The second order rate of change of  
193 the likelihood ( $\Delta K$ ) peaked at K=6. The additional  
194 population distinguished at K=6 compared to K=5 how-  
195 ever encompassed only two strains, that were not much  
196 differentiated from others in the splitsTree (MUCL 14462  
197 and CBS 9194, isolated from non-cheese substrates and  
198 clustering with the wild population at K=5; FIGURE 1B).  
199 We therefore chose to consider only the five largest popu-  
200 lations in the following, as we could not run phenotypic  
201 tests on a population with only two strains.

202 Some cheese strains could not be assigned to any ge-  
203 netic cluster with the admixture analysis and were placed  
204 on the PCA between the three well-delimited cheese ge-  
205 netic clusters (Figure 1C-D), suggesting that they resulted  
206 from admixture events. To test this hypothesis, we in-  
207 vestigated whether these strains had mosaic genomes,  
208 with different genomic regions assigned to distinct clus-  
209 ters. We calculated pairwise identity between unassigned  
210 strains and the other strains, computing mean identity to  
211 the different genetic clusters along scaffolds using sliding  
212 windows. For all unassigned strains in the cheese clade,  
213 we observed shifts in identity values along scaffolds, con-  
214 firming that these strains are the results of admixture  
215 between clusters (Figure S2). In contrast, the three unas-  
216 signed strains outside of the cheese clade did not show  
217 changes in similarity level to the different clusters along  
218 their genome; these strains may belong to yet additional  
219 genetic clusters that could not be distinguished by the  
220 analyses because too few strains belonged to these clus-  
221 ters in the sampling (Figure 1B).

222 We tested for an over-representation in the three cheese  
223 clusters of the five types of cheeses from which strains  
224 were sampled, i.e., soft, soft natural rind, pressed un-  
225 cooked, pressed cooked and blue cheeses (Table S1). We  
226 found an over-representation only in the cheese\_1 popu-  
227 lation ( $\chi^2 = 36.6$  ;  $df = 12$  ;  $p < 0.001$ ), that encompassed  
228 all strains isolated from soft natural rind goat cheeses,  
229 these kinds of cheeses having a characteristic convoluted  
230 aspect.

231 The wild clade was the most differentiated population  
232 from all other *G. candidum* populations with  $F_{ST}$  values  
233 above 0.70 and  $d_{xy}$  of 1.04E-02 (Table S2). The percentage  
234 of private SNPs in the five populations was also high  
235 (Table S3).  $F_3$  tests based on the number of shared sites  
236 (Table S4) supported the genetic differentiation between  
237 these populations and the lack of gene flow or migration.





**Fig. 1. Phylogenetic relationships and population structure of 98 strains of *Geotrichum candidum*, based on whole-genome data** (A) Neighbor-net analysis based on a single nucleotide polymorphism (SNP) distance matrix. The scale bar represents 0.01 substitutions per site for branch lengths.

(B) Genetic relationships among strains and population structure in *G. candidum* based on 699,755 SNPs. a) Maximum likelihood tree showing phylogenetic relationships among the 98 isolates used in this study. All nodes are supported by bootstrap support >98% (bootstrap analysis with 1000 resampled datasets). The scale bar represents 0.05 substitutions per site for branch lengths. We used the midpoint rooting method to root the tree. The “\$” symbol pinpoints commercial starter strains and “\*” the PacBio sequences. Genomes used as references are written in bold. b) Population subdivision inferred for K = 5. Colored bars represent the coefficients of membership in the five gene pools based on SNP data.

(C) Principal component analysis (PCA) based on 699,755 SNPs and 98 strains. Genetic clusters are represented by the same colors on all panels: light blue for Cheese\_1, dark blue for Cheese\_2, pink for Cheese\_3, light grey for the mixed-origin population and dark grey for the wild population. The symbols represent the environment from which strains were sampled: circle for cheese/dairy, square for food and triangle for wild environment. Symbols are circled in red when multiple strains are overlapping due to clonal lineages (with a threshold set at <1,200 SNP for defining clonal strains).

(D) PCA based on the 323,385 SNPs when the dataset was restricted to the 78 strains from the cheese clade.

## B. High nucleotide diversity within cheese populations and footprints of recombination

The overall diversity in the cheese clade ( $\pi = 2.82E-3$ ) was higher than in the wild population ( $\pi = 2.12E-03$ ). Each of the three cheese populations of *G. candidum* had however reduced nucleotide diversities compared to wild and mixed-origin populations (Table S5), by at least a factor of two. The Cheese\_2 population showed the lowest genetic diversity ( $\pi = 4.82E-04$ ), by a factor of four compared to the two other cheese populations (Table S5).

*Geotrichum candidum* is a heterothallic fungus, meaning that sexual reproduction can only occur between two haploid cells carrying different mating types. Two mating types have been described in *G. candidum* (Morel et al., 2015): MATA, encoded by a HMG box gene homolog to the MATA2 *Kluyveromyces lactis* allele, present in CLIB 918 (sequence id: HF558448.1), and MATB, encoded by an alphabox gene homolog to the MAT $\alpha$ 1 *S. cerevisiae* allele, present in the strain CBS 615.84 (sequence id: HF558449.1). In the Cheese\_2 population, we found a significant departure from the 1:1 mating-type ratio expected under regular sexual reproduction; all the 12 strains carried the MATB allele, suggesting that this population is at least partly clonal (Table S6). The absence of linkage disequilibrium decay with physical distance between two SNPs (Figure S3), together with the absence of reticulation in the neighbor-net (Figure 1), are also consistent with a lack of recombination in the Cheese\_2 population. However, pairwise homology index (PHI) tests, testing with permutations the null hypothesis of no recombination by looking at the genealogical association among adjacent sites, were significant in all the *G. candidum* populations (Table S7); this indicates that recombination did occur at least in a recent past, even in the Cheese\_2 population.

We did not detect any accumulation of nonsense or missense mutations in any population compared to silent mutations (Table S8), while degeneration can be expected to be particularly strong in clonally replicated populations as recombination allows more efficient selection. None of the genes that presented nonsense mutation had predicted functions that could be detected as specific either to the wild or cheese environments (Figure S4). This suggests that the absence of sexual reproduction in the Cheese\_2 population may be too recent to observe an accumulation of deleterious mutations.

In contrast to the Cheese\_2 population, we found both mating-type alleles in balanced proportions in both Cheese\_1 and Cheese\_3 populations (Table S6) and we observed sharp decays in linkage disequilibrium (LD) with genomic distance, although LD levels remained higher than in the mixed-origin and wild populations (Figure S3). We observed reticulations in the neighbor-net network within populations and, to a lesser extent, between populations (Figure 1A).

As previously mentioned, the 16 commercial starter strains in our *G. candidum* dataset were scattered in the maximum likelihood tree (“\$” symbol, Figure 1B.a.) and

we detected above footprints of recombination in the Cheese\_1 and Cheese\_3 populations (Figure 1A, Figure S3). We nevertheless detected a few groups of clonemates, by the lack of branches in the trees and the presence of fewer than 1,200 SNPs between strains (Figure 1; Table S1, clonal group column). As strains within these clonal lineages were isolated from different cheeses and from various French regions, it indicates that these lineages are likely clonally cultivated and sold for cheese making. Some of the commercial starter strains were, in fact, placed within these clonal groups (“\$” symbol on Figure 1B). Among the admixed cheese strains, 19 out of 23 were part of clonal lineages.

## C. Copy number variation: lack of two tandem beta lactamase-like genes in the cheese populations and higher repeat content

Expansions of gene families involved in specific metabolism pathways, of transposable elements and loss of genes no longer required in the new environment can be involved in adaptation to new environments. For example, variations in gene copy number were associated with the adaptation of *S. cerevisiae* to beer making, with duplications of genes involved in maltose and maltotriose metabolism specifically in domesticated beer strains (Gallone et al., 2016; Giannakou et al., 2020; Gonçalves et al., 2016). We therefore looked for gene copy-number variation (CNV) that differentiated wild and cheese populations, using 500 bp sliding windows and two reference genomes, belonging to the Cheese\_3 and the wild populations, respectively (Table S9). Using the Cheese\_3 reference, we found 61 CNV regions (mean length of 1515 bp and 45 non-genic CNVs), encompassing in total 16 genes, half having predicted functions, none being obviously related to cheese adaptation (e.g. methylglyoxal reductase and tRNAs, Table S9). Using the wild genome reference, we found 132 CNV regions (mean length of 1664 bp and 105 non genic CNVs), encompassing 29 genes (seven with unknown functions).

One of these regions, 20 kb long, included only two genes, both matching the Pfam hidden markov model for beta-lactamases; these two genes (g5112 and g5113) were absent from all cheese populations, and were present in most wild strains (except one that had partially the region) and in four strains belonging to the mixed-origin population (Figure 2). The nucleotide identity between the two beta lactamase-like genes was 93%. A third beta lactamase-like gene (g5111) was found immediately next to this CNV region in all *G. candidum* strains, and displayed a nucleotide identity of 87% with the two other beta lactamase-like genes within the CNV. Surrounding these different genes, we found several Tc1/mariner, a LINE/Tad1 and other DNA transposons, that may have contributed to the beta-lactamase-like gene deletion (Figure 2). Fungal beta lactamase-like genes are known to contribute to hydrolysis of microbial and plant xenobiotics, and thus may be important in the wild environment

351 to compete with other microorganisms (Gao et al., 2017).  
352 If the ancestral state in *G. candidum* is the presence of three  
353 beta-lactamase genes, the cheese populations may have  
354 lost these two copies of the beta-lactamase genes due to  
355 relaxed selection; indeed, these functions may not be use-  
356 ful in the cheese environment if *G. candidum* is inoculated  
357 in high quantity compared to potential competitors.

358 De novo detection of repeats using the wild strain LMA-  
359 244 yielded a library containing 107 types of repeated el-  
360 ements (including 15 types of DNA transposons and 11  
361 of retroelements and 3 rolling-circles). We identified 14  
362 types of repeated elements present in at least one other  
363 *G. candidum* genome with five times more copies than in  
364 the LMA-244 wild strain (this threshold was set based on  
365 the fat tail of the distribution; Figure S5). Among these 14  
366 types of repeated elements, several DNA transposons of  
367 the Tc1/mariner repeat family showed a cheese-clade spe-  
368 cific expansion (Table S10, Figure 3). Several unidentified,  
369 Tad1 and Helitron repeat types also showed expansions  
370 in the cheese clade, alongside a milder expansion in the  
371 mixed origin clade. Such higher contents of transposable  
372 elements in the cheese clade could result from expansions  
373 due to relaxed selection in the cheese clade (Baduel et al.,  
374 2019).

#### 375 **D. Genomic footprints of adaptation: genomic islands** 376 **of differentiation and genes under positive selection**

377 We looked for genomic regions with a greater differentia-  
378 tion or a lower genetic diversity than the genomic back-  
379 ground when comparing each of the three cheese pop-  
380 ulations to the wild population, to detect footprints of  
381 divergent selection and recent selective sweeps, respec-  
382 tively. We scanned the whole genome in each cheese  
383 cluster using non-overlapping windows and identified  
384 the windows within both the 1% highest differentiation  
385 with the wild population (dxy) and the 5% lowest within-  
386 population diversity ( $\pi$ ). We also performed an analysis  
387 using SweeD on each genetic cluster to detect selective  
388 sweeps based on the site frequency allelic spectrum, keep-  
389 ing the windows with the 1% highest likelihood. The  
390 putative functions of outlier genes is given per cluster and  
391 analysis in the Table S11. . In the Cheese\_1 population,  
392 four genes were outliers in all three selection analyses  
393 (1% highest likelihood sweeD, 1% highest differentiation  
394 and 5% lowest diversity), with one coding for a  $\text{Ca}^{2+}$ -  
395 dependent cystein protease (Table S11). This function  
396 was also represented in outliers in the other cheese pop-  
397 ulations, but only shared between the sweeD and dXY  
398 analyses. Proteases are important in cheese making as  
399 the breakdown of milk caseins greatly contributes to cheese  
400 texture and decreases water activity by degrading pro-  
401 teins into molecules with free carboxyl and amino groups  
402 (McSweeney, 2004). *Geotrichum candidum* is prevalent dur-  
403 ing the amino-acid catabolism ripening step of Pelardon  
404 fresh cheese (Penland et al., 2021), suggesting that *G. can-*  
405 *didum* plays an important role in proteolysis in cheese-  
406 making. Among the genes shared between the dXY and

407 sweeD analyses, we also detected functions related to iron  
408 uptake, in particular the homolog of the Schizosaccha-  
409 romyces pombe Ctr4 in Cheese\_1 and a high-affinity iron  
410 permease in Cheese\_2. Ctr4 is a copper-sensing transcrip-  
411 tion factor regulating iron uptake genes in yeasts (Labbé  
412 1999). Iron is limited in cheese and it may therefore be  
413 advantageous for cheese strains to better regulate iron  
414 uptake (Monnet 2012).

415 We also searched for genes evolving under positive  
416 selection in terms of high rates of non-synonymous sub-  
417 stitutions by performing McDonald and Kreitman (MK)  
418 tests (Table S12), comparing the mixed-origin population  
419 to each cheese population and to the cheese clade as a  
420 whole. We detected 25 genes as evolving under positive  
421 selection in at least one cheese population (9 for Cheese\_1,  
422 18 for Cheese\_2, two in Cheese\_3 and one in all three  
423 cheese populations at once; Table S12). Among them, a  
424 metalloendopeptidase evolved under positive selection in  
425 all three cheese populations, likely playing a role in casein  
426 degradation through cell lysis (Dugat-Bony et al., 2015;  
427 Kumura et al., 2002). A Glucan 1,3-beta-glucosidase was  
428 also detected as evolving under positive selection in the  
429 Cheese\_2 population; this enzyme could be involved in  
430 fungal inhibition through fungal cell degradation (Adams,  
431 2004). The other genes under positive selection had either  
432 no predicted function or putative functions that could not  
433 be related to cheese adaptation (Table S12).

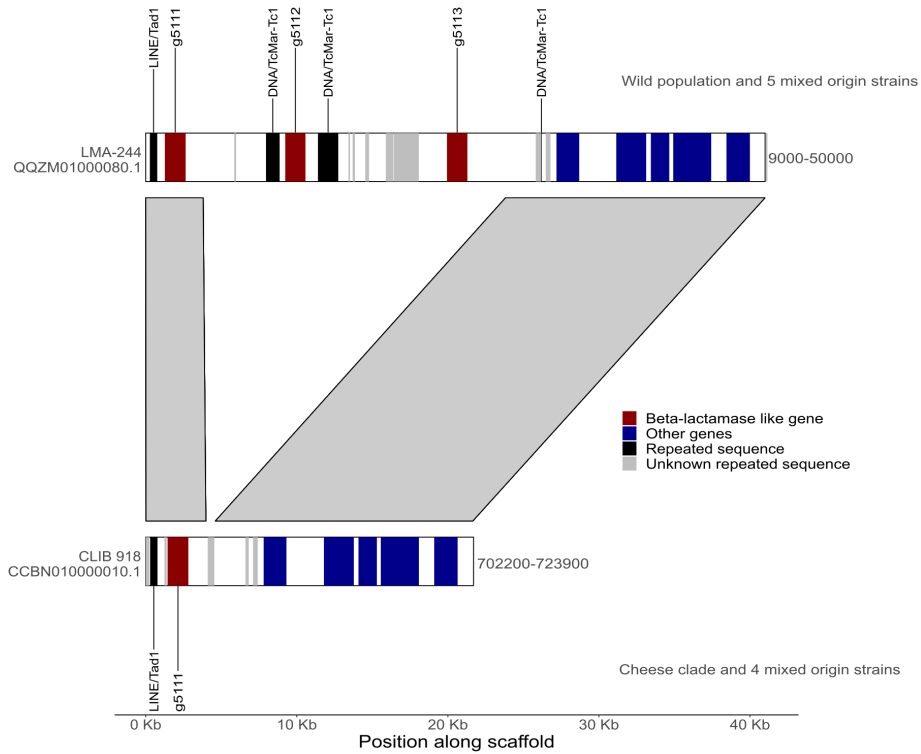
#### 434 **E. Phenotypic differentiation between cheese and wild** 435 **populations**

##### 436 **E.1. Denser mycelial growth and/or faster proteolysis in cheese** 437 **populations of *Geotrichum candidum***

438 Cheese fungi selected by humans are expected to display  
439 specific traits beneficial for cheese making, such as faster  
440 growth in cheese at cave temperature or colonies of at-  
441 tractive aspect or color. In contrast, the ability to grow in  
442 harsh conditions, e.g., with little nutrients, may have been  
443 lost in cheese strains due to relaxed selection, as often re-  
444 ported for unused traits in human-made environments in  
445 domesticated organisms (Gallone et al., 2018; Price, 2002;  
446 Ropars et al., 2015).

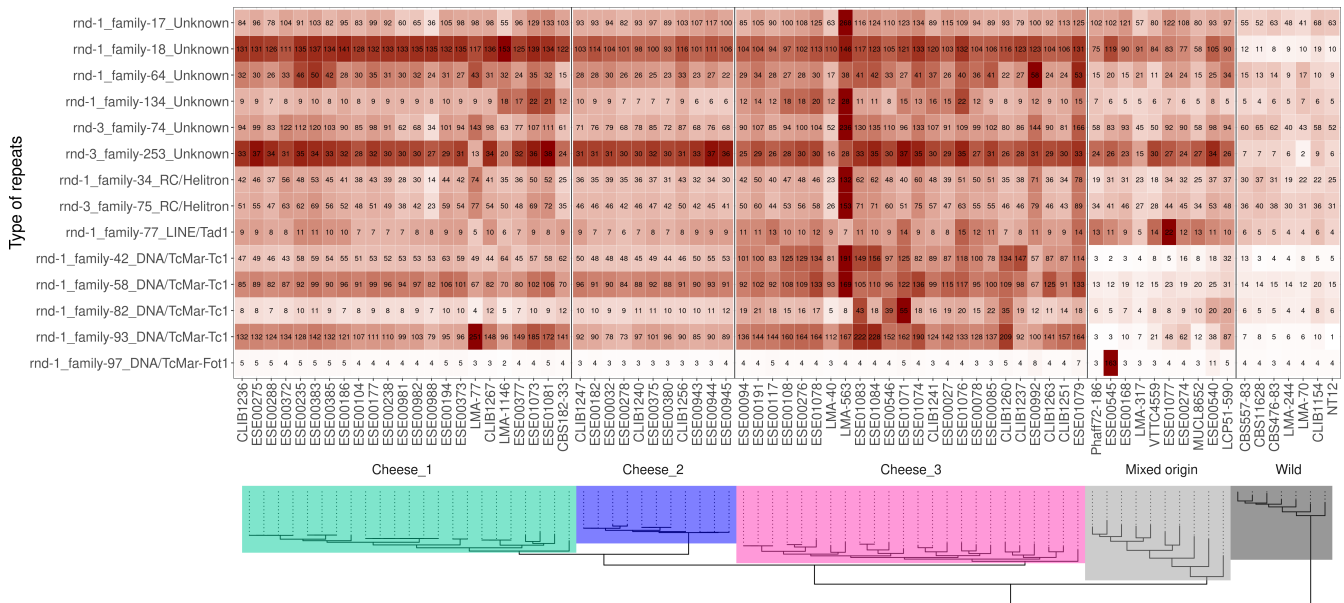
447 We therefore measured colony radial growth of 31  
448 strains from the five *G. candidum* populations on different  
449 agar media (cheese, rich and poor media) at different tem-  
450 peratures. The wild population grew faster than cheese  
451 populations on all media (on a poor medium contain-  
452 ing only inorganic salts and a low-concentration carbon  
453 source, but also on cheese media) and at all temperatures,  
454 with a more pronounced difference at 25°C (Table S13, Fig-  
455 ure S8). This may result from trade-offs with other traits,  
456 such as a fluffier mycelium, i.e. more vertical growth at  
457 the expense of less radial growth.

458 To test whether cheese populations had a denser  
459 mycelium or had become whiter and/or fluffier, we com-  
460 pared the opacity of populations on cheese agar at cave  
461 temperature (10°C), which integrates the brightness and  
462 fluffiness of a colony. The Cheese\_1 and Cheese\_3 popula-



**Fig. 2. Lack of the beta lactamase-like genes in the cheese clade of *Geotrichum candidum***

Synteny between parts of the scaffold QQZM01000080.1 of the LMA-244 (wild) strain against the scaffold CCBN010000010.1 of the cheese CLIB 918 (Cheese\_3) strain. Beta-lactamase-like genes are annotated in red while other genes are displayed in blue. Black triangles indicate positions with repeated sequences. All strains from the cheese clade and five strains from the mixed-origin populations (LMA-317, ESE00274, MUCL8652 and ESE00540) lacked the g5112 and g5113 genes, both encoding for beta-lactamase like.



**Fig. 3. Heatmap of repeats in the cheese populations**

The total number of copies is indicated in the center of each cell and by the grey to red color scale. The maximum likelihood (ML) tree from the figure 1 (without admixed strains) is plotted below strain names.

463 tions were not more opaque than wild populations (Figure  
464 4B). The Cheese\_2 population had a significantly higher  
465 opacity than all other *G. candidum* populations, except  
466 the mixed-origin population (post-hoc Tukey test in Table  
467 S13). Lipolysis and proteolysis are crucial biochemical  
468 processes during cheese ripening, as they influence flavor  
469 and texture of the final product. However, too fast proteolysis  
470 or lipolysis can lead to degraded products. The wild  
471 and mixed-origin populations degraded a significantly  
472 higher amount of proteins than the cheese populations  
473 while we did not detect any proteolysis in the Cheese\_2  
474 population in our experiment (Figure S9; Table S14). All  
475 populations had similar lipolysis rates.

### 476 **E.2. No adaptation to high salt concentration or milk origin in** 477 **cheese populations**

478 Cheese is a salty medium, with the percentage of salt  
479 varying from 0.5 g / 100 g for Emmentaler to 3 g / 100 g  
480 for Roquefort. Salt is added on the surface of cheeses to  
481 prevent the growth of contaminants, and cheese popula-  
482 tions of *G. candidum* may thus have adapted to high salt  
483 concentrations. Cheeses display a wide range of salt con-  
484 centrations so we tested four cheese media: unsalted, 1%  
485 salt as St Nectaire and cream cheeses, 2% as Camembert  
486 and goat cheeses and 4% as Roquefort blue cheeses. Wild  
487 populations grew faster than cheese populations in all salt  
488 concentrations tested, as on YPD and minimal media (Fig-  
489 ure S10A ;Table S13). Because all strains sampled from  
490 goat cheeses belonged to the Cheese\_1 population, we  
491 tested whether this population was able to grow faster  
492 on goat cheese medium (1% salt) compared to other pop-  
493 ulations. We however found no significant interaction  
494 between population and media on radial growth effects,  
495 i.e. no specific adaptation to any particular kind of milk  
496 by the different populations (Figure S10B).

### 497 **E.3. Contrasting volatile compound production between wild** 498 **and cheese populations**

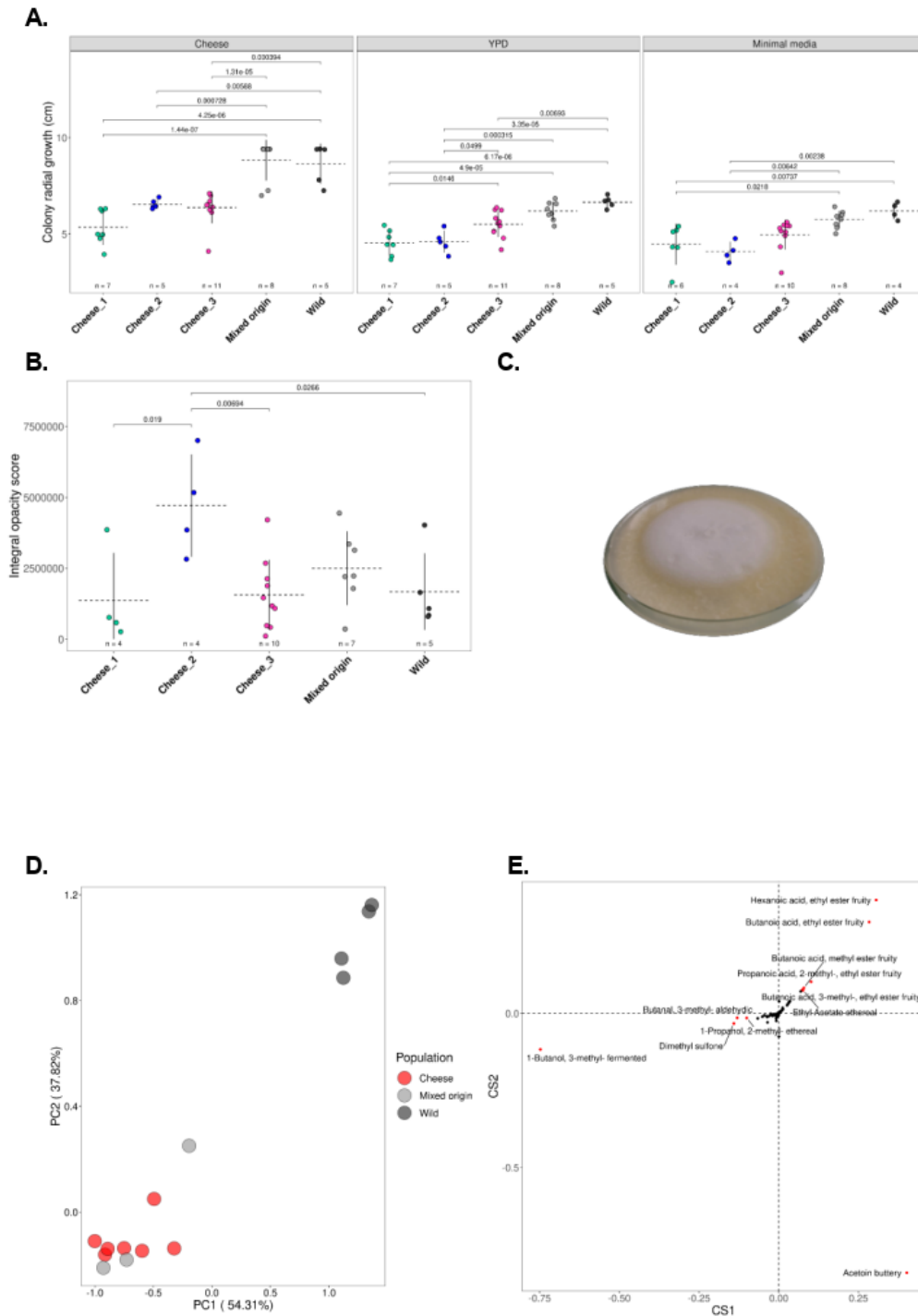
499 Cheese ripening fungi, including *G. candidum*, contribute  
500 to cheese flavor through the production of volatile com-  
501 pounds (McSweeney, 2004). Flavor is a crucial criterion  
502 for cheese consumers and the cheese populations may  
503 have been selected for desirable and specific volatile com-  
504 pounds. We grew 14 *G. candidum* strains on a sterilized  
505 Camembert curd for 21 days at 10°C, i.e., the ripening con-  
506 ditions of a Camembert. On average across compounds,  
507 the wild population produced five times higher quan-  
508 tities of volatiles than cheese populations. In order to  
509 compare the relative proportions of the different com-  
510 pounds, which is also an important aspect for flavor, we  
511 standardized the values by dividing all compound quan-  
512 tities by the total quantity of volatiles per sample. The  
513 PCA indicated a differentiation between wild and cheese  
514 strains in terms of volatile relative proportions (Figure  
515 4D). The wild population thus produced combinations  
516 of volatile compounds different from cheese populations,  
517 with a high proportion of ethyl esters and ethyl acetates  
518 (Figure S10C). However, the impact of ethyl acetate on fla-

519 vor is rather negative because it brings solvent type notes.  
520 In cheese, these esters are never predominant (Liu et al.,  
521 2004; Urbach, 1997). Ethyl esters are involved in anaero-  
522 bic metabolism and may be important for survival in the  
523 wild. By contrast, cheese strains produced many alcohols,  
524 ketones, aldehydes and sulfur compounds (Figure S10C),  
525 known for producing flavors typical of cheeses such as  
526 buttery, cheesy, fermented and aldehydic notes (Curioni  
527 and Bosset, 2002). These cheese-associated volatile com-  
528 pounds were present in similar absolute quantities in wild  
529 strains but were in minor proportions compared to other  
530 volatile compounds (Table S13), suggesting that cheese  
531 populations evolved a lower production of undesirable  
532 and unused volatiles. The overall balance between dif-  
533 ferent volatile compounds is as important as volatile ab-  
534 solute quantities for flavor perception (Liu et al., 2004).  
535 The dimethyl sulfone, a compound previously reported to  
536 be produced during the catabolism of L-methionine in *G.*  
537 *candidum*, is actually specifically produced by the cheese  
538 populations (Bonnarme et al., 2001; Penland et al., 2021),  
539 with no difference between the cheese populations.

### 540 **E.4. Cheese populations inhibit more the growth of food spoil-** 541 **ers than wild populations**

542 Cheese is a protein- and fat-rich medium, where many mi-  
543 croorganisms, including desired microbes, but also spoil-  
544 ers, can thrive and thus compete for nutrients; for exam-  
545 ple, iron is limiting in cheese (Mayo et al., 2021; Mon-  
546 net et al., 2015, 2012). Cheese *G. candidum* populations  
547 may have been selected for excluding competitors by in-  
548 hibiting their growth (Boutrou and Gueguen, 2005). This  
549 fungus is known to inhibit fungal and bacterial food spoil-  
550 ers, such as *Aspergillus* species and *Listeria monocyto-*  
551 *genes*, but these inhibitory activities have only been inves-  
552 tigated in cheese *G. candidum* strains so far (Dieuleveux  
553 et al., 1998; Nielsen et al., 1998; Omeike et al., 2021). We  
554 therefore tested whether cheese populations displayed  
555 better growth inhibition abilities than the wild popula-  
556 tion, using common fungal food spoilers as competitors:  
557 *Debaryomyces hansenii*, *Penicillium bifforme*, *P. roqueforti*  
558 and *Scopulariopsis asperula*. We also tested whether growth  
559 inhibition of challengers occurred via secreted and/or  
560 volatile compounds.

561 **Inhibition by a mycelium lawn** - In the first experi-  
562 ment, we grew challengers in a central spot for 24h, alone  
563 or after spreading out *G. candidum* to let it grow as a lawn;  
564 growth inhibition could occur in this setting by secreted  
565 molecules in the medium, volatile compounds and/or  
566 a physical barrier to reach nutrients. The growth of *D.*  
567 *hansenii* was completely inhibited by all populations of *G.*  
568 *candidum*. *Penicillium roqueforti* was strongly inhibited by  
569 *G. candidum*, in particular by the Cheese\_2 population that  
570 completely prevented *P. roqueforti* growth (Figure 5; Table  
571 S13). The growth of *Scopulariopsis asperula* and *P. bifforme*  
572 was also inhibited by *G. candidum*, with a significant dif-  
573 ference between competitor growth when spread alone  
574 or on a *G. candidum* lawn; the Cheese\_2 population again



**Fig. 4. Differences in growth, opacity and volatile compounds between the five populations of *Geotrichum candidum***

Each point represents a strain, horizontal dotted lines and vertical lines represent the mean and the standard deviation of the phenotype in the population, respectively. The number *n* at the bottom of plots indicates the number of strains used per population for measuring the corresponding phenotypes. The pairwise Tukey tests performed to assess whether there were mean differences between populations are indicated between brackets with their *p*-values. (A) Mean radial growth of the three cheeses, mixed-origin and wild populations on cheese (1% salt), yeast peptone dextrose (YPD) and minimal media at 25°C. (B) Differences in opacity between the three cheese populations, mixed-origin and wild populations on cheese medium (1% salt) at 10°C. Integral opacity is defined as the sum of the brightness values for all the pixels within the fungal colony bounds and measures the whiteness and density of the mycelium. (C) Strain ESE00182 from the Cheese\_2 population showing the fluffiness of the colony (D) First two PCA axes of the principal component analysis (PCA) of *G. candidum* strains based on their relative proportions of different volatile compounds produced. (E) Contribution of each volatile compound to the first two PCA axes. The compounds contributing the most to the differentiation were colored in red and labeled (i.e., those distant from 0 by an Euclidean distance  $\geq 0.1$ ).

575 inhibited better competitors than any other population  
576 (Figure 5; Table S13). **Inhibition by volatile compounds**  
577 - In a second experiment, we used splitted Petri dishes  
578 (the two parts being separated by a plastic barrier) to test  
579 whether cheese populations inhibited competitors to a  
580 greater extent than the wild population when only volatile  
581 compounds can reach challengers. No significant growth  
582 difference was observed between the growth alone and  
583 at the side of *G. candidum* for neither *S. asperula* nor *P.*  
584 *roqueforti* (Figure 5B, Table S13). Only *P. bifforme* showed a  
585 significant growth inhibition by *G. candidum* in this setup  
586 (Table S13); such a growth inhibition by *G. candidum* from  
587 an isolated Petri dish compartment indicates that volatile  
588 compounds produced by *G. candidum* are able to impair  
589 the growth of some competitors.

### 590 3. DISCUSSION

#### 591 A. Domestication of the cheese-making fungus 592 *Geotrichum candidum*, with three varieties displaying 593 contrasting phenotypes

594 We found three differentiated clades based on genomic  
595 analyses of 98 *G. candidum* strains isolated from different  
596 kinds of cheeses, dairy products (e.g., raw milk), other  
597 food substrates (e.g., sausage) and other environments  
598 (e.g., plants). One clade was specific to cheeses, one to  
599 wild environments and one had mixed origins (dairy and  
600 other environments). Although the dairy environment  
601 was over-represented in our sampling, the 10 wild strains  
602 available captured a substantial diversity both in terms  
603 of substrates (e.g. soil, flower and polyurethane) and of  
604 geographic origins (i.e., Thailand, UK, French Guiana,  
605 Brazil, Egypt, Senegal, South Africa, Belgium, Spain and  
606 Sweden).

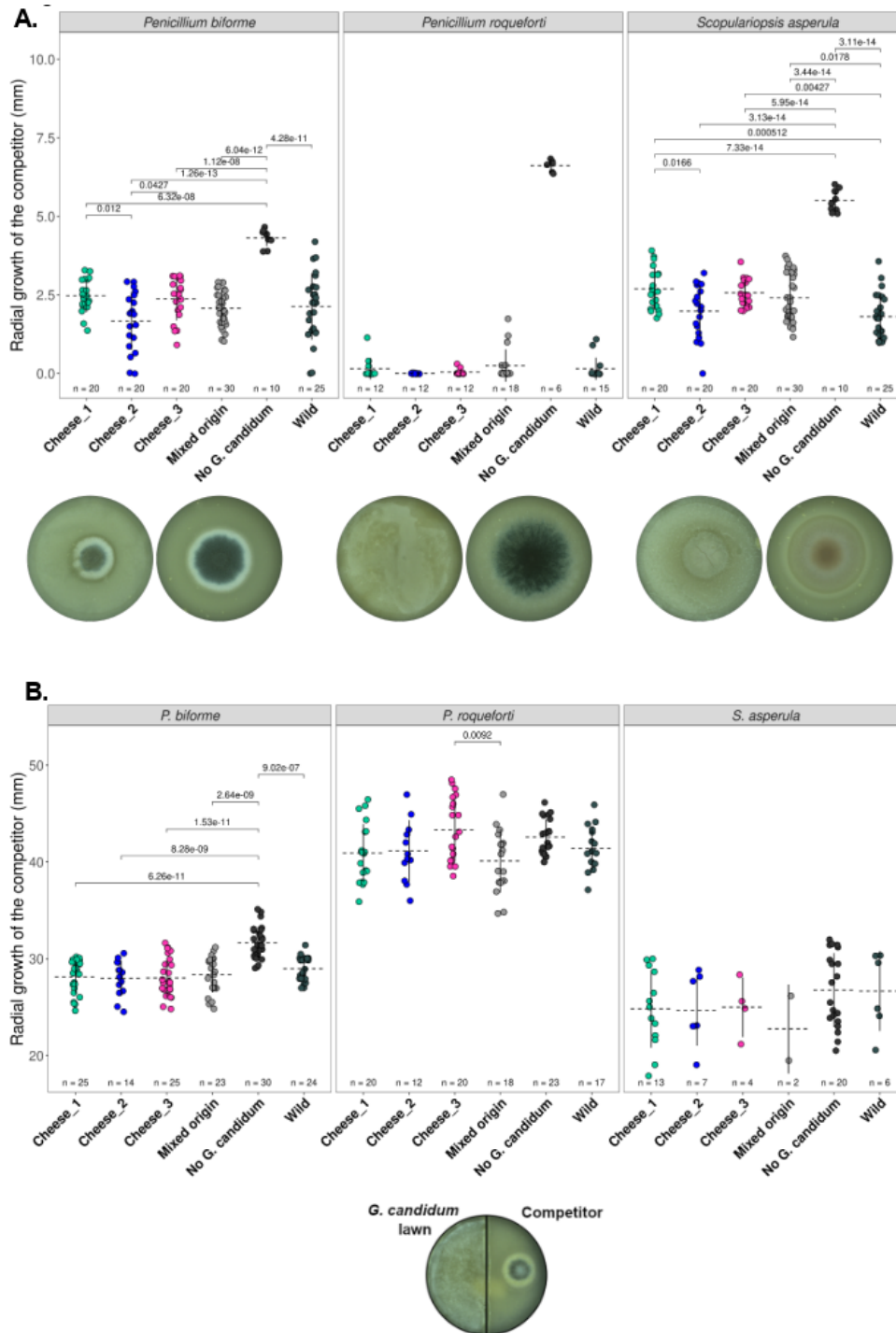
607 Within the cheese clade, we revealed the existence of  
608 three distinct genetic clusters and several admixed strains.  
609 The cheese\_1 population included all strains isolated from  
610 goat and soft cheeses, such as Sainte Maure de Touraine  
611 or crottin de chavignol. The other cheese populations cor-  
612 responded to strains isolated from other types of cheeses,  
613 such as pressed uncooked cheeses (e.g. Reblochon) or  
614 cooked (e.g. Comté) cheeses. This suggests that cheese-  
615 making practices have led to genetic differentiation by  
616 divergent selection. We did not detect in the present study  
617 any specific trait in the Cheese\_1 population compared  
618 to the Cheese\_3 population, but our experiments did not  
619 capture all aspects of cheese-making; in particular, a fea-  
620 ture of goat cheeses caused by *G. candidum* that could  
621 be investigated is their convoluted aspect. Alternatively,  
622 the clustering of soft goat cheese strains may be due to  
623 migration by strain sharing among cheese makers with  
624 similar products, as shown for example in bread yeasts  
625 (von Gastrow et al., 2022).

626 The cheese populations each had twice as low genetic  
627 diversity as the wild population, suggesting bottlenecks  
628 due to human selection. We did not, however, find that a  
629 single clonal lineage was used for most of cheeses, as is the

630 case in *P. roqueforti* and in *P. camemberti* (Dumas et al 2020;  
631 Ropars et al 2020), indicating the lack of such extreme  
632 bottlenecks in *G. candidum*. The overall diversity in the  
633 cheese clade with the three populations pooled was even  
634 higher than in the wild population, which may be due to  
635 the relatively low number of wild strains available. Alter-  
636 natively, the high diversity in the cheese clade as a whole  
637 may be due to i) the diversification into three populations  
638 under human divergent selection, ii) regular migration  
639 from a wild pool thanks to the natural occurrence of *G.*  
640 *candidum* in raw milk, thus preserving the diversity of  
641 strains isolated from cheeses, or iii) the divergence of each  
642 of the three cheese clusters from distinct but unsampled  
643 wild populations. According to this third hypothesis, the  
644 three cheese clusters would correspond to a diversifica-  
645 tion event that predated domestication, with selection of  
646 strains for cheese making from three already differenti-  
647 ated populations. In fact, the cheese clade was not nested  
648 within the wild clade, as could be expected if we had  
649 sampled the population-of-origin of the cheese clade. The  
650 traits in the studied wild clade may therefore not corre-  
651 spond exactly to those of the ancestral population(s) of  
652 the cheese clade.

653 We nevertheless found some evidence of domestication  
654 in the cheese clusters. We detected genomic footprints of  
655 adaptation to cheese, with the presence of genomic is-  
656 lands of differentiation and selective sweeps in cheese  
657 populations, in particular on a gene involved in iron up-  
658 take, that is limiting in cheese (Monnet et al., 2015), and  
659 on a metalloprotease involved in casein degradation (Mc-  
660 Sweeney, 2004). The cheese populations further lacked  
661 genes not required in the human-made environment, i.e.  
662 tandem beta lactamase-like genes. The cheese popula-  
663 tions may therefore have lost these genes due to relaxed  
664 selection, although we cannot exclude a gain in the wild  
665 clade. The cheese populations also carried a higher load  
666 of transposable elements, which may result from an ex-  
667 pansion following relaxed selection; we cannot rule out a  
668 reduction in repetitive elements in the wild clade instead,  
669 although it seems less likely.

670 We also found evidence of phenotypic adaptation to  
671 cheese making in *G. candidum* cheese populations, with  
672 shared traits specific to cheese populations, beneficial for  
673 cheese making and different from wild and mixed-origin  
674 populations. Cheese populations produced a greater  
675 amount of volatiles typical of cheese aromas and attrac-  
676 tive to consumers compared to wild strains. The cheese  
677 populations also displayed slower growth on all media,  
678 even cheese, and lower proteolysis activity, which may  
679 prevent a too fast degradation of products during matu-  
680 ration, as found in the Roquefort *P. roqueforti* population  
681 (Dumas et al. 2020) and in the dry-cured meat *Penicillium*  
682 fungi (Lo et al., 2022). Alternatively, the slower growth  
683 could represent degeneration due to the accumulation  
684 of deleterious mutations because of genetic drift. How-  
685 ever, the genetic diversity is not that low in the cheese  
686 clade, and slower growth seems a feature of cheese fungi



**Fig. 5. Competitive abilities of the different populations of *Geotrichum candidum* against *Penicillium bifforme*, *P. roqueforti* and *Scopulariopsis asperula* challengers**

(A) At the top, radial growth abilities of the competitors on lawns of *Geotrichum candidum* belonging to different populations (the three cheese populations, the mixed-origin and the wild populations). Each point represents a combination of the growth of a competitor strain on a lawn of a *G. candidum* strain. Horizontal dotted lines and vertical lines represent the mean and the standard deviation of the competitor growth in the population, respectively. The number n at the bottom of plots indicates the number of combinations of competitor-lawn used per population. The competitor was inoculated in a central point 24h later on a lawn of *G. candidum*. At the bottom, from left to right, are shown pictures of *P. biforme* ESE00023 on a *G. candidum* ESE00186 lawn and without any lawn, *P. roqueforti* ESE00645 on a *G. candidum* ESE00186 lawn and without any lawn, and *S. asperula* ESE01324 on a *G. candidum* ESE00198 lawn and without any lawn, all on a salted cheese medium.

B: At the top, radial growth abilities of competitors, with various *G. candidum* strains belonging to different populations being grown on the other side of splitted Petri dishes. The competitor was inoculated in a central spot on one side and the *G. candidum* strain was spread on the other side of the splitted Petri dish (a picture is shown as an example below the figure). The medium is not contiguous between the two sides of Petri dishes, so that inhibition can only occur by volatile compounds. Horizontal dotted lines and vertical lines represent the mean and the standard deviation of the competitor growth in the focal population, respectively. The number n at the bottom of plots indicates the number of combinations of competitor-lawn used per population.



(Ropars and Giraud 2022). These shared traits may have been acquired in their common (already domesticated) ancestor, or represent convergence in the case of repeated domestication from three different wild populations.

We did not detect higher salt tolerance or higher lipolysis ability in cheese populations. Higher salt tolerance and lipolysis rates have been reported in the non-Roquefort *P. roqueforti* population compared to non-cheese populations, but not in the Roquefort *P. roqueforti* population (Dumas et al 2020; Crequer 2023). The dry-cured meat *Penicillium* fungi even had lower lipolysis rates (Lo et al., 2022). This may be due to a lack of selection if fast growth of fungi is not beneficial for cheese making, for example if it leads to degraded products. There may also be evolutionary constraints for salt adaptation. Another hypothesis is that selection has not been efficient enough for changing multiple traits at the same time, and/or because of migration. As *G. candidum* is naturally present in raw milk, regular migration may occur between cheese and wild populations in artisanal cheeses that do not use commercial starters. We did not detect evidence of wild-to-cheese gene flow except in the admixed strains, but we may not have identified all wild populations.

Within the cheese clade, we found phenotypic differences between the three populations, in terms of colony density, volatile compounds produced and the ability of competitor inhibition. They thus correspond to different varieties in the sense of genetically differentiated clusters with contrasting phenotypes. These differences may result from selection for different traits, for example for making different cheese types or cheese-making practices. Alternatively, but not exclusively, part of the differences may correspond to pre-existing differentiated traits if the three cheese populations derived from different, unsampled, wild populations.

The *G. candidum* Cheese\_2 genetic cluster was phenotypically the most differentiated, with in particular a denser and fluffier mycelium and a higher competitive ability against challengers compared to the other populations. The Cheese\_2 population had a stronger inhibition ability than the other *G. candidum* populations when molecules could diffuse in the medium and the mycelium could act as a barrier. The Cheese\_2 population was also the most fluffy population on cheese and had a beta-glucanase gene under positive selection. Therefore, the challenger inhibition of Cheese\_2 may be due to either mycelium density as a physical barrier or degradation of competitor cell walls.

Our two sets of experiments enabled us to distinguish two mechanisms of competitor inhibition by *G. candidum*, and their action depended on the identity of the competitors. Indeed, *P. bifforme* growth was inhibited by *G. candidum* in the splitted Petri dishes, suggesting that volatile compounds are able to impair its growth. On the contrary, *P. roqueforti* and *S. asperula* were only inhibited by *G. candidum* when grown together, i.e., when molecules could diffuse in their medium and *G. candidum* mycelium could

form a physical barrier.

The genetic and phenotypic differentiation between the *G. candidum* clusters suggest that domestication occurred in several steps, with an ancient domestication event separating the mixed-origin and the wild clades, then the cheese and the mixed-origin clades, and yet more recently the three cheese clusters. Domestication also occurred stepwise in the *P. camemberti* clade (Ropars et al. 2020): the first step has led to the domestication of *P. bifforme* from the wild relative *P. fuscoglaucum*, and the second step to the divergence of *P. camemberti* from a *P. bifforme*-like population. Such stepwise domestication has also been reported in other domesticated fungi, i.e. Cachaça yeasts, which are derived from wine yeasts (Barbosa et al., 2018) and also in crops (Janzen and Hufford, 2016; Lin et al., 2014; Purugganan, 2019). Alternatively, the divergence of the different cheese populations may have predated domestication; they would then correspond to unsampled wild populations, from which strains would have been independently isolated for cheese making.

Nevertheless, the Cheese\_2 population appeared to represent a more advanced stage of domestication than the other cheese populations, with much lower genetic diversity, stronger genetic differentiation from other populations, and more differentiated traits from the wild population, with features beneficial for cheese making. The Cheese\_2 population indeed likely corresponds to an asexually cultivated line for cheese making, displaying low diversity, a single mating type, a high level of linkage disequilibrium, an absence of reticulation in the neighbor-net network and a non-significant PHI test. The Cheese\_2 population further display traits resembling other cheese fungi, with a denser and fluffier mycelium, a higher competitive ability and a complete lack of proteolysis activity. This may thus correspond to a stronger domestication syndrome, strikingly resembling the *P. camemberti* case. Indeed, *P. camemberti* is also a clonal lineage, recently derived from a cheese population with less extreme phenotypes, and having evolved drastic changes, with a white and fluffy mycelium (Ropars et al. 2020).

In contrast to the situations in *P. roqueforti* and *P. camemberti* (Dumas et al 2020, Ropars et al. 2020), commercial starter strains did not represent a single clonal lineage, being instead scattered throughout the cheese clade, some of them being admixed and/or belonged to small clonal groups (i.e. small terminal clusters in the tree with no branch length). These clonal lineages have each been isolated from different cheese types and different French regions, indicating asexual multiplication due to strain culture and commercialization, and migration between cheese makers (von Gastrow et al., 2022). The existence of admixed commercial starter strains, belonging to clonal lineages, suggests that admixed lineages have been selected for beneficial traits for cheese making. Hybrids have been selected for their particular traits beneficial for food production in other domesticated fungi, such as cold tolerance in the hybrid yeast *Saccharomyces pastorianus*

801 used for the production of lager beer (Gallone et al. 2018). 854

## 802 **B. Phenotypic convergence but contrasted genetic di-** 803 **versity patterns among cheese fungi**

804 We found evidence of convergence between *G. candidum*  
805 and other cheese-making fungi. As seen above, dense and  
806 white mycelial growth leading to a fluffy aspect at the  
807 expense of less rapid radial growth has strikingly evolved  
808 both in the Cheese\_2 *G. candidum* and the *P. camemberti* var.  
809 *camemberti* lineages (Ropars et al., 2020b), thus represent-  
810 ing a convergent phenotype between two distantly related  
811 cheese fungi (Ropars et al 2022; Ropars and Giraud 2022).  
812 *Geotrichum candidum* is increasingly inoculated in milk in  
813 the place of *P. camemberti* for industrial soft cheese pro-  
814 duction, as it provides the white and fluffy desired aspect  
815 without the disadvantage of *P. camemberti* that browns the  
816 surface of Camembert cheeses at the end of the ripening  
817 process (Carreira et al., 2002). Another convergence con-  
818 cerned proteolysis activity, which is low in *G. candidum*  
819 cheese populations as well as in the Roquefort *P. roque-*  
820 *forti* population, which may be crucial to avoid degraded  
821 cheeses (Dumas et al. 2020). In terms of volatile com-  
822 pounds, we found specific and attractive flavors for cheese  
823 making in cheese strains, different from wild strains, as  
824 also documented in *P. roqueforti* (Caron et al., 2021; Dumas  
825 et al., 2020).

826 In contrast to the striking convergence of phenotypes  
827 (Figure 4C), we found different patterns of diversity in  
828 *G. candidum* compared to *P. camemberti*. While *P. camem-*  
829 *berti* and the non-Roquefort *P. roqueforti* populations are  
830 each a clonal lineage with virtually no genetic diversity  
831 (Dumas et al. 2020; Ropars et al. 2020), the *G. candidum*  
832 populations harboured each some diversity, even the less  
833 diverse, the Cheese\_2 population. The Cheese\_2 popu-  
834 lation of *G. candidum* indeed exhibited four times as low  
835 genetic diversity as the two other cheese populations, be-  
836 ing of the same order of magnitude as in the Roquefort *P.*  
837 *roqueforti* population (Dumas et al., 2020). The Cheese\_1  
838 population of *G. candidum* displayed the same level of  
839 diversity as the cheese-making species *P. biforme* and the  
840 Cheese\_3 population as its wild relative *P. fuscoglaucum*.  
841 The different patterns of diversity may be due to different  
842 cheese-making practices, as *P. camemberti* and *P. roqueforti*  
843 are inoculated at the beginning of the cheesemaking pro-  
844 cess, while *P. biforme* and *G. candidum* are also naturally  
845 present in the cheese environment, either in raw milk or  
846 in barns.

847 The differentiation level between the whole cheese  
848 clade and the mixed-origin population in *G. candidum*  
849 was similar to that found between the domesticated *P.*  
850 *camemberti* mold and its wild closest relative species, *P.*  
851 *fuscoglaucum* (FST = 0.83; dxy = 6E-03), further support-  
852 ing the view that the three cheese populations have been  
853 domesticated.

## 854 **4. CONCLUSION**

855 Our study shows that fungi are excellent models to study  
856 domestication and independent adaptation events to sim-  
857 ilar environments and usage (Ropars and Giraud 2022).  
858 This is an important topic in evolutionary biology as it is  
859 important to understand whether independent adaptation  
860 events to similar environments leads to convergence in  
861 phenotypes, i.e., whether evolution is repeatable (Alberto  
862 et al., 2018; Cresko et al., 2004; Dyer et al., 2012; Elmer  
863 et al., 2014, 2010; Lin et al., 2012; Macías et al., 2021; O'Quin  
864 et al., 2010; Thorpe et al., 2015). We found here both sim-  
865 ilarities (convergence) and differences in the adaptation  
866 of *G. candidum* to cheese compared to other cheese fungi.  
867 One of the most striking convergent traits was the fluffy  
868 and white mycelium in *G. candidum* as in *P. camemberti*  
869 with a similar trade-off with radial growth (Ropars et  
870 al., 2020b). Greater competitive ability and lower prote-  
871 olysis activity, as well as greater production of positive  
872 volatile compounds, also represent interesting convergent  
873 phenotypes between *G. candidum* and *Penicillium* cheese  
874 fungi, which are very distant fungal lineages. It will be  
875 interesting in future studies to investigate the genomic  
876 mechanisms underlying convergence, to assess whether  
877 the evolution of similar traits arose by similar or different  
878 de novo mutations, or by introgression or by the hori-  
879 zontal transfer of the same genomic regions, as already  
880 reported in *Penicillium* cheese fungi (Ropars et al. 2015,  
881 Cheeseman et al 2014) and dry-cured meat *Penicillium*  
882 fungi (Lo et al., 2022).

883 Our findings also have industrial implications, reveal-  
884 ing a high genetic diversity and genetic subdivision in a  
885 fungus widely used in the cheese industry, and the exist-  
886 ence of genetically and phenotypically different popula-  
887 tions used for cheese making, with specific and contrasted  
888 traits beneficial for cheese making. The most fluffy and  
889 most competitive cheese population corresponded to a  
890 clonal lineage which may represent the most recent and  
891 strongest selection event. The occurrence of recombina-  
892 tion between cheese strains is highly relevant for cheese  
893 producers as it opens possibilities for further improve-  
894 ment for cheese making. It is crucial to maintain a large  
895 genetic diversity in cheese *G. candidum* populations as this  
896 is essential for variety improvement and diversification  
897 and to avoid degeneration (Harlan et al., 2012).

## 898 **5. MATERIAL AND METHODS**

### 899 **A. Sampling**

900 We isolated 53 strains from different kinds of cheeses (e.g.  
901 Camembert, Brie, Saint Nectaire, Ossau Iraty, comté, bleu  
902 de chèvre) from five European countries, Canada and  
903 the USA. Cheese crusts were left in the freezer for 24h  
904 to kill acarians. Then, we diluted a piece of each crust  
905 in sterile water and spread 50  $\mu$ L of the suspension on a  
906 malt agar Petri dish. When colonies appeared on the Petri  
907 dish, typically after three days, we isolated the different  
908 morphotypes with a sterile toothpick and inoculated them

909 on new Petri dishes. After seven days of growth, we  
910 performed monospore isolation by several dilution steps,  
911 in order to obtain separated colonies arising each from  
912 a single spore. We identified the species of these pure  
913 strains after DNA extraction by sequencing the 5' end of  
914 the nuclear ribosomal large subunit (LSU rDNA) using the  
915 LROR/LR6 oligonucleotide primers (Vilgalys and Hester,  
916 1990). We also gathered 24 strains from INRAE, isolated  
917 from cheeses but also other environments (e.g. sand, hay,  
918 rainforest) and 15 strains from a French spore seller. We  
919 gathered all the wild strains available in public collections.  
920 For each strain, single spore cultures were generated to  
921 ensure the presence of a single genotype before DNA  
922 extraction.

923 The LMA-244 strain were inoculated on Yeast Extract  
924 Glucose (YEG) agar plates (10 g.L-1 of yeast extract (Fis-  
925 cher Scientific), 10 g.L-1 of D-glucose (EMD Chemicals)  
926 and 15 g.L-1 of Bacto agar (BD Diagnostics)) directly from  
927 15% glycerol (v/v) stock cultures stored at -80°C. The  
928 plates were incubated in the dark for five days at 25°C.

## 929 B. DNA extraction, genome sequencing, assembly, an- 930 notation and mapping

931 We used the Nucleospin Soil Kit (Macherey-Nagel, Düren,  
932 Germany) to extract DNA from 88 *G. candidum* strains  
933 cultured for five days on malt agar. Sequencing was per-  
934 formed with Illumina HiSeq 3000 paired-end technology  
935 (Illumina Inc.), 2x150 bp. For the eight LMA strains, se-  
936 quencing was performed using the Illumina HiSeq paired-  
937 end technology.

938 All Illumina reads were trimmed and adapters cleaned  
939 with Trimmomatic v0.36 (Bolger et al., 2014). Leading or  
940 trailing low quality or N bases below a quality score of  
941 three were removed. For each read, only parts that had  
942 an average quality score higher than 20 on a four base  
943 window are kept. After these steps, only reads with a  
944 length of at least 36 bp were kept.

945 Cleaned Illumina reads were assembled with SPAdes  
946 v3.15.3 not using unpaired reads with “-careful” param-  
947 eter.

948 For the LMA-244 strain, Genomic DNA was extracted  
949 using the Fungi/Yeast Genomic DNA Isolation Kit (Nor-  
950 gen Biotek Corp.) with the following modifications.  
951 Thirty milligrams of frozen grounded mycelium were  
952 thawed and homogenized in 500 µL of a 0.9% NaCl so-  
953 lution. The elution buffer was replaced by a Tris 10 mM  
954 buffer (pH 8). Following the extraction step, gDNA sus-  
955 pensions were purified and concentrated using Agencourt  
956 AMPure XP magnetic beads (Beckman-Coulter), accord-  
957 ing to the manufacturer’s protocol.

958 DNA concentration and purity were measured using a  
959 NanoDrop ND-1000 spectrophotometer (Thermo Fisher  
960 Scientific Inc., Wilmington, U.S.A.) and a Qubit Fluorome-  
961 ter 3.0 (Thermo Fisher Scientific Inc., Wilmington, U.S.A.).

962 The DNA library was prepared following the Pacific  
963 Biosciences 20 kb template preparation using BluePippin  
964 Size-Selection System protocol and the Pacific Biosciences

965 Procedure & Checklist – Preparing Multiplexed Microbial  
966 Libraries Using SMRTbell Express Template Prep Kit 2.0  
967 protocol. No DNA shearing was performed. The DNA  
968 damage repair, end repair and SMRT bell ligation steps  
969 were performed as described in the template preparation  
970 protocol with the SMRTbell Template Prep Kit 1.0 reagents  
971 and the SMRTbell Express Template Prep Kit 2.0 reagents  
972 (Pacific Biosciences, Menlo Park, CA, USA). The DNA  
973 library was size selected on a BluePippin system (Sage  
974 Science Inc., Beverly, MA, USA) using a cut-off range of  
975 10 kb to 50 kb. The sequencing primer was annealed at a  
976 final concentration of 0.83 nM and the P6 v2 polymerase  
977 was bound at 0.50 nM while the sequencing primer was  
978 annealed with sequencing primer v4 at a final concentra-  
979 tion of 1 nM and the Sequel 3.0 polymerase was bound  
980 at 0.5 nM. The libraries were sequenced on a PacBio RS  
981 II instrument at a loading concentration (on-plate) of 160  
982 pM using the MagBead OneCellPerWell loading protocol,  
983 DNA sequencing kit 4.0 v2, SMRT cells v3 and 4 hours  
984 movies.

985 Raw PacBio reads were corrected using Illumina reads  
986 already available and described in a previous article  
987 (Perkins et al., 2020), with the default parameters of the  
988 LoRDEC software and trimmed with Canu v1.6 (Koren  
989 et al., 2017; Salmela and Rivals, 2014). Corrected and  
990 trimmed PacBio reads were then assembled using Canu  
991 v1.6. Illumina polishing of the Canu assembly was per-  
992 formed using Pilon v1.22 (Walker et al., 2014). A final  
993 assembly step was then performed with the hybrid assem-  
994 bler SPAdes v3.11.1 using the trimmed PacBio reads, the  
995 Illumina reads and the Pilon corrected assembly as trusted  
996 contigs (Antipov et al., 2016; Prjibelski et al., 2020). Ad-  
997 ditionally, the CLIB 918 assembly (Bioproject PRJEB5752)  
998 was used as a reference in the SPAdes script for the as-  
999 sembly of each *G. candidum* genome (Morel et al., 2015).  
1000 Scaffolds were filtered using the khmer software with a  
1001 length cut-off of 1,000 bp (Crusoe et al., 2015).

1002 The LMA-244 PacBio assembly and reads have been de-  
1003 posited in GenBank: nbPROJECT. To annotate short read  
1004 assemblies and the LMA-244 genome, gene prediction  
1005 was performed using Augustus v3.4.0 (Stanke et al., 2008).  
1006 The training annotation file “saccharomyces” was used,  
1007 with parameters as follows: “-gff3=on”, “-protein=on”,  
1008 “-codingseq=on”, “-exonnames=on”, “-cds=on” and “-  
1009 uniqueGeneId=true”. The output of Augustus and the  
1010 CLIB 918 gff was provided to Funannotate v1.8.9 (Palmer  
1011 and Stajich, 2020) for functional annotation. InterProScan  
1012 was used under Funannotate pipeline locally (Blum et al.,  
1013 2021). Funannotate then searched in the Pfam database  
1014 v34.0 and dbCAN database version 10.0 with Hmmer  
1015 v3.3.2 (Eddy, 2011; Huang et al., 2018; Mistry et al.,  
1016 2021), in database UniProt version 2021\_03 and database  
1017 MEROPS version 12.0 with diamond blastp v2.0.11 (Rawl-  
1018 ings et al., 2018; The UniProt Consortium et al., 2021),  
1019 eggNOG-mapper v2 on the database eggNOG 5.0 (Canta-  
1020 lapiedra et al., n.d.; Huerta-Cepas et al., 2019).

1021 Cleaned reads were mapped on the reference genomes

1022 CLIB 918 and LMA-244 using Bowtie2 v2.4.2 (Langmead 1076  
1023 and Salzberg, 2012). Maximum fragment length was set 1077  
1024 to 1000 and the preset “very-sensitive-local” was used. 1078

1025 SAMtools v1.7 (Li et al., 2009) was used to filter out 1079  
1026 duplicate reads and reads with a mapping quality score 1080  
1027 above ten for SNP calling and above one for CNV analy- 1081  
1028 ses. 1082

1029 In total, we have a dataset of 98 genomes, 88 being se- 1083  
1030 quenced (Bioproject PRJNA866540), eight from the Univer- 1084  
1031 sity of Laval (LMA strains: Bioproject PRJNA482576, PR- 1085  
1032 JNA482605, PRJNA482610, PRJNA482613, PRJNA482616, 1086  
1033 PRJNA482619, PRJNA490507, PRJNA490528), one strain 1087  
1034 CLIB 918 from the Collection de Levures d’Intérêt Biotech- 1088  
1035 nologique (Bioproject PRJEB5752), and one of the strain 1089  
1036 Phaff72-186 from the 1000 Fungal Genomes project (Bio- 1090  
1037 project PRJNA334358 NCBI). 1091

### 1038 C. SNP calling 1093

1039 Single nucleotide polymorphisms (SNPs) were called 1094  
1040 using GATK v4.1.2.0 HaplotypeCaller, which provides 1095  
1041 one gVCF per strain (option -ERC GVCF). GVCFs were 1096  
1042 combined using GATK CombineGVCFs, genotypes with 1097  
1043 GATK GenotypeGVCFs, SNPs were selected using GATK 1098  
1044 SelectVariants (option -select-type SNP). SNPs were fil- 1099  
1045 tered using GATK VariantFiltration and options QUAL 1100  
1046 < 30, DP < 10, QD < 2.0, FS > 60.0, MQ < 40.0, SOR > 1101  
1047 3.0, QRankSum < -12.5, ReadPosRankSum < -8.0. All pro- 1102  
1048 cesses from cleaning to variant calling were performed 1103  
1049 with Snakemake v5.3.0 (script available at [https://github.](https://github.com/BastienBennetot/Article_Geotrichum_2022) 1104  
1050 [com/BastienBennetot/Article\\_Geotrichum\\_2022](https://github.com/BastienBennetot/Article_Geotrichum_2022)). 1105

### 1051 D. Phylogenetic analysis 1107

1052 We inferred phylogenetic relationships among the 98 iso- 1108  
1053 lates using the dataset of 699,755 SNPs in a maximum 1109  
1054 likelihood framework using IQ-Tree2 v2.1.1 (Minh et al., 1110  
1055 2020). The tree has been midpoint rooted. The best-fit 1111  
1056 model chosen according to Bayesian information crite- 1112  
1057 rion (BIC) was TVMe+R2. Branch supports are ultrafast 1113  
1058 bootstrap support (1000 bootstrap replicates, Minh et al., 1114  
1059 2013). 1115

### 1060 E. Genetic structure 1116

1061 We used the dataset of 699,755 SNPs to infer population 1117  
1062 structure based on the mapping on the CLIB 918 reference 1118  
1063 genome. We used Splitstree v4.16.2 (Huson and Bryant, 1119  
1064 2006) for the neighbor-net analysis. We used the R pack- 1120  
1065 age Ade4 (Bougeard and Dray, 2018; Chessel et al., 2004; 1121  
1066 Dray et al., 2007; Dray and Dufour, 2007; Thioulouse et 1122  
1067 al., 2018) for principal component analyses (PCA, cen- 1123  
1068 tered and unscaled). We used NGSadmix v.33 (Jørsboe 1124  
1069 et al., 2017) from the ANGSD (Korneliussen et al., 2014) 1125  
1070 package (version 0.933-110-g6921bc6) to infer individual 1126  
1071 ancestry from genotype likelihoods based on realigned 1127  
1072 reads, by assuming a given number of populations. A Bea- 1128  
1073 gle file was first prepared from bam using ANGSD with 1129  
1074 the following parameters: “-uniqueOnly 1 -remove\_bads 1130  
1075 1 -only\_proper\_pairs 1 -GL 1 -doMajorMinor 1 -doMaf 1 1131

-doGlf 2 -SNP\_pval 1e-6”. The Beagle file was used to run  
NGSadmix with 4 as the minimum number of informative  
individuals. Given the high number of strains genetically  
highly similar among cheese strains (that may represent  
clonal lineages), we randomly sampled one of the individ-  
uals for each group of clonemates identified on the ML  
tree as having fewer than 1,200 SNPs and filtered out the  
other strains (N=64 strains kept) to avoid biasing the anal-  
ysis. The analysis was run for different K values, ranging  
from 2 to 10. A hundred independent runs were carried  
out for each number of clusters (K).

The nucleotide diversity  $\pi$  (Nei’s  $P_i$ ; Hudson et al.,  
1992; Nei and Li, 1979), the Tajima’s  $D$ , the Watter-  
son’s  $\theta$  (Watterson, 1975), the fixation index  $F_{ST}$  (Hud-  
son et al., 1992) and the absolute divergence  $d_{XY}$  (Nei  
and Li, 1979) were calculated using the popgenome  
package in R (Pfeifer et al., 2014). Fixed, private and  
shared sites were counted using custom scripts available  
at [https://github.com/BastienBennetot/fixd\\_shared\\_private\\_](https://github.com/BastienBennetot/fixd_shared_private_count)  
[count](https://github.com/BastienBennetot/fixd_shared_private_count), with bcftools version 1.11 (using htlib 1.13+ds).  
 $F_3$  tests were computed using the admixr package v0.9.1.  
The pairwise homology index (PHI) test was performed  
using PhiPack v1.1 and CLIB 918 genome as reference.

Linkage disequilibrium was calculated using vcftools  
v0.1.17 with the -hap-r2 parameter and a minimum dis-  
tance between SNPs of 15,000 bp. Values were averaged  
when SNPs had the same distance.

Pairwise identity between an admixed strain and each  
non-admixed strain was calculated using overlapping slid-  
ing windows of 30 kb span and 5 kb step. Admixed  
clusters are indicated in Table S1. The custom script  
is available on [https://github.com/BastienBennetot/Article\\_](https://github.com/BastienBennetot/Article_Geotrichum_2022)  
[Geotrichum\\_2022](https://github.com/BastienBennetot/Article_Geotrichum_2022)

### 1117 F. Copy number variation and identification of premature 1118 stop codons in CDS

Copy number variation (CNV) was analyzed using  
Control-FREEC v11.6 with the following parameters:  
ploidy was set to 1, non-overlapping windows of 500  
bp, telomeric and centromeric regions were excluded, ex-  
pected GC content was set between 0.25 and 0.55, mini-  
mum of consecutive windows to call a CNV set to 1. This  
analysis was performed using as references the CLIB 918  
(cheese\_3) and LMA-244 (wild) genome sequences. CNVs  
were classified in different groups when the median of  
copy number was different between populations. We de-  
fined three groups: regions for which copy number was  
different between wild and cheese populations, between  
mixed-origin and cheese populations and when at least  
one cheese population differed from another population.  
For each InterPro term present in these regions, we per-  
formed enrichment tests, i.e., a fisher exact test compar-  
ing the number of a particular InterPro term found in these  
regions and the whole genome (Table S9).

We used snpeff (Cingolani et al., 2012) to assess how  
each SNP affected the coding sequence of predicted pro-  
teins, in the vcf file containing all SNPs and all genomes

1132 of our dataset. We detected premature stop codons in the 1185  
1133 7,150 CDS of the CLIB 918 genome and the 5,576 CDS of 1186  
1134 the LMA-244 genome using a custom script and bcftools 1187  
1135 v1.11. 1188

## 1136 G. Analyzing the repeat landscape

1137 In order to de novo detect repeats within *G. candidum*, 1191  
1138 RepeatModeler (v2.0.2; Flynn et al., 2020), using the ncbi 1192  
1139 engine (-engine ncbi) and the option -LTRStruct, was run 1193  
1140 on the pacbio genome assembly of LMA 244 generating 1194  
1141 a library of 176 repeats. The repeat redundancy was re- 1195  
1142 duced using cd-hit-est, as described in Goubert et al., giv- 1196  
1143 ing a final library of 108 repeats (Goubert et al., 2022). 1197  
1144 To estimate the per strain copy number of each repeat, 1198  
1145 illumina reads were aligned using bwa mem (v0.7.17; Li, 1199  
1146 2013) to the repeat library and the median coverage for 1200  
1147 each repeat was then normalized by the LMA 244 genome 1201  
1148 wide median coverage. 1202

## 1149 H. Detection of selective sweeps

1150 We used SweeD (Pavlidis et al., 2013) to detect selective 1203  
1151 sweeps based on site frequency spectra. We set the ploidy 1204  
1152 to 1 and the number of positions in the alignment where 1205  
1153 the CLR was computed to 1000 (gris option). We only kept 1206  
1154 the windows with the 1% highest likelihood per scaffold 1207  
1155 for each cheese population. 1208

## 1156 I. Detecting positive selection

1157 The assemblies LMA-317, LMA-77 and LMA-563 have 1209  
1158 been excluded for this analysis because of a N50 under 1210  
1159 ten kb. All the 437441 predicted protein sequences from 1211  
1160 the 66 genomes of all cheese clades and mixed-origin clade 1212  
1161 were searched against each other with BLASTP using dia- 1213  
1162 mond v0.9.36 and clustered into orthologous groups using 1214  
1163 Orthogog v1.0.3 (Ekseth et al., 2014). For these analyses, 1215  
1164 we only kept single-copy orthologs shared between two 1216  
1165 populations. We compared the mixed origin population 1217  
1166 to each cheese population and the cheese clade. Multiple 1218  
1167 nucleotide sequence alignments with predicted gene se- 1219  
1168 quences were then constructed using MACSE v2.0.3 with 1220  
1169 default parameters (Ranwez et al., 2018). We performed 1221  
1170 an approximative McDonald Kreitman tests using the R 1222  
1171 package PopGenome (Pfeifer et al., 2014). The approxima- 1223  
1172 tion comes from the fact that only codons with a single 1224  
1173 SNP are examined. The assumption of this version of the 1225  
1174 test is that the probability that two SNPs will appear in 1226  
1175 the same codon is very low. To identify genes evolving un- 1227  
1176 der positive selection in *G. candidum* genomes,  $\alpha$ , i.e. the 1228  
1177 representation of the proportion of substitutions driven 1229  
1178 by positive selection was used. Genes with an alpha un- 1230  
1179 der 0 were filtered out. Of these genes, only those with a 1231  
1180 Fisher's test p-value under 0.05 were kept. 1232

## 1181 J. Phenotypic characterization

### 1182 J.1. Sampling and strain calibration

1183 We used 36 *Geotrichum candidum* strains for laboratory 1233  
1184 experiments: seven from the Cheese\_1 population, five 1234

1185 from the Cheese\_2 population, eleven from the Cheese\_3 1186  
1187 population, eight from the mixed-origin population and 1188  
1189 five from the wild population (Table S1). This set encom- 1190  
1191 passed 26 strains isolated from dairies, one from other 1192  
1193 food environments and nine isolated from environments 1194  
1195 other than food. Experiments were initiated with spore 1196  
1197 suspensions calibrated to  $1.10^6$  spores/mL with a hemo- 1198  
1199 cytometer. 1200

### 1193 J.2. Media preparation

1194 All media were sterilized in an autoclave at 121°C for 20 1195  
1196 minutes except those with cheese or milk for which the au- 1197  
1198 toclave was run at 110°C for 15 minutes to avoid curdling. 1199  
1200 Each 94mm-diameter Petri dish was filled with 25mL of 1201  
1202 the appropriate medium. Cheese medium was prepared 1203  
1204 as follows for 800mL: 300g of unsalted cream cheese from 1205  
1206 La Doudou farm in Cheptainville, 16g agar, 8g NaCl dis- 1207  
1208 solved in 200mL of deionized water. Deionized water was 1208  
1209 added to reach 800mL. pH was adjusted to 6.5 and drops 1209  
1210 of blue food dyes were added to enable fungal colony 1210  
1211 measures (white medium and white colonies are not dis- 1211  
1212 tinguable). Yeast Peptone Dextrose (YPD) medium was 1212  
1213 prepared as follows for 1L: 10g Yeast extract, 10g Bacto 1213  
1214 Peptone, 10g glucose, 14g agar powder. Minimal medium 1214  
1215 was prepared as described in "Improved protocols for 1215  
1216 *Aspergillus* minimal medium: trace element and mini- 1216  
1217 mal medium salt stock solutions", Terry W. Hill, Rhodes 1217  
1218 College, Etta Kafer, Simon Fraser University. Tributyrin 1218  
1219 agar was prepared as follows: Tributyrin medium 33 g/L, 1219  
1220 neutral Tributyrin 10 g/L, Bacto Agar 15 g/L. Ingredients 1220  
1221 were bought at Nutri-Bact company, Québec, Canada. Ca- 1221  
1222 seinate agar was prepared according to Frazier and Rupp, 1222  
1223 modified as follows: Calcium caseinate medium 37.2 g/L, 1223  
1224 Bacto Agar 15 g/L. Ingredients were bought at Nutri- 1224  
1225 Bact company, Québec, Canada. For yogurt media we 1225  
1226 used three different types of raw milk, i.e. sheep, goat 1226  
1227 and cow milks, coming from d'Armenon farm near Les 1227  
1228 Molières (Esonne, France), Noue farm in Celle les Bordes 1228  
1229 and Coubertin farm in Saint-Rémy-lès-Chevreuse respec- 1229  
1230 tively. Each medium was prepared following the same 1230  
1231 procedure: 1L of milk was mixed with 62.5g of Danone 1231  
1232 brand yogurt, heated for 5 hour at 43°C and stored in a 1232  
1233 fridge before use. A subset of 300g of this preparation 1233  
1234 was used with 16g of agar powder, 8g of NaCl, 4 drops of 1234  
1235 blue food dye and filled up with deionized water to reach 1235  
1236 800mL. 1236

### 1230 J.3. Growth in different conditions and different media

1231 Petri dishes were inoculated with 10 $\mu$ L of the 1.106 1231  
1232 cells/mL in a 10% glycerol solution. Inoculated Petri 1232  
1233 dishes were wrapped with plastic film before letting them 1233  
1234 grow in the dark. A millimeter rule was used to mea- 1234  
1235 sure two opposite diameters of fungal colonies to estimate 1235  
1236 their growth. Means of these two measures were used for 1236  
1237 statistical analyses. 1237

1238 To test media and temperature effect on growth, *G.* 1238  
1239 *candidum* strains were grown on minimal, YPD and cheese 1239  
1240 media. We took pictures and measured their growth at 1240

1241 seven, 11 and 14 days for minimal, YPD and cheese media 1297  
1242 at 10°C (ripening cellar temperature), at seven and 11 1298  
1243 days for the cheese medium at 15°C and at seven days for 1299  
1244 minimal, YPD and cheese media at 25°C (Figure S8). 1300

1245 To test salt tolerance, *G. candidum* strains were grown at 1301  
1246 10°C on cheese media of different salt concentrations: un- 1302  
1247 salted media, 1% salt as St Nectaire and cream cheeses, 2% 1303  
1248 as Camembert and goat cheeses and 4% as Roquefort blue 1304  
1249 cheeses. We took pictures and measured colony diameters 1305  
1250 after 14 days of growth. 1306

1251 To test adaptation of *G. candidum* populations to dif- 1307  
1252 ferent milk origins, growth was measured on different 1308  
1253 yogurt media made from goat, sheep and cow raw milk  
1254 for seven days at 25°C.

1255 To test lipolytic and proteolytic activities of *G. candidum* 1309  
1256 populations, we grew strains on tributyrin agar and ca- 1310  
1257 caseinate agar, respectively. Each strain was inoculated in 1311  
1258 triplicate Petri dishes that were let grown at 25°C for 14 1312  
1259 days. The radius of lysis was measured and the mean 1313  
1260 between triplicates was used for the analysis. 1314

1261 Pictures were taken using a Scan 1200 (Interscience). 1315  
1262 Petri dishes grown on cheese were analyzed using IRIS 1316  
1263 (Kritikos et al., 2017) which measured Integral opacity 1317  
1264 scores, defined as the sum of the brightness values for all 1318  
1265 the pixels within the colony bounds. 1319

#### 1266 **J.4. Volatile compounds analysis using Gas-chromatography** 1320 1267 **mass-spectrometry (GC-MS)** 1321

1268 Volatile compounds produced by *G. candidum* were ana- 1322  
1269 lyzed using gas-chromatography mass-spectrometry (GC- 1323  
1270 MS). Compounds were extracted and concentrated by 1324  
1271 using a dynamic headspace (DHS) combined with a ther- 1325  
1272 mal desorption unit (TDU). Strains were grown for 21 1326  
1273 days at 10°C (minimum Camembert ripening time) on a 1327  
1274 cheese agar medium made with Camembert-type curds. 1328  
1275 After 21 days, each Petri dish content, with its medium 1329  
1276 and *G. candidum* mycelium, was mixed with a fork for 1330  
1277 one minute, gathered in vials and immediately frozen in 1331  
1278 liquid nitrogen. For each sample, three grams of frozen 1332  
1279 cultured media were weighted and stored in vials with 1333  
1280 septum caps at -80°C. Sixteen hours before analysis, sam- 1334  
1281 ples were stored at 4°C. The Cheese\_2 population was not 1335  
1282 tested in this experiment because population delineation 1336  
1283 was not known at this time. 1337

1284 Dynamic headspace (DHS) conditions were as follows: 1338  
1285 Inert gas: He; Incubation: 30°C for 3min; Needle tem- 1339  
1286 perature: 120°C; Trap: nature tenax, 30°C, 450 mL He; 1340  
1287 He flow: 30 mL/min; Dry purge : temperature 30°C, 850 1341  
1288 mL He, He flow 50 mL/ min. Thermal Desorption Unit 1342  
1289 (TDU) conditions were as follows: inert gas : He; Initial 1343  
1290 temperature: 30°C, then 60°C/min until 290°C kept for 1344  
1291 7 minutes; Transfer temperature: 300°C. Cool Injection 1345  
1292 System (CIS) conditions were as follows: inert gas : He; 1346  
1293 Initial temperature: -100°C, then 12°C/s until 270°C kept 1347  
1294 for 5 minutes. Gas chromatograph (brand Agilent 7890B) 1348  
1295 was used with a polyethylene glycol (PEG) type polar 1349  
1296 phase column (HP-Innowax, ref. Agilent 19091N-116I, 1350

60m×0.32mm, 0.25µm film thickness). Helium flow was  
set at 1.6mL/min. Samples were injected in splitless mode  
with a holding time of 1 minute. To optimize separation  
of compounds, a specific program of the gas chromatog-  
raphy oven was used, with initial temperature at 40°C for  
5 minutes, rising temperature from 40°C to 155°C with a  
slope of 4°C/min, rising temperature from 155°C to 250°C  
with a slope of 20°C/min and then temperature was kept  
at 250°C for 5 minutes. A single quadrupole mass spec-  
trometer was used to determine m/z of sample molecules  
(Agilent, référence 5977B MSD). Molecules were identified  
using NIST libraries (NIST 2017 Mass Spectral Library).

#### 1309 **J.5. Competition experiments**

1310 To test the abilities of *G. candidum* populations to ex- 1310  
1311 clude other fungi by secreting molecules or volatile com- 1311  
1312 pounds, we compared the growth of competitors when 1312  
1313 grown alone and on a lawn of an already grown *G. can-* 1313  
1314 *didum* mycelium. We inoculated a cheese medium with 1314  
1315 150µL of a *G. candidum* calibrated spore solution (1.10<sup>6</sup> 1315  
1316 spores/mL), spread evenly on the Petri dish. After 24h of 1316  
1317 growth, we inoculated 10µL of a competitor spore solution 1317  
1318 (1.10<sup>6</sup> spores/mL) in a single spot, in the middle of the 1318  
1319 Petri dish. We used as competitors the following species 1319  
1320 and strains: *Penicillium bifforme* (ESE00018, ESE00023, 1320  
1321 ESE00125, ESE00222), *Penicillium roqueforti* (ESE00645, 1321  
1322 ESE00925, LCP06040), *Scopulariopsis asperula* (ESE00044, 1322  
1323 ESE00102, ESE00835, ESE01287, ESE01324) and *Debary-* 1323  
1324 *omyces hansenii* (ESE00284, ESE00561, ESE00576; Table 1324  
1325 S15). For each competitor, two Petri dishes were inocu- 1325  
1326 lated without any *G. candidum* as controls for measuring 1326  
1327 growth without a lawn. 1327

1328 We took pictures of the Petri dishes at 6 days, when 1328  
1329 the competitor mycelium grown alone was near the Petri 1329  
1330 dishes border; we measured colony size at 7 days for *P.* 1330  
1331 *bifforme* and *P. roqueforti* and at 19 days for *D. hansenii*, 1331  
1332 which grows more slowly. 1332

1333 To test the abilities of *G. candidum* populations to ex- 1333  
1334 clude other microorganisms by producing volatile com- 1334  
1335 pounds, we set up an experiment with splitted Petri dishes 1335  
1336 where only air can be shared between the two parts. In 1336  
1337 one part of the Petri dish, we spread 75µL of a *G. candidum* 1337  
1338 spore solution (1.10<sup>6</sup> spores/mL) and let it grow during 1338  
1339 24 hours before adding on the other part of the Petri 1339  
1340 dish a drop of 5µL of a competitor spore solution (1.10<sup>6</sup> 1340  
1341 spores/mL). For each competitor, two Petri dishes were 1341  
1342 inoculated without any *G. candidum* as controls. We used 1342  
1343 as competitors the following species and strains: *Penicil-* 1343  
1344 *lium bifforme* (ESE00018, ESE00023, ESE00125, ESE00222, 1344  
1345 ESE00423), *Penicillium roqueforti* (ESE00250, ESE00631, 1345  
1346 ESE00640, ESE00925) and *Scopulariopsis asperula*(ESE00044, 1346  
1347 ESE00102, ESE00835, ESE01287, ESE01324; Table S15). 1347  
1348 Petri dishes were grown at 10°C, measured and pictured 1348  
1349 at 11 days for *P. bifforme* and *P. roqueforti* and 19 days for 1349  
1350 *Scopulariopsis asperula*. 1350

## J.6. Graphics and statistical analyses

Plots and statistical analyses were made using ggplot2 (Wickham, 2016), rstatix and ggpubr packages in the R environment. For ANOVAs, we used standard linear models in which all explanatory variables were discrete, with explained variables being radial growth for growth conditions (for media, temperature, salt content and adaptation to milk experiments), integral opacity score (for opacity experiment), relative proportions of volatiles compounds (for volatile compounds experiment) and radial growth of the competitor (for competition experiments). The explanatory variable common for all analyses was the ‘population’ of *G. candidum*. The variables ‘medium’, ‘day’ and ‘temperature’ were explanatory variables specific to the growth analysis. The ‘competitor species’ variable was specific to competition analyses. 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 honest significant difference (HSD) tests. The data normality of residuals was checked; when residues deviated from normality (only for the opacity experiment), we also ran non-parametric tests (Wilcoxon ranking tests) using R. Radius of lysis for lipolytic and proteolytic activities experiments was often discrete, strains either showing lytic activity or not at all. This is why we decided to transform these data into qualitative discrete data in order to fit a generalized linear model with a binomial function as logit. Growth time (7, 14 and 21 days) and temperature (15 and 25°C) were taken as random variables because no fit could be achieved with little data and we wanted to test for population effect. Tukey contrasts were used to compare population means of populations when population effect was significant.

## 6. ACKNOWLEDGMENTS

We thank everyone who sent cheese crusts and Riwanon Lemee for providing 15 strains from Laboratoires STANDA, Caen, France. We thank Laura Prugneau, Jérémy Raynaud, Thomas Mari, Véronique Bougie and Jules Larouche for their help in the laboratory experiments. This work was funded by the Artifice ANR-19-CE20-0006-01 ANR grant to J.R. In addition, a NSERC-Discovery Grant supported the scholarship of V.P and data generation, and is held by S. Labrie. (RGPIN-2017-06388).

## 7. AUTHOR CONTRIBUTIONS

J.R. designed and supervised the study, and obtained funding. B.B., V.P., J.R. and A.S. generated the data. C.G. provided strains from the CIRM-Levures INRAE collection. B.B., J.-P.V., R.C.d.I.V. and S.O. analyzed the genomes. B.B., S.H., J.R. and A.S. performed the experiments. B.B. analyzed the data from laboratory experiments. M.H.L. and St.L. supervised the lipolysis and proteolysis analyses. B.B., So.L. and A.-C.P. performed the volatile compound experiment. T.G. contributed to interpretation; B.B., T.G.

and J.R. wrote the manuscript, with contributions from all the authors.

## 8. REFERENCES

1. Diamond, J. Evolution, consequences and future of plant and animal domestication. *Nature* 418, 700–707 (2002).
2. Frantz, L. A. F. et al. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat. Genet.* 47, 1141–1148 (2015).
3. Petersen, J. L. et al. Genome-Wide Analysis Reveals Selection for Important Traits in Domestic Horse Breeds. *PLoS Genet.* 9, e1003211 (2013).
4. Warmuth, V. et al. European Domestic Horses Originated in Two Holocene Refugia. *PLoS ONE* 6, e18194 (2011).
5. Hufford, M. B. et al. Comparative population genomics of maize domestication and improvement. *Nat. Genet.* 44, 808–811 (2012).
6. Mabry, M. E. et al. The Evolutionary History of Wild, Domesticated, and Feral Brassica oleracea (Brassicaceae). *Mol. Biol. Evol.* 38, 4419–4434 (2021).
7. Peng, J. H., Sun, D. & Nevo, E. Domestication evolution, genetics and genomics in wheat. *Mol. Breed.* 28, 281 (2011).
8. Liu, X. et al. A single-nucleotide mutation within the TBX3 enhancer increased body size in Chinese horses. *Curr. Biol.* 32, 480–487.e6 (2022).
9. Plassais, J. et al. Natural and human-driven selection of a single non-coding body size variant in ancient and modern canids. *Curr. Biol.* 32, 889–897.e9 (2022).
10. Qanbari, S. et al. Classic Selective Sweeps Revealed by Massive Sequencing in Cattle. *PLOS Genet.* 10, e1004148 (2014).
11. Cornille, A., Giraud, T., Smulders, M. J. M., Roldán-Ruiz, I. & Gladieux, P. The domestication and evolutionary ecology of apples. *Trends Genet.* 30, 57–65 (2014).
12. Purugganan, M. D. Evolutionary Insights into the Nature of Plant Domestication. *Curr. Biol.* 29, R705–R714 (2019).
13. Marsden, C. D. et al. Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proc. Natl. Acad. Sci.* 113, 152–157 (2016).
14. Zhu, Q., Zheng, X., Luo, J., Gaut, B. S. & Ge, S. Multilocus Analysis of Nucleotide Variation of *Oryza sativa* and Its Wild Relatives: Severe Bottleneck during Domestication of Rice. *Mol. Biol. Evol.* 24, 875–888 (2007).
15. Steensels, J., Gallone, B. & Verstrepen, K. J. Interspecific hybridization as a driver of fungal evolution and adaptation. *Nat. Rev. Microbiol.* 1–16 (2021) doi:10.1038/s41579-021-00537-4.
16. Gladieux, P. et al. Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Mol. Ecol.* 23, 753–773 (2014).
17. Bai, F.-Y., Han, D.-Y., Duan, S.-F. & Wang, Q.-M. The Ecology and Evolution of the Baker’s Yeast *Saccharomyces cerevisiae*. *Genes* 13, 230 (2022).
18. Barbosa, R. et al. Multiple Rounds of Artificial Selection Promote Microbe Secondary Domestication—The Case of Cachaça Yeasts. *Genome Biol. Evol.* 10, 1939–1955 (2018).
19. Bigey, F. et al. Evidence for Two Main Domestication Trajectories in *Saccharomyces cerevisiae* Linked to Distinct Bread-Making Processes. *Curr. Biol.* 31, 722–732.e5 (2021).
20. Duan, S.-F. et al. The origin and adaptive evolution of domesticated populations of yeast from Far East Asia. *Nat. Commun.* 9, 2690 (2018).
21. Fay, J. C. & Benavides, J. A. Evidence for Domesticated and Wild Populations of *Saccharomyces cerevisiae*. *PLoS Genet.* 1, e5 (2005).

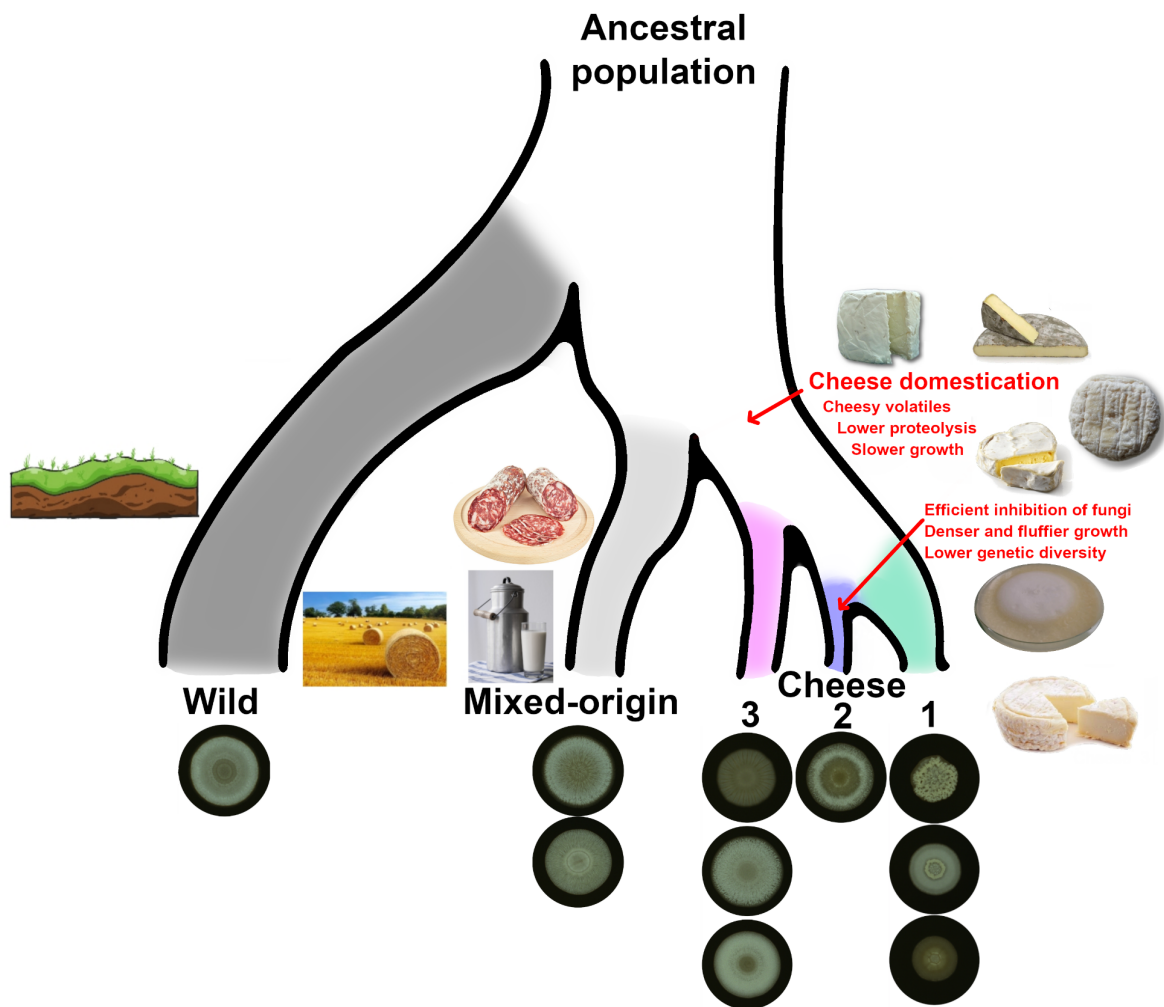
- 1465 22. Gallone, B. et al. Domestication and Divergence of *Saccharomyces* 1530  
1466 *cerevisiae* Beer Yeasts. *Cell* 166, 1397–1410.e16 (2016). 1531
- 1467 23. Gonçalves, M. et al. Distinct Domestication Trajectories in Top- 1532  
1468 Fermenting Beer Yeasts and Wine Yeasts. *Curr. Biol. CB* 26, 1533  
1469 2750–2761 (2016). 1534
- 1470 24. Lahue, C., Madden, A., Dunn, R. & Smukowski Heil, C. History 1535  
1471 and Domestication of *Saccharomyces cerevisiae* in Bread Baking. 1536  
1472 *Front. Genet.* 11, (2020). 1537
- 1473 25. Legras, J.-L. et al. Adaptation of *S. cerevisiae* to Fermented Food 1538  
1474 Environments Reveals Remarkable Genome Plasticity and the Foot- 1539  
1475 prints of Domestication. *Mol. Biol. Evol.* 35, 1712–1727 (2018). 1540
- 1476 26. Libkind, D. et al. Microbe domestication and the identification of 1541  
1477 the wild genetic stock of lager-brewing yeast. *Proc. Natl. Acad. Sci.* 108, 14539–14544 (2011). 1542  
1478 1543
- 1479 27. Liti, G. et al. Population genomics of domestic and wild yeasts. 1544  
1480 *Nature* 458, 337–341 (2009). 1545
- 1481 28. Ludlow, C. L. et al. Independent Origins of Yeast Associated with 1546  
1482 Coffee and Cacao Fermentation. *Curr. Biol.* 26, 965–971 (2016). 1547
- 1483 29. Peter, J. et al. Genome evolution across 1,011 *Saccharomyces cere-* 1548  
1484 *visiae* isolates. *Nature* 556, 339–344 (2018). 1549
- 1485 30. Galagan, J. E. et al. Sequencing of *Aspergillus nidulans* and compar- 1550  
1486 ative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438, 1551  
1487 1105–1115 (2005). 1552
- 1488 31. Gibbons, J. G. et al. The Evolutionary Imprint of Domestication 1553  
1489 on Genome Variation and Function of the Filamentous Fungus 1554  
1490 *Aspergillus oryzae*. *Curr. Biol.* 22, 1403–1409 (2012). 1555
- 1491 32. Machida, M. et al. Genome sequencing and analysis of *Aspergillus* 1555  
1492 *oryzae*. *Nature* 438, 1157–1161 (2005). 1556  
1493 1557
- 1494 33. Ropars, J. et al. Domestication of the Emblematic White Cheese- 1558  
1495 Making Fungus *Penicillium camemberti* and Its Diversification into 1559  
1496 Two Varieties. *Curr. Biol.* (2020) doi:10.1016/j.cub.2020.08.082. 1560
- 1497 34. Cheeseman, K. et al. Multiple recent horizontal transfers of a large 1561  
1498 genomic region in cheese making fungi. *Nat. Commun.* 5, 1–9 1562  
1499 (2014). 1563
- 1500 35. Dumas, E. et al. Independent domestication events in the blue- 1564  
1501 cheese fungus *Penicillium roqueforti*. *Mol. Ecol.* 29, 2639–2660 1565  
1502 (2020). 1566
- 1503 36. Ropars, J. et al. Adaptive Horizontal Gene Transfers between 1567  
1504 Multiple Cheese-Associated Fungi. *Curr. Biol.* 25, 2562–2569 1568  
1505 (2015). 1569
- 1506 37. Bigey, F. et al. Evidence for Two Main Domestication Trajectories in 1570  
1507 *Saccharomyces cerevisiae* Linked to Distinct Bread-Making Processes. 1571  
1508 *Curr. Biol.* 31, 722–732.e5 (2021). 1572
- 1509 38. Caron, T. et al. Strong effect of *Penicillium roqueforti* populations on 1573  
1510 volatile and metabolic compounds responsible for aromas, flavor 1574  
1511 and texture in blue cheeses. *Int. J. Food Microbiol.* 354, 109174 1575  
1512 (2021). 1576
- 1513 39. Almeida, P. et al. A Gondwanan imprint on global diversity and 1577  
1514 domestication of wine and cider yeast *Saccharomyces uvarum*. 1578  
1515 *Nat. Commun.* 5, 4044 (2014). 1579
- 1516 40. Barros Lopes, M., Bellon, J. R., Shirley, N. J. & Ganter, P. F. Evidence 1580  
1517 for multiple interspecific hybridization in *Saccharomyces sensu* 1581  
1518 *stricto* species. *FEMS Yeast Res.* 1, 323–331 (2002). 1582
- 1519 41. Borneman, A. R., Forgan, A. H., Kolouchova, R., Fraser, J. A. 1583  
1520 & Schmidt, S. A. Whole Genome Comparison Reveals High 1584  
1521 Levels of Inbreeding and Strain Redundancy Across the Spec- 1585  
1522 trum of Commercial Wine Strains of *Saccharomyces cerevisiae*. *G3* 1586  
1523 *GenesGenomesGenetics* 6, 957–971 (2016). 1587
- 1524 42. Naumova, E. S., Naumov, G. I., Masneuf-Pomarède, I., Aigle, M. & 1588  
1525 Dubourdieu, D. Molecular genetic study of introgression between 1589  
1526 *Saccharomyces bayanus* and *S. cerevisiae*. *Yeast* 22, 1099–1115 1590  
1527 (2005). 1591
- 1528 43. Novo, M. et al. Eukaryote-to-eukaryote gene transfer events re- 1590  
1529 vealed by the genome sequence of the wine yeast *Saccharomyces* 1591  
1592 *cerevisiae* EC1118. *Proc. Natl. Acad. Sci.* 106, 16333–16338 (2009). 1593
44. Gillot, G. et al. Functional diversity within the *Penicillium roqueforti* 1594  
1595 species. *Int. J. Food Microbiol.* 241, 141–150 (2017). 1596
45. Ropars, J., Caron, T., Lo, Y.-C., Bennetot, B. & Giraud, T. La domes- 1597  
1598 tication des champignons *Penicillium* du fromage. *C. R. Biol.* 343, 1599  
1600 155–176 (2020). 1601
46. Alper, I., Frenette, M. & Labrie, S. Genetic diversity of dairy 1602  
1603 *Geotrichum candidum* strains revealed by multilocus sequence typ- 1604  
1605 ing. *Appl. Microbiol. Biotechnol.* 97, 5907–5920 (2013). 1606
47. Jacques, N., Mallet, S., Laaghouti, F., Tinsley, C. R. & Casaregola, 1607  
1608 S. Specific populations of the yeast *Geotrichum candidum* revealed 1609  
1610 by molecular typing. *Yeast* 34, 165–178 (2017). 1611
48. Perkins, V. et al. Phenotypic and Genetic Characterization of the 1612  
1613 Cheese Ripening Yeast *Geotrichum candidum*. *Front. Microbiol.* 11, 1614  
1615 (2020). 1616
49. Tinsley, C. R. et al. Molecular Genetic Analysis with Microsatellite- 1617  
1618 like Loci Reveals Specific Dairy-Associated and Environmental 1619  
1620 Populations of the Yeast *Geotrichum candidum*. *Microorganisms* 10, 1621  
1622 103 (2022). 1623
50. Boutrou, R. & Gueguen, M. Interests in *Geotrichum candidum* for 1624  
1625 cheese technology. *Int. J. Food Microbiol.* 102, 1–20 (2005). 1626
51. Morel, G. et al. Differential gene retention as an evolutionary 1627  
1628 mechanism to generate biodiversity and adaptation in yeasts. *Sci. 1629  
1630 Rep.* 5, 11571 (2015). 1631
52. Felsenstein, J. The evolutionary advantage of recombination. *Ge- 1632  
1633 netics* 78, 737–756 (1974). 1634
53. Giannakou, K., Cotterrell, M. & Delneri, D. Genomic Adaptation of 1635  
1636 *Saccharomyces* Species to Industrial Environments. *Front. Genet.* 11, 1637  
1638 (2020). 1639
54. Gao, M., Glenn, A. E., Blacutt, A. A. & Gold, S. E. Fungal Lac- 1640  
1641 tamases: Their Occurrence and Function. *Front. Microbiol.* 8, 1642  
1643 (2017). 1644
55. Baduel, P., Quadrana, L., Hunter, B., Bomblies, K. & Colot, V. Re- 1645  
1646 laxed purifying selection in autopolyploids drives transposable 1647  
1648 element over-accumulation which provides variants for local adap- 1649  
1650 tation. *Nat. Commun.* 10, 5818 (2019). 1651
56. McSweeney, P. L. H. Biochemistry of cheese ripening. *Int. J. Dairy 1652  
1653 Technol.* 57, 127–144 (2004). 1654
57. Bonnarne, P. et al. Sulfur compound production by *Geotrichum 1655  
1656 candidum* from l-methionine: importance of the transamination 1657  
1658 step. *FEMS Microbiol. Lett.* 205, 247–252 (2001). 1659
58. Penland, M. et al. Linking Pélardon artisanal goat cheese microbial 1660  
1661 communities to aroma compounds during cheese-making and 1662  
1663 ripening. *Int. J. Food Microbiol.* 345, 109130 (2021). 1664
59. Collins, Y. F., McSweeney, P. L. H. & Wilkinson, M. G. Lipolysis and 1665  
1666 free fatty acid catabolism in cheese: a review of current knowledge. 1667  
1668 *Int. Dairy J.* 13, 841–866 (2003). 1669
60. Dugat-Bony, E. et al. Overview of a Surface-Ripened Cheese Com- 1670  
1671 munity Functioning by Meta-Omics Analyses. *PLOS ONE* 10, 1672  
1673 e0124360 (2015). 1674
61. Kumura, H. et al. Casein Digestion by *Debaryomyces hansenii* Iso- 1675  
1676 lated from Cheese. *Biosci. Biotechnol. Biochem.* 66, 1370–1373 1677  
1678 (2002). 1679
62. Adams, D. J. Fungal cell wall chitinases and glucanases. *Microbiol- 1680  
1681 ogy* 150, 2029–2035 (2004). 1682
63. Pitt, J. I., Cruickshank, R. H. & Leistner, L. *Penicillium commune*, *P.* 1683  
1684 *camemberti*, the origin of white cheese moulds, and the production 1685  
1686 of cyclopiazonic acid. *Food Microbiol.* 3, 363–371 (1986). 1687
64. Gallone, B. et al. Origins, evolution, domestication and diversity 1688  
1689 of *Saccharomyces* beer yeasts. *Curr. Opin. Biotechnol.* 49, 148–155 1690  
1691 (2018). 1692
65. Animal domestication and behavior. (CABI, 2002). 1693  
1694 doi:10.1079/9780851995977.0000. 1695



- 1592 66. Liu, S.-Q., Holland, R. & Crow, V. L. Esters and their biosynthesis 1655  
1593 in fermented dairy products: a review. *Int. Dairy J.* 14, 923–945 1656  
1594 (2004). 1657
- 1595 67. Urbach, G. The flavour of milk and dairy products: II. Cheese: 1658  
1596 contribution of volatile compounds. *Int. J. Dairy Technol.* 50, 1659  
1597 79–89 (1997). 1660
- 1598 68. Curioni, P. M. G. & Bosset, J. O. Key odorants in various cheese 1661  
1599 types as determined by gas chromatography-olfactometry. *Int.* 1662  
1600 *Dairy J.* 12, 959–984 (2002). 1663  
1664
- 1601 69. Mayo, B., Rodríguez, J., Vázquez, L. & Flórez, A. B. Microbial 1665  
1602 Interactions within the Cheese Ecosystem and Their Application 1666  
1603 to Improve Quality and Safety. *Foods* 10, 602 (2021). 1667
- 1604 70. Monnet, C., Landaud, S., Bonnarme, P. & Swennen, D. Growth 1668  
1605 and adaptation of microorganisms on the cheese surface. *FEMS* 1669  
1606 *Microbiol. Lett.* 362, 1–9 (2015). 1670
- 1607 71. Monnet, C., Back, A. & Irlinger, F. Growth of Aerobic Ripening 1671  
1608 Bacteria at the Cheese Surface Is Limited by the Availability of Iron. 1672  
1609 *Appl. Environ. Microbiol.* 78, 3185–3192 (2012). 1673  
1674
- 1610 72. Dieuleveux, V., Van Der Pyl, D., Chataud, J. & Gueguen, M. Purifi- 1675  
1611 cation and characterization of anti-*Listeria* compounds produced 1676  
1612 by *Geotrichum candidum*. *Appl. Environ. Microbiol.* 64, 800–803 1677  
1613 (1998). 1678
- 1614 73. Nielsen, M. S., Frisvad, J. C. & Nielsen, P. V. Colony interaction 1679  
1615 and secondary metabolite production of cheese-related fungi in 1680  
1616 dual culture. *J. Food Prot.* 61, 1023–1029 (1998). 1681
- 1617 74. Omeike, S. O. et al. Purification, De Novo Characterization and 1682  
1618 Antibacterial Properties of a Novel, Narrow-Spectrum Bacterio- 1683  
1619 static Tripeptide from *Geotrichum candidum* OMON-1. *Arab. J. Sci.* 1684  
1620 *Eng.* 46, 5275–5283 (2021). 1685
- 1621 75. von Gastrow, L. et al. Microbial community dispersal from wheat 1686  
1622 grains to sourdoughs: A contribution of participatory research. 1687  
1623 *Mol. Ecol.* (2022) doi:10.1111/mec.16630. 1688
- 1624 76. Lo, Y.-C. et al. Domestication in dry-cured meat *Penicillium* fungi: 1689  
1625 convergent specific phenotypes and horizontal gene transfers 1690  
1626 without strong genetic subdivision. 2022.03.25.485132 Preprint 1691  
1627 at <https://doi.org/10.1101/2022.03.25.485132> (2022). 1692
- 1628 77. Ropars, J. & Giraud, T. Convergence in domesticated fungi used for 1693  
1629 cheese and dry-cured meat maturation: beneficial traits, genomic 1694  
1630 mechanisms, and degeneration. *Curr. Opin. Microbiol.* 70, 102236 1695  
1631 (2022). 1696
- 1632 78. Crequer, E. et al. A new cheese population in *Penicillium roqueforti* 1697  
1633 and adaptation of the five populations to their ecological niche. 1698  
1634 <http://biorxiv.org/lookup/doi/10.1101/2023.01.21.524518> (2023) 1699  
1635 doi:10.1101/2023.01.21.524518. 1700
- 1636 79. Janzen, G. M. & Hufford, M. B. Crop Domestication: A Sneak-Peek 1701  
1637 into the Midpoint of Maize Evolution. *Curr. Biol.* 26, R1240–R1242 1702  
1638 (2016). 1703
- 1639 80. Lin, T. et al. Genomic analyses provide insights into the history of 1704  
1640 tomato breeding. *Nat. Genet.* 46, 1220–1226 (2014). 1705
- 1641 81. Carreira, A. et al. Influence of selected factors on browning of 1706  
1642 Camembert cheese. *J. Dairy Res.* 69, 281–292 (2002). 1707
- 1643 82. Alberto, F. J. et al. Convergent genomic signatures of domestication 1708  
1644 in sheep and goats. *Nat. Commun.* 9, 813 (2018). 1709
- 1645 83. Cresko, W. A. et al. Parallel genetic basis for repeated evolution 1710  
1646 of armor loss in Alaskan threespine stickleback populations. *Proc.* 1711  
1647 *Natl. Acad. Sci.* 101, 6050–6055 (2004). 1712
- 1648 84. Dyer, A. G. et al. Parallel evolution of angiosperm colour signals: 1713  
1649 common evolutionary pressures linked to hymenopteran vision. 1714  
1650 *Proc. R. Soc. B Biol. Sci.* 279, 3606–3615 (2012). 1715
- 1651 85. Elmer, K. R. et al. Parallel evolution of Nicaraguan crater lake 1716  
1652 cichlid fishes via non-parallel routes. *Nat. Commun.* 5, 1–8 (2014). 1717
- 1653 86. Lin, Z. et al. Parallel domestication of the *Shattering1* genes in 1718  
1654 cereals. *Nat. Genet.* 44, 720–724 (2012). 1719
87. Macías, L. G. et al. Convergent adaptation of *Saccharomyces* 1720  
1721 *uvarum* to sulfite, an antimicrobial preservative widely used in 1722  
1723 human-driven fermentations. *PLOS Genet.* 17, e1009872 (2021). 1724
88. Oin, K. E., Hofmann, C. M., Hofmann, H. A. & Carleton, K. L. 1725  
1726 Parallel Evolution of Opsin Gene Expression in African Cichlid 1727  
1728 Fishes. *Mol. l. Evol.* 27, 2839–2854 (2010). 1729
89. Thorpe, R. S., Barlow, A., Malhotra, A. & Surget-Groba, Y. 1730  
1731 Widespread parallel population adaptation to climate variation 1732  
1733 across a radian: implications for adaptation to climate change. *Mol.* 1734  
1735 *Ecol.* 24, 1019–1030 (2015). 1736
90. Vilgalys, R. & Hester, M. Rapid genetic identification and mapping 1737  
1738 of enzymatically amplified ribosomal DNA from several *Crypto-* 1739  
1740 *coccus* species. *J. Bacteriol.* 172, 4238–4246 (19 1741
91. 1742
92. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible 1743  
1744 trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120 1745  
1746 (2014). 1747
93. Koren, S. et al. Canu: scalable and accurate long-read assembly 1748  
1749 via adaptive k-mer weighting and repeat separation. *Genome Res.* 1750  
1751 27, 722–736 (2017). 1752
94. Salmela, L. & Rivals, E. LoRDEC: accurate and efficient long read 1753  
1754 error correction. *Bioinformatics* 30, 3506–3514 ( 1755  
1756 95. 4). 1757
96. Walker, B. J. et al. Pilon: An Integrated Tool for Comprehensive 1758  
1759 Microbial Variant Detection and Genome Assembly Improvement. 1760  
1761 *PLoS ONE* 9, e112963 (2 1762  
1763 97. ). 1764
98. Antipov, D., Korobeynikov, A., McLean, J. S. & Pevzner, P. A. 1765  
1766 hybridSPAd 1767  
1768 an algorithm for hybrid assembly of short and long reads. *Bioin-* 1769  
1770 *formatics* 32, 1009–1015 (2016). 1771
100. Pribelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynik 1772  
1773 101. A. Using SPAdes De Novo Assembler. *Curr. Protoc. Bioinforma.* 1774  
1775 70, e102 (2020). 1776
102. Cruseo 1777
103. . R. et al. The khmer software package: enabling efficient nu- 1778  
1779 cleotide sequence analysis. *F1000Research* 4, 900 (2015). 1780
104. Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. Using native 1781  
1782 and syntenic mapped cDNA alignments to improve de novo gene 1783  
1784 finding. *Bioinformatics* 24, 637–644 (2008). 1785
105. Palmer, J. M. & Stajich, J. Funannotate v1.8.1: Eukaryotic genome 1786  
1787 annotation. (2020) doi:10.5281/zenodo.405. 1788
106. Blum, M. et al. The InterPro protein families and domains database: 1789  
1790 20 years on. *Nucleic Acids Res.* 49, D344–D354 (2021). 1791
107. Eddy, S. R. Accelerated Profile HMM Searches. *PLoS Comput. Biol.* 1792  
1793 7, e1002195 (2011). 1794
108. . Huang, L. et al. dbCAN-seq: a database of carbohydrate-active 1795  
1796 enzyme (CAZyme) sequence and annotation. *Nucleic Acids Res.* 1797  
1798 46, D516–D521 (2018). 1799
109. Mistry, J. et al. Pfam: The protein families database in 2021. *Nucleic* 1800  
1801 *Acids Res.* 49, D412–D419 (20 1802
110. . 1803
111. Rawlings, N. D. et al. The MEROPS database of proteolytic en- 1804  
1805 zymes, their substrates and inhibitors in 2017 and a comparison 1806  
1807 with peptidases in the PANTHER database. *Nucleic Acids Res.* 46, 1808  
1809 D624–D632 (2018). 1810
112. The UniProt Consortium et al. UniProt: the universal protein 1811  
1812 knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489 (2 1813  
1814 113. ). 1815
114. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & 1816  
1817 Huerta-Cepas, J. eggNOG-mapper v2: Functional Annotationol- 1818  
1819 ogy Assignments, and Domain Prediction at the Metagenomic 1819  
1820 Scale I. *Biol. Evol.* 38, 5825–5829 (2021). 1821

- 1718 115. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally 1780  
1719 and phylogenetically annotated orthology resource based on 5090 1781  
1720 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314 1782  
1721 (2019). 1783
- 1722 116. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with 1784  
1723 Bowtie 2. *Nat. Methods* 9, 357–359 (2012). 1785
- 1724 117. Li, H. et al. The Sequence Alignment/Map format and SAMtools. 1786  
1725 *Bioinformatics* 25, 2078–2079 (2009). 1787
- 1726 118. Minh, B. Q. et al. IQ-TREE 2: New Models and Efficient Methods 1788  
1727 for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37, 1789  
1728 1530–1534 (2020). 1790
- 1729 119. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. Ultrafast Approx-  
1730 imation for Phylogenetic Bootstrap. *Mol. Biol. Evol.* 30, 1188–1195  
1731 (2013).
- 1732 120. Huson, D. H. & Bryant, D. Application of Phylogenetic Networks  
1733 in Evolutionary Studies. *Mol. Biol. Evol.* 23, 254–267 (2006).
- 1734 121. Bougeard, S. & Dray, S. Supervised Multiblock Analysis in R with  
1735 the ade4 Package. *J. Stat. Softw.* 86, 1–17 (2018).
- 1736 122. Chessel, D., Dufour, A.-B. & Thioulouse, J. The ade4 Package – I:  
1737 One-Table Methods. *R News* 4, 5–10 (2004).
- 1738 123. Dray, S., Dufour, A.-B. & Chessel, D. The ade4 Package – II: Two-  
1739 Table and K-Table Methods. *R News* 7, 47–52 (2007).
- 1740 124. Dray, S. & Dufour, A.-B. The ade4 Package: Implementing the  
1741 Duality Diagram for Ecologists. *J. Stat. Softw.* 22, 1–20 (2007).
- 1742 125. Thioulouse, J. et al. Multivariate Analysis of Ecological Data with  
1743 ade4. (Springer, 2018). doi:10.1007/978-1-4939-8850-1.
- 1744 126. Jørsboe, E., Hanghøj, K. & Albrechtsen, A. fastNGSadmix: admix-  
1745 ture proportions and principal component analysis of a single NGS  
1746 sample. *Bioinformatics* 33, 3148–3150 (2017).
- 1747 127. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analy-  
1748 sis of Next Generation Sequencing Data. *BMC Bioinformatics* 15,  
1749 356 (2014).
- 1750 128. Hudson, R. R., Slatkin, M. & Maddison, W. P. Estimation of levels  
1751 of gene flow from DNA sequence data. *Genetics* 132, 583–589  
1752 (1992).
- 1753 129. Nei, M. & Li, W. H. Mathematical model for studying genetic  
1754 variation in terms of restriction endonucleases. *Proc. Natl. Acad.*  
1755 *Sci. U. S. A.* 76, 5269–5273 (1979).
- 1756 130. Watterson, G. A. On the number of segregating sites in genetical  
1757 models without recombination. *Theor. Popul. Biol.* 7, 256–276  
1758 (1975).
- 1759 131. Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E. & Lercher, M.  
1760 J. PopGenome: An Efficient Swiss Army Knife for Population  
1761 Genomic Analyses in R. *Mol. Biol. Evol.* 31, 1929–1936 (2014).
- 1762 132. Cingolani, P. et al. A program for annotating and predicting the  
1763 effects of single nucleotide polymorphisms, SnpEff: SNPs in the  
1764 genome of *Drosophila melanogaster* strain w 1118 ; iso-2; iso-3. *Fly*  
1765 (Austin) 6, 80–92 (2012).
- 1766 133. Flynn, J. M. et al. RepeatModeler2 for automated genomic discov-  
1767 ery of transposable element families. *Proc. Natl. Acad. Sci. U. S.*  
1768 *A.* 117, 9451–9457 (2020).
- 1769 134. Goubert, C. et al. A beginner’s guide to manual curation of trans-  
1770 posable elements. *Mob. DNA* 13, 7 (2022).
- 1771 135. Li, H. Aligning sequence reads, clone sequences and assembly  
1772 contigs with BWA-MEM. <http://arxiv.org/abs/1303.3997> (2013)  
1773 doi:10.48550/arXiv.1303.3997.
- 1774 136. Pavlidis, P., Živkovic, D., Stamatakis, A. & Alachiotis, N. SweeD:  
1775 likelihood-based detection of selective sweeps in thousands of  
1776 genomes. *Mol. Biol. Evol.* 30, 2224–2234 (2013).
- 1777 137. Ekseth, O. K., Kuiper, M. & Mironov, V. orthAgogue: an agile tool  
1778 for the rapid prediction of orthology relations. *Bioinformatics* 30,  
1779 734–736 (2014).
138. Ranwez, V., Douzery, E. J. P., Cambon, C., Chantret, N. & Delsuc,  
F. MACSE v2: Toolkit for the Alignment of Coding Sequences  
Accounting for Frameshifts and Stop Codons. *Mol. Biol. Evol.* 35,  
2582–2584 (2018).
139. Kritikos, G. et al. A tool named Iris for versatile high-throughput  
phenotyping in microorganisms. 24 (2017).
140. Wickham, H. Getting Started with ggplot2. in *ggplot2: Elegant  
Graphics for Data Analysis* (ed. Wickham, H.) 11–31 (Springer  
International Publishing, 2016). doi:10.1007/978-3-319-24277-4\_2.
141. Harlan, J. R. et al. Biodiversity in agriculture: domestication, evo-  
lution, and sustainability. (Cambridge University Press, 2012).

### III. Synthèse de la domestication de *Geotrichum candidum*



**Figure 18 : Histoire de la domestication de *G. candidum*.** La domestication de *G. candidum* s'est déroulée en plusieurs étapes successives : la divergence entre la population sauvage et celle d'origine mixte (foin, sol et produits alimentaires) puis la divergence entre la population d'origine mixte et le clade contenant les souches isolées de fromage. Le clade contenant les souches du fromage se subdivise en trois populations génétiques: Cheese\_1, Cheese\_2 et Cheese\_3. La population Cheese\_2 est la population du fromage ayant la plus faible diversité génétique, indiquant un goulet d'étranglement plus fort. Les boîtes de Petri présentent l'aspect morphologique de chaque population sur un milieu de gélose d'agar contenant du fromage salé après 14 jours de croissance à 10°C. Lorsque plusieurs morphologies sont observables dans la population, plusieurs photos sont disposées verticalement pour rendre compte de cette diversité morphologique. Les populations du fromage ont une croissance radiale plus faible, plus de composés volatiles souhaitables et une plus faible protéolyse comparé aux autres populations. Plus particulièrement, la population Cheese\_2 est caractérisée par une croissance plus dense et une meilleure inhibition des contaminants.

L'étude de la génomique des populations de *G. candidum* nous fournit une histoire différente de celle observée chez *P. camemberti*. Tout comme chez *P. camemberti* on retrouve une domestication en deux étapes, la première marquant l'invasion de la niche agro-alimentaire puis une seconde étape menant à une population spécifique de l'environnement fromager (Figure 18). De même, trois populations du fromage aux phénotypes contrastés ont divergé. En revanche, ces populations du fromage sont nettement plus diversifiées génétiquement que *P. camemberti*. Les goulets d'étranglements sont bien moins intenses. *Penicillium camemberti* a été sélectionné puis propagé à large échelle pour inoculer les fromages de type Brie et Camembert, alors que *G. candidum* est présent naturellement dans le lait. Étonnement, certains phénotypes convergents ont été identifiés entre *G. candidum* et *P. camemberti* tel que la densité du mycélium et les capacités d'exclusion des compétiteurs. Nous discuterons donc ces différents points dans la partie Discussion. *Geotrichum candidum* est un nouveau modèle, il reste beaucoup à faire maintenant que la génomique des populations a été explorée. Les souches du fromage proviennent principalement de la France et des États-Unis d'Amérique. Peut-être qu'en accentuant l'effort d'échantillonnage dans d'autres pays, il serait possible d'identifier de nouvelles populations. Chez *S. cerevisiae*, l'échantillonnage en Afrique et en Asie a permis de découvrir des populations naturelles plus variées (Lee et al. 2022). Les analyses de génomique des populations et de phénotypage de *G. candidum* ont été effectués sur des souches dont l'appartenance à une des cinq populations

identifiées était certaine. Or, parmi les souches de *G. candidum*, on retrouve des souches qui ne correspondent à aucune des cinq populations définies. Ces souches sont sûrement introgressées et ont été mises de côté alors qu'elles pourraient correspondre à une population avec un échantillonnage plus conséquent. Analyser ces souches permettrait d'approfondir l'histoire de la domestication de *G. candidum*. Ce point sera abordé dans la partie perspective.

# Discussion générale

## I. Les conséquences de la domestication des champignons du fromage

### A) Des niveaux de diversité génétique variés entre les espèces domestiquées

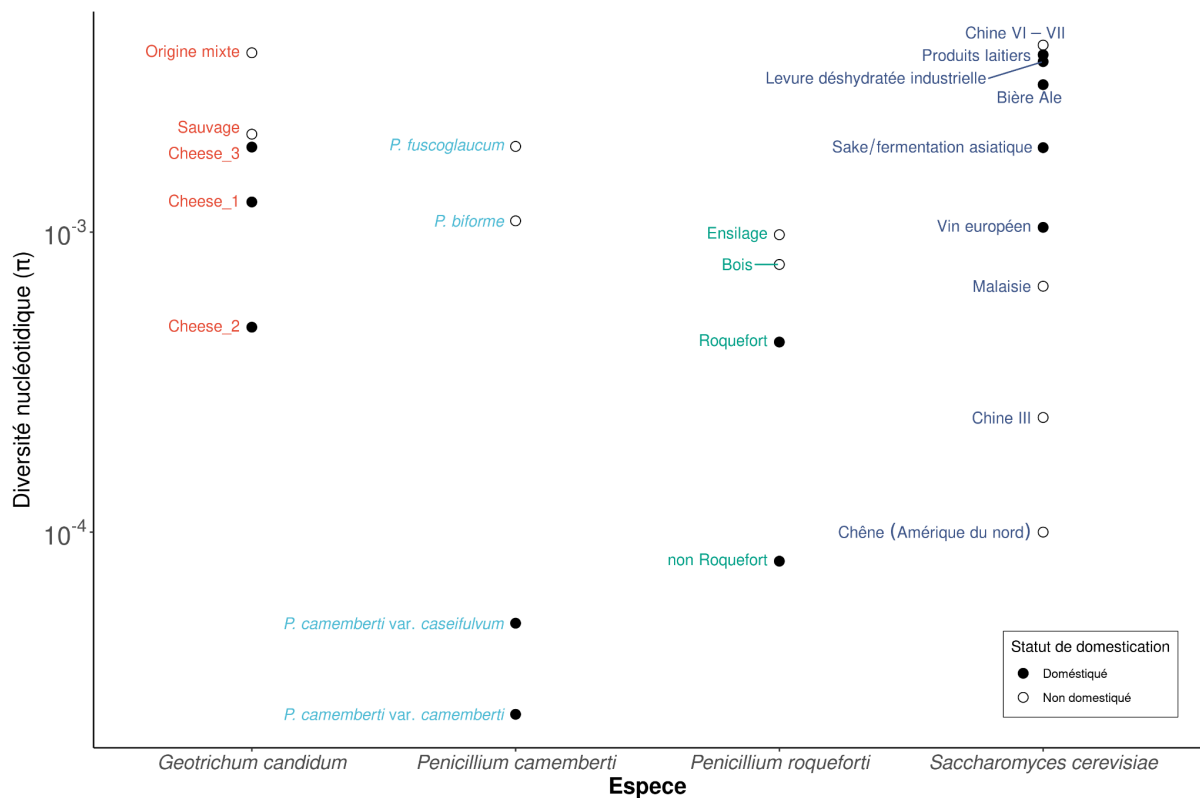
Dans le cadre de la domestication, la sélection forte pour des caractères adaptatifs entraîne une réduction de la diversité.

Les populations domestiquées de *P. camemberti* et *P. roqueforti* ont subi de forts goulets d'étranglement. Une unique souche clonale a été sélectionnée chez *P. camemberti* et cela semble être le cas aussi pour la population Roquefort de *P. roqueforti*. La sélection intense des populations du fromage a aussi mené à une baisse de la fertilité (Ropars, Didiot, et al. 2020; Ropars et al. 2016). Le brassage génétique aurait pu compenser la perte de diversité.

Chez les populations du fromage de *G. candidum* et *P. biforme*, on observe une diversité génétique bien plus élevée que chez *P. camemberti* et *P. roqueforti* (figure 19) (Dumas et al. 2020; Ropars, Didiot, et al. 2020). Ces populations partagent plusieurs caractéristiques. Les populations Cheese\_1 et Cheese\_3 de *G. candidum* et *P. biforme* sont spécifiques de l'environnement fromager. Les adaptations au milieu anthropique de ces populations sont moins marquées. Le brassage génétique inter et intra-populationnel est encore efficace au sein de ces populations, ce qui permet de conserver la diversité génétique. En effet, plusieurs indices vont en ce sens : la présence des deux types sexuels dans les populations, des traces de recombinaison dans les génomes et le nombre élevé de mutations partagées avec les populations proches de l'environnement fromager (38% de mutations partagées entre Cheese\_1 et Cheese\_3 contre 8-9% entre Cheese\_2 et Cheese\_1/Cheese\_3 ;

*P. biforme* partage 7 et 14 % de ses mutations avec *P. camemberti* et *P. fuscoglaucum* respectivement, contre au 4% entre *P. fuscoglaucum* et *P. camemberti*).

De plus, le fait que les populations du fromage de *G. candidum* et *P. biforme* soient présentes naturellement dans le lait cru ou dans les caves d'affinage, même si certaines d'entre elles sont inoculées, évite l'utilisation d'une souche unique pour la fabrication des fromages, contrairement à *P. camemberti* et la population non-Roquefort de *P. roqueforti*.



**Figure 19 : Diversité nucléotidique des populations de *G. candidum*, *P. camemberti*, *P. roqueforti* et *S. cerevisiae*.** Source : (Bai et al. 2022; Dumas et al. 2020; Ropars, Didiot, et al. 2020)

Le cas des populations de *Saccharomyces* est particulier car la diversité des populations domestiquées est similaire à celle des populations naturelles. Lors de leur domestication, *S. cerevisiae* a subi des hybridations inter- et intra-spécifique, ce



qui a un impact non négligeable sur la diversité génétique (Duan et al. 2018; Scopel et al. 2021). De plus, les populations naturelles de *S. cerevisiae* ont été peu échantillonnées, nous n'avons donc qu'une sous-estimation de la diversité génétique réelle (Lee et al. 2022; Bai et al. 2022).

## B) Une diversité phénotypique des souches qui reflète plusieurs usages

La domestication peut mener à la différenciation en plusieurs lignées domestiquées adaptées chacune à des usages spécifiques. Par exemple, dans le cas de la domestication du chien, diverses lignées ont été créées en fonction des besoins (chiens de bergers, d'agrément ou de chasse) (Wayne et vonHoldt 2012). Chez les plantes, la domestication des Brassicaceae a mené à plusieurs espèces dont les parties comestibles sont différentes (chou-fleur, choux de bruxelles, chou kale, etc) (Mabry et al. 2021)

Chez *S. cerevisiae* (vin, bière, pain ou fromage) et *B. bruxellensis* (vin, bière, kombucha, téquila), les populations correspondent à différents aliments et boissons (Bai et al. 2022; Duan et al. 2018; Lahue et al. 2020; Giannakou, Cotterrell, et Delneri 2020; Harrouard et al. 2022; Eberlein et al. 2021). Les adaptations liées au différents usages peuvent impliquer des gènes différents. Par exemple, chez *B. bruxellensis* le gène de la vinyl phenol réductase a été sélectionné dans les populations du vin (Crauwels et al. 2017) alors que ce sont les gènes du métabolisme du maltose qui sont sélectionnés dans les populations de la bière (Eberlein et al. 2021; Bai et al. 2022).

Chez *P. camemberti*, les variétés *caseifulvum* et *camemberti* sont utilisées pour des fromages différents (rigotte de Condrieu et st marcelin contre camembert et brie). Il

en est de même pour *P. biforme* que l'on retrouve fréquemment sur les fromages à base de laits de chèvre.

Chez *G. candidum*, on retrouve trois populations de l'environnement fromager qui pourraient potentiellement être adaptés à différents fromages. D'autant plus, que la majorité des souches isolées de fromage de chèvre sont dans la population Cheese\_1. Cependant, les souches de la population Cheese\_1 ne sont pas spécifiques des fromages de chèvre. Nos résultats n'ont pas permis de savoir si les trois populations de *G. candidum* correspondaient à différents types de fromage, ni à différentes origines géographiques.

### C) La population Cheese\_2 de *G. candidum* : une trajectoire évolutive similaire à *P. camemberti*

La population Cheese\_2 est la moins diversifiée des populations de *G. candidum*. Elle est caractérisée par un mycélium plus dense et une inhibition plus efficace des compétiteurs. Cette population ne possède plus qu'une seule version du type sexuel et son déséquilibre de liaison ne correspond pas à celui d'une population dont le brassage génétique est encore actif. De plus, seuls 9% des mutations sont partagées avec les populations Cheese\_1 et Cheese\_3, là où ces deux dernières partagent entre elles 38% de leurs mutations. La population Cheese\_2 est donc dans une dynamique d'isolation génétique et de propagation clonale. Cette trajectoire évolutive d'une population fortement sélectionnée pour des caractères spécifiques, propagées clonalement, isolées génétiquement, rappelle le cas de *P. camemberti*. Cependant, la population Cheese\_2 possède encore une diversité génétique dix fois supérieure à celle de *P. camemberti*, et similaire à la population Roquefort de *P. roqueforti* (figure 19). La structure de population de *P. roqueforti*

n'est pas liée à un type de fromage en particulier. C'est l'AOP Roquefort qui a permis la protection de la diversité de la population Roquefort. La population Cheese\_2 n'est pas associée à une situation géographique particulière. Sans effort de conservation de la diversité génétique, il est probable que cette diversité soit perdue au cours du temps. La préservation de cette diversité peut permettre d'éviter la dégénérescence des souches et la conservation du potentiel adaptatif à des environnements changeant sur le long terme. Pour cela il faudrait déterminer les pratiques qui favorisent la conservation de la diversité naturelle (souche présente initialement dans le lait).

## II. Processus génétiques impliqués dans l'adaptation

### A) Transferts horizontaux

Lors de leur domestication, les *Penicillium* du fromage ont acquis indépendamment par transfert horizontal les régions *CheesyTer* et *Wallaby*. Ces régions contiennent des gènes en lien avec l'adaptation au fromage car ils sont impliqués dans le métabolisme du lactose et l'inhibition de contaminants (Ropars et al. 2015). On retrouve aussi des transferts horizontaux chez *Saccharomyces* où la population domestiquée originaire du lait possède des régions provenant de *Torulaspota delbrueckii* (Duan et al. 2018).

Chez *Geotrichum candidum*, à partir des analyses de variations de copies, de grandes régions présentes uniquement chez certaines populations ont été mises en évidence. Cependant il faudra des analyses plus approfondies de chacune d'entre elles pour savoir si elles sont source d'adaptation.

### B) Réarrangements chromosomiques

Chez *Saccharomyces*, les réarrangements chromosomiques sont des mécanismes adaptatifs majeurs. Les souches utilisées pour la fermentation de bière lager sont des hybrides (Monerawela et Bond 2018; Rainieri et al. 2006; Libkind et al. 2011) tandis que les souches de *S. cerevisiae* utilisées pour le pain et la bière ont subi des aneuploidies voire des duplications du génome entier (Scopel et al. 2021; Gallone et al. 2018). Des analyses sont actuellement menées au sein de l'équipe pour tester la présence de réarrangements chromosomiques et, si oui, analyser leurs implications dans l'adaptation des souches domestiquées de *G. candidum* et *P. camemberti*. Chez *B. bruxellensis*, les populations domestiquées sont caractérisées par des

changements de ploïdie. Cependant, la ploïdie de *G. candidum* est majoritairement haploïde aussi bien au sein des populations domestiquées que sauvages.

### C) Variations du nombre de copie de gènes

Chez *Saccharomyces* et *B. bruxellensis*, les gènes du métabolisme impliqués dans l'adaptation à la nouvelle niches écologique sont souvent dupliqués voire fusionnés (Bai et al. 2022; Giannakou, Cotterrell, et Delneri 2020; Eberlein et al. 2021). Ces adaptations jouent sur la quantité et l'efficacité de ces enzymes. Chez *Kluyveromyces lactis* par exemple, la duplication de gènes a mené à une néofonctionnalisation du transporteur de cellobiose à lactose, utile dans la nouvelle niche écologique (Varela et al. 2019). Chez *P. camemberti* et *G. candidum*, nos analyses n'ont pas permis de mettre en évidence de variations du nombre de copies de gènes ayant un rôle adaptatif. Cependant, il est possible que certains gènes dupliqués dont la fonction n'est pas connue soient adaptatifs dans l'environnement fromager.

Chez *G. candidum*, une des copies d'une bêta-lactamase dupliquée en tandem a été perdue chez toutes les souches de l'environnement fromager. Les bêta-lactamases sont des enzymes pouvant dégrader des antibiotiques de la famille des bêta lactamines. Les organismes produisant des bêta-lactamases peuvent croître dans un milieu où d'autres espèces fongiques produisent ce type d'antibiotique. Cette perte pourrait être due à un relâchement de la pression de sélection dans le milieu anthropique.

### D) Mutations ponctuelles

Différentes mutations ponctuelles peuvent être impliquées dans l'adaptation en changeant la séquence protéique ou en modifiant la transcription de gènes

importants. Chez *S. cerevisiae*, une mutation dans un gène codant pour un transporteur de maltotriose engendre un gain d'efficacité de cette protéine dans la population utilisée pour la bière (Gallone et al. 2018). Nous avons mis en évidence qu'une mutation dans un gène du cluster de la biosynthèse de l'acide cyclopiazonique rendait la toxine non fonctionnelle chez *P. camemberti* var. *caseifulvum*. Dans la population non-Roquefort l'acide mycophénolique n'est plus produit à cause d'une délétion de 174 pb dans le gène *mpaC* (Gillot et al. 2017). Chez *G. candidum*, des mutations entraînent des codons stop dans certains gènes mais aucun d'entre eux n'a pu être mis en relation avec une adaptation au fromage.

### III) Phénotypes sélectionnés pour la production des fromages

#### A) Phénotypes acquis au cours de la domestication

En étudiant la domestication des champignons du fromage, plusieurs convergences phénotypiques ont été mises en évidence. *Penicillium roqueforti* et *P. camemberti* ont tous les deux subi une sélection sur la couleur et la perte de mycotoxines. *Geotrichum candidum* est connu pour produire une mycotoxine, l'acide Phényllactique, dont l'activité antimicrobienne a été étudiée (Dieuleveux et al. 1998; Kawtharani et al. 2020). Cependant, nous ne connaissons pas les gènes de synthèse de l'acide Phényllactique. Je n'ai donc pas pu vérifier si cette mycotoxine est produite dans certaines populations spécifiquement. *Penicillium camemberti*, *P. roqueforti* et une des populations de *G. candidum* ont été sélectionnés pour un mycélium plus dense et une inhibition plus efficace des contaminants (Ropars, Didiot, et al. 2020; Ropars et al. 2015). A partir des expériences de compétition de *G. candidum*, on sait que l'inhibition se fait soit par des molécules diffusives (peptides, protéines ou acides organiques) soit par occupation du milieu (avec un mycélium plus dense). Il serait donc intéressant de réaliser des études protéomiques pour identifier les protéines sécrétées lors de l'interaction et ainsi déterminer si les protéines sont les mêmes chez *P. camemberti*, *P. roqueforti* et *G. candidum*. De même, on pourrait détecter les acides organiques via une approche de chromatographie en phase liquide à haute performance et spectrométrie de masse (HPLC-MS).

## B) Caractéristiques facilitant l'adaptation à la niche écologique du fromage

Il est possible d'établir une liste des espèces fongiques que l'on trouve sur les fromages : les levures *Geotrichum candidum*, *Debaryomyces hansenii*, *Saccharomyces cerevisiae* et *Kluyveromyces lactis* ainsi que les champignons filamenteux des genres *Mucor*, *Scopulariopsis*, *Fusarium*, *Sporendonema* et *Penicillium*. Ces espèces sont retrouvées dans l'environnement fromager à l'échelle mondiale. On peut se questionner sur ce qui fait que ces champignons sont systématiquement retrouvés dans l'environnement fromager. Chez certaines de ces espèces, il existe des populations domestiquées mais il y a dû y avoir des caractères qui ont permis l'adaptation de la niche écologique du fromage. Cela rappelle le cas de la domestication du chien : les loups ont d'abord été en contact proche avec l'être humain avant même que la domestication ne se soit mise en place (Range et Marshall-Pescini 2022; Galeta et al. 2021).

Dans le cas des champignons du fromage, cela pourrait être une proximité géographique avec une niche écologique naturelle proche des animaux allaitants (sur la peau ou la fourrure des brebis/vaches/chèvres, l'herbe, le sol, les déjections etc), de l'être humain fabriquant le fromage voire des caves d'affinage. Une autre possibilité serait que ces champignons soient pré-adaptés aux conditions de la niche écologique du fromage (fermentation lactique, tolérance au sel, capacités enzymatiques). Pour comprendre ce qui différencie ces espèces domestiquées des autres champignons, il faut s'intéresser aux niches écologiques naturelles d'origine des populations aujourd'hui domestiquées. Pour ce faire, il faut accentuer l'effort d'échantillonnage des espèces proches provenant du milieu naturel (*P.*



*fuscoglaucum*, la population sauvage de *G. candidum* et les populations ensilage et bois de *P. roqueforti*).

## Conclusion & Perspectives

En comparant les souches du milieu naturel et du milieu fromage via des approches génomiques et phénotypiques, nous avons pu mettre en évidence leur différenciation. Alors que certaines populations domestiquées ont subi des goulots d'étranglement drastiques (*P. camemberti* et *P. roqueforti*) avec la sélection d'un individu maintenant repiqué par les industriels, d'autres possèdent encore une grande diversité génétique (*P. biforme*, *Saccharomyces* et les populations du fromage de *G.candidum*).

Bien qu'il existe une grande diversité de moyens d'adaptations, on constate un certain nombre de convergences évolutives au sein des champignons du fromage. En effet, *P. camemberti* et *P. roqueforti* possèdent des régions génomiques transférées horizontalement en commun et ont été sélectionnés pour des phénotypes similaires (couleur, aspect des colonies et perte de certaines mycotoxines). Une des populations du fromage de *G. candidum* montre des traces d'un état plus avancé de domestication. Elle est caractérisée par une diversité génétique moins importante que les autres populations, un mycélium plus dense, et une capacité à inhiber plus efficacement des contaminants, ce qui est similaire à *P. camemberti*.

En plus de l'intérêt fondamental de mes recherches, la connaissance démographique de *P. camemberti* et *G. candidum* a des intérêts appliqués. Les industriels vont pouvoir adapter leurs catalogues de souches à la réalité biologique mais surtout être sensibilisés à la diversité génétique et phénotypique présente chez *G. candidum*. Il faut dès maintenant conserver la biodiversité des différentes

populations de *G. candidum* pour ne pas se retrouver dans l'impasse où est *P. camemberti* dont l'usage est limité par la diversité génétique. Cela représente un défi, car cela demande de s'émanciper de la vision fonctionnelle où l'on cherche à obtenir la souche parfaite. *Geotrichum candidum* étant naturellement présent dans le lait, on peut facilement préserver sa diversité en évitant de stériliser le lait. Les méthodes de fabrication du fromage pourraient aussi avoir un impact sur la présence de ces souches non inoculées. Des études liant sociologie et analyse de diversité seraient donc intéressantes pour mieux comprendre comment les pratiques de fabrication influencent la persistance de ces souches de *G. candidum* mais aussi d'autres espèces fongiques.

En perspectives :

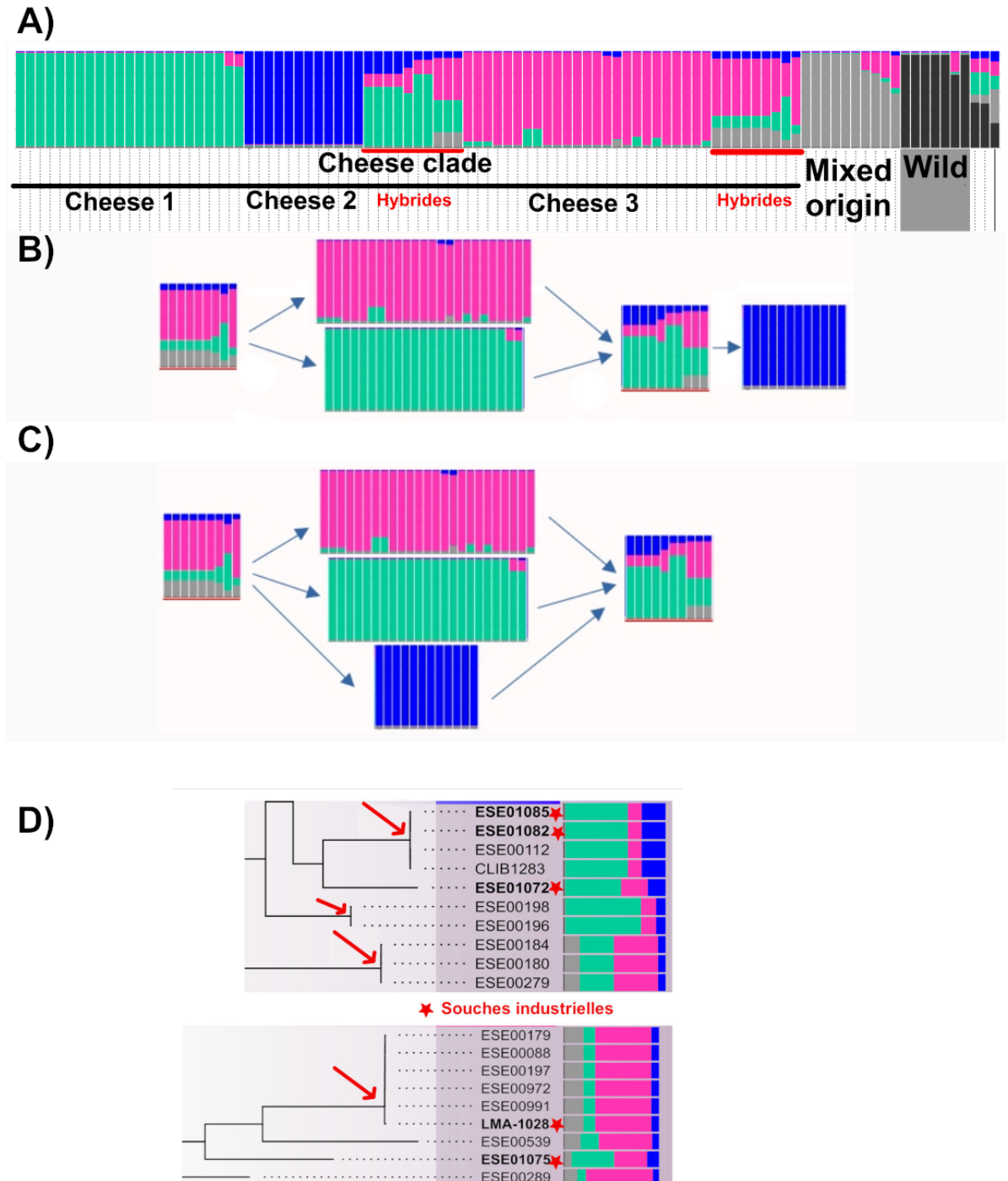
1. Parmi les souches de *G. candidum*, on retrouve qui ne font pas partie d'une des cinq populations de *G. candidum*. Ces souches sont le résultat d'hybridation intra ou interspécifique. Deux groupes d'hybrides se situent au sein du groupe monophylétique formé par toutes les souches du fromage de *G. candidum* (Figure 20A). Un premier groupe d'hybrides se situe à la racine des populations du fromage alors que le second groupe d'hybrides est proche génétiquement de la population Cheese\_2. On peut proposer deux scénarios pour expliquer ces deux groupes d'hybrides :
  - Scénario 1 (figure 20B) : les populations Cheese\_1 et Cheese\_3 sont issues de la différenciation du premier groupe d'hybrides puis par un second évènement d'hybridation, le second groupe d'hybrides est formé. La population Cheese\_2 pourrait être issue de la différenciation

à partir de ce second groupe d'hybrides, expliquant sa proximité génétique.

- Scénario 2 (figure 20C) : Le premier groupe d'hybrides aurait mené par différenciation aux trois populations du fromage et par des réhybridations, le second groupe d'hybrides aurait été formé.

A partir d'approches de coalescence telles que l'Approximate Bayesian Clustering (ABC), il est possible de connaître la plausibilité de ces différents scénarios. Ainsi, on pourrait approfondir l'histoire de la domestication de *G. candidum*.

Parmi les hybrides, on retrouve de nombreuses souches inoculées par les industriels de l'agro-alimentaire (Figure 20D). Les industriels ont pu sélectionner ces souches car elles possèdent des phénotypes uniques ou atypiques du fait qu'elles sont issues d'hybridation. Cette piste pourra être explorée par le phénotypage des hybrides.



**Figure 20 : Origine des souches introgressées de *G. candidum* et potentiel intérêt de celles-ci pour l'agro-alimentaire.** (A) Analyse d'admixture attribuant les variations génétiques des souches de *G. candidum* parmi cinq populations. L'appartenance aux cinq populations définies est ensuite inféré pour chaque souche. Quand une souche a une inférence ambiguë, alors la barre verticale correspondant à la souche est coloré proportionnellement à l'inférence de chaque population. Les souches dont l'inférence ne correspond pas à une unique population de *G. candidum* sont indiquées avec une barre rouge et la mention « hybrides ». Deux clusters génétiques de souches hybrides se distinguent, le premier est phylogénétiquement proche de la population Cheese\_2 alors que le second se situe à l'embranchement des populations du fromage. Plusieurs scénarios démographiques peuvent être avancés : (A) Les hybrides à l'embranchement des

populations du fromage ont donnés par divergence la population Cheese\_1 et Cheese\_3 puis un second évènement d'hybridation a permis de former le second cluster génétique proche de Cheese\_2 qui en divergeant aurait donné à son tour la population Cheese\_2. (C) Un second scénario met en jeu la divergence des trois populations du fromage à partir du premier groupe d'hybrides puis par réhybridation, on aurait la formation du second groupe d'hybrides. (D) Parmi ces souches hybrides, on recense de nombreuses souches industrielles (indiquées par une étoile rouge) ainsi que des clusters de souches quasi identiques génétiquement, bien qu'isolés de fromages différents, ce qui indique que ceux-ci ont été inoculés par la même souche. Ces souches hybrides pourraient posséder des phénotypes d'intérêts en fromagerie.

2. Il serait intéressant de tester la fertilité de *P. camemberti* et *G. candidum* tout comme cela a été fait avec *P. roqueforti*. L'isolation génétique et la fertilité sont deux caractéristiques importantes pour mieux comprendre l'évolution des différentes populations. De plus, pouvoir réaliser des croisements au laboratoire permettrait de générer de nouveaux génotypes et ainsi faire de la génétique d'association. Ainsi, on pourrait faire le lien entre génotype et phénotype observés au sein de ces populations. Les conditions de reproduction sexuée de *P. roqueforti* étant déjà connues, on peut appliquer des conditions similaires à *P. biforme* et *P. fuscoglaucum*. En revanche, chez *G. candidum*, nous ne connaissons ni les structures reproductives (s'il y en a), ni les conditions requises à la reproduction sexuée.
3. Au sein des collections, certaines souches de *G. candidum* sont polyploïdes, avec dans certains cas, les deux types sexuels. Pour mon étude, nous avons fait le choix de ne pas les analyser. Contrairement à *S. cerevisiae* et *B. bruxellensis* où des populations domestiquées entières sont polyploïdes, seulement 8 souches de notre jeux de données (de 106 génomes) étaient potentiellement polyploïdes. J'ai pu les détecter à partir des sites hétérozygotes (*G. candidum* est haploïde) mais cela reste à vérifier par fluorescence. Les analyser serait l'occasion de mieux comprendre pourquoi chez *G. candidum* il n'y a pas eu de population polyploïde adaptée au milieu

fromage. Cependant, au vu de leur faible nombre et de leur répartition aussi bien dans les populations d'environnement naturel que celles du fromage, il sera compliqué de réaliser des analyses statistiques.

4. Jusqu'à maintenant les expériences de phénotypage ont été réalisées en boîte de Petri avec une seule espèce (à part pour la compétition). Il serait intéressant de voir si les résultats sont reproductibles *in caseo* et avec plus d'espèces en interaction (ferments lactiques et d'autres champignons du fromage). En particulier, *G. candidum* et *P. camemberti* sont inoculés ensemble au début du processus de fabrication du camembert et du brie. On peut se questionner sur leurs interactions et comment cela influence leur phénotype (composés volatils produits, morphologie, etc.)

## Références

- Aidoo, Kofi E., M. J. Rob Nout, et Prabir K. Sarkar. 2006. « Occurrence and Function of Yeasts in Asian Indigenous Fermented Foods ». *FEMS Yeast Research* 6 (1): 30-39. <https://doi.org/10.1111/j.1567-1364.2005.00015.x>.
- Alberto, Florian J., Frédéric Boyer, Pablo Orozco-terWengel, Ian Streeter, Bertrand Servin, Pierre de Villemereuil, Badr Benjelloun, et al. 2018. « Convergent Genomic Signatures of Domestication in Sheep and Goats ». *Nature Communications* 9 (1): 813. <https://doi.org/10.1038/s41467-018-03206-y>.
- Amrane, A., et Y. Prigent. 1997. « Comparison of growth parameters for *Geotrichum candidum* Geo17 and *Penicillium camemberti* LV2 in solid and liquid cultures. » *Lait* 77 (5): 641-48.
- Bai, Feng-Yan, Da-Yong Han, Shou-Fu Duan, et Qi-Ming Wang. 2022. « The Ecology and Evolution of the Baker's Yeast *Saccharomyces Cerevisiae* ». *Genes* 13 (2): 230. <https://doi.org/10.3390/genes13020230>.
- Baptiste, Eric, Maureen A O'Malley, Robert G Beiko, Marc Ereshefsky, J Peter Gogarten, Laura Franklin-Hall, François-Joseph Lapointe, et al. 2009. « Prokaryotic Evolution and the Tree of Life Are Two Different Things ». *Biology Direct* 4 (1): 34. <https://doi.org/10.1186/1745-6150-4-34>.
- Bell, Michael A., et Foster A Susan. 1994. *The Evolutionary Biology of the Threespine Stickleback*. Oxford, New York: Oxford University Press.
- Berger, Lee, Alexandra A. Roberts, Jamie Voyles, Joyce E. Longcore, Kris A. Murray, et Lee F. Skerratt. 2016. « History and Recent Progress on Chytridiomycosis in Amphibians ». *Fungal Ecology, Aquatic Fungi*, 19 (février): 89-99. <https://doi.org/10.1016/j.funeco.2015.09.007>.
- Bergström, Anders, David W. G. Stanton, Ulrike H. Taron, Laurent Frantz, Mikkel-Holger S. Sinding, Erik Ersmark, Saskia Pfrengle, et al. 2022. « Grey Wolf Genomic History Reveals a Dual Ancestry of Dogs ». *Nature* 607 (7918): 313-20. <https://doi.org/10.1038/s41586-022-04824-9>.
- Bigey, Frédéric, Diego Segond, Anne Friedrich, Stéphane Guezennec, Aurélie Bourgeois, Lucie Huyghe, Nicolas Agier, Thibault Nidelet, et Delphine Sicard. 2021. « Evidence for Two Main Domestication Trajectories in *Saccharomyces Cerevisiae* Linked to Distinct Bread-Making Processes ». *Current Biology* 31 (4): 722-732.e5. <https://doi.org/10.1016/j.cub.2020.11.016>.
- Blomqvist, Donald, Angela Pauliny, Mikael Larsson, et Lars-Åke Flodin. 2010. « Trapped in the Extinction Vortex? Strong Genetic Effects in a Declining Vertebrate Population ». *BMC Evolutionary Biology* 10 (1): 33. <https://doi.org/10.1186/1471-2148-10-33>.
- Boutrou, R., et M. Gueguen. 2005. « Interests in *Geotrichum Candidum* for Cheese Technology ». *International Journal of Food Microbiology* 102 (1): 1-20. <https://doi.org/10.1016/j.ijfoodmicro.2004.12.028>.
- Carmichael, J. W. 1957. « *Geotrichum Candidum* ». *Mycologia* 49 (6): 820-30. <https://doi.org/10.1080/00275514.1957.12024694>.
- Caron, Thibault, Mélanie Le Piver, Anne-Claire Péron, Pascale Lieben, René Lavigne, Sammy Brunel, Daniel Roueyre, et al. 2021. « Strong Effect of *Penicillium Roqueforti* Populations on Volatile and Metabolic Compounds Responsible for Aromas, Flavor and Texture in Blue Cheeses ». *International Journal of Food Microbiology* 354 (septembre): 109174.



- <https://doi.org/10.1016/j.ijfoodmicro.2021.109174>.
- Ceballos, Gerardo, Paul R. Ehrlich, Anthony D. Barnosky, Andrés García, Robert M. Pringle, et Todd M. Palmer. 2015. « Accelerated modern human-induced species losses: Entering the sixth mass extinction ». *Science Advances* 1 (5): e1400253. <https://doi.org/10.1126/sciadv.1400253>.
- Cheeseman, Kevin, Jeanne Ropars, Pierre Renault, Joëlle Dupont, Jérôme Gouzy, Antoine Branca, Anne-Laure Abraham, et al. 2014. « Multiple Recent Horizontal Transfers of a Large Genomic Region in Cheese Making Fungi ». *Nature Communications* 5 (1): 1-9. <https://doi.org/10.1038/ncomms3876>.
- Cole, Nicholas J., Mikiko Tanaka, Alan Prescott, et Cheryll Tickle. 2003. « Expression of Limb Initiation Genes and Clues to the Morphological Diversification of Threespine Stickleback ». *Current Biology* 13 (24): R951-52. <https://doi.org/10.1016/j.cub.2003.11.039>.
- Comai, Luca. 2005. « The Advantages and Disadvantages of Being Polyploid ». *Nature Reviews Genetics* 6 (11): 836-46. <https://doi.org/10.1038/nrg1711>.
- Corbett-Detig, Russell B, Shelbi L Russell, Rasmus Nielsen, et Jonathan Losos. 2020. « Phenotypic Convergence Is Not Mirrored at the Protein Level in a Lizard Adaptive Radiation ». *Molecular Biology and Evolution* 37 (6): 1604-14. <https://doi.org/10.1093/molbev/msaa028>.
- Cornille, Amandine, Tatiana Giraud, Marinus J. M. Smulders, Isabel Roldán-Ruiz, et Pierre Gladieux. 2014. « The Domestication and Evolutionary Ecology of Apples ». *Trends in Genetics* 30 (2): 57-65. <https://doi.org/10.1016/j.tig.2013.10.002>.
- Crauwels, Sam, Filip Van Opstaele, Barbara Jaskula-Goiris, Jan Steensels, Christel Verreth, Lien Bosmans, Caroline Paulussen, et al. 2017. « Fermentation assays reveal differences in sugar and (off-) flavor metabolism across different *Brettanomyces bruxellensis* strains ». *FEMS Yeast Research* 17 (1): fow105. <https://doi.org/10.1093/femsyr/fow105>.
- Cresko, William A., Angel Amores, Catherine Wilson, Joy Murphy, Mark Currey, Patrick Phillips, Michael A. Bell, Charles B. Kimmel, et John H. Postlethwait. 2004. « Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations ». *Proceedings of the National Academy of Sciences* 101 (16): 6050-55. <https://doi.org/10.1073/pnas.0308479101>.
- De Gracia, Marie, Mathilde Cascales, Pascale Expert, Marie-Noëlle Bellanger, Bruno Le Cam, et Christophe Lemaire. 2015. « How Did Host Domestication Modify Life History Traits of Its Pathogens? » Édité par Richard A Wilson. *PLOS ONE* 10 (6): e0122909. <https://doi.org/10.1371/journal.pone.0122909>.
- De Vuyst, L., G. Vrancken, F. Ravyts, T. Rimaux, et S. Weckx. 2009. « Biodiversity, Ecological Determinants, and Metabolic Exploitation of Sourdough Microbiota ». *Food Microbiology* 26 (7): 666-75. <https://doi.org/10.1016/j.fm.2009.07.012>.
- De Vuyst, Luc, et Patricia Neysens. 2005. « The Sourdough Microflora: Biodiversity and Metabolic Interactions ». *Trends in Food Science & Technology* 16 (1-3): 43-56. <https://doi.org/10.1016/j.tifs.2004.02.012>.
- Diamond, Jared. 2002. « Evolution, Consequences and Future of Plant and Animal Domestication ». *Nature* 418 (6898): 700-707. <https://doi.org/10.1038/nature01019>.
- Dieuleveux, V., D. Van Der Pyl, J. Chataud, et M. Gueguen. 1998. « Purification and Characterization of Anti-Listeria Compounds Produced by *Geotrichum Candidum* ». *Applied and Environmental Microbiology* 64 (2): 800-803.

- <https://doi.org/10.1128/AEM.64.2.800-803.1998>.
- Duan, Shou-Fu, Pei-Jie Han, Qi-Ming Wang, Wan-Qiu Liu, Jun-Yan Shi, Kuan Li, Xiao-Ling Zhang, et Feng-Yan Bai. 2018. « The Origin and Adaptive Evolution of Domesticated Populations of Yeast from Far East Asia ». *Nature Communications* 9 (1): 2690. <https://doi.org/10.1038/s41467-018-05106-7>.
- Dugat-Bony, Eric, Lucille Garnier, Jeremie Denonfoux, Stéphanie Ferreira, Anne-Sophie Sarthou, Pascal Bonnarme, et Françoise Irlinger. 2016. « Highlighting the Microbial Diversity of 12 French Cheese Varieties ». *International Journal of Food Microbiology* 238 (décembre): 265-73. <https://doi.org/10.1016/j.ijfoodmicro.2016.09.026>.
- Dugat-Bony, Eric, Cécile Straub, Aurélie Teissandier, Djamila Onésime, Valentin Loux, Christophe Monnet, Françoise Irlinger, et al. 2015. « Overview of a Surface-Ripened Cheese Community Functioning by Meta-Omics Analyses ». *PLOS ONE* 10 (4): e0124360. <https://doi.org/10.1371/journal.pone.0124360>.
- Dumas, Emilie, Alice Feurtey, Ricardo C. Rodríguez de la Vega, Stéphanie Le Prieur, Alodie Snirc, Monika Coton, Anne Thierry, et al. 2020. « Independent Domestication Events in the Blue-cheese Fungus *Penicillium Roqueforti* ». *Molecular Ecology* 29 (14): 2639-60. <https://doi.org/10.1111/mec.15359>.
- Dyer, Adrian G., Skye Boyd-Gerny, Stephen McLoughlin, Marcello G. P. Rosa, Vera Simonov, et Bob B. M. Wong. 2012. « Parallel evolution of angiosperm colour signals: common evolutionary pressures linked to hymenopteran vision ». *Proceedings of the Royal Society B: Biological Sciences* 279 (1742): 3606-15. <https://doi.org/10.1098/rspb.2012.0827>.
- Eberlein, Chris, Omar Abou Saada, Anne Friedrich, Warren Albertin, et Joseph Schacherer. 2021. « Different Trajectories of Polyploidization Shape the Genomic Landscape of the *Brettanomyces Bruxellensis* Yeast Species ». *Genome Research* 31 (12): 2316-26. <https://doi.org/10.1101/gr.275380.121>.
- Elmer, Kathryn R., Shaohua Fan, Henrik Kusche, Maria Luise Spreitzer, Andreas F. Kautt, Paolo Franchini, et Axel Meyer. 2014. « Parallel Evolution of Nicaraguan Crater Lake Cichlid Fishes via Non-Parallel Routes ». *Nature Communications* 5 (1): 1-8. <https://doi.org/10.1038/ncomms6168>.
- Elmer, Kathryn R., Henrik Kusche, Topi K. Lehtonen, et Axel Meyer. 2010. « Local variation and parallel evolution: morphological and genetic diversity across a species complex of neotropical crater lake cichlid fishes ». *Philosophical Transactions of the Royal Society B: Biological Sciences* 365 (1547): 1763-82. <https://doi.org/10.1098/rstb.2009.0271>.
- Elmer, Kathryn R., et Axel Meyer. 2011. « Adaptation in the Age of Ecological Genomics: Insights from Parallelism and Convergence ». *Trends in Ecology & Evolution* 26 (6): 298-306. <https://doi.org/10.1016/j.tree.2011.02.008>.
- Fages, Antoine, Kristian Hanghøj, Naveed Khan, Charleen Gaunitz, Andaine Seguin-Orlando, Michela Leonardi, Christian McCrory Constantz, et al. 2019. « Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series ». *Cell* 177 (6): 1419-1435.e31. <https://doi.org/10.1016/j.cell.2019.03.049>.
- Fay, Justin C, et Joseph A Benavides. 2005. « Evidence for Domesticated and Wild Populations of *Saccharomyces Cerevisiae* ». Édité par James E Haber. *PLoS Genetics* 1 (1): e5. <https://doi.org/10.1371/journal.pgen.0010005>.
- Feurtey, Alice, Ellen Guitton, Marie De Gracia Coquerel, Ludovic Duvaux, Jason Shiller, Marie-Noëlle Bellanger, Pascale Expert, et al. 2020. « Threat to Asian Wild Apple Trees Posed by Gene Flow from Domesticated Apple Trees and

- Their “Pestified” Pathogens ». *Molecular Ecology* 29 (24): 4925-41.  
<https://doi.org/10.1111/mec.15677>.
- Frantz, Laurent A F, Joshua G Schraiber, Ole Madsen, Hendrik-Jan Megens, Alex Cagan, Mirte Bosse, Yogesh Paudel, Richard P M A Crooijmans, Greger Larson, et Martien A M Groenen. 2015. « Evidence of Long-Term Gene Flow and Selection during Domestication from Analyses of Eurasian Wild and Domestic Pig Genomes ». *Nature Genetics* 47 (10): 1141-48.  
<https://doi.org/10.1038/ng.3394>.
- Gabaldón, Toni. 2020. « Hybridization and the origin of new yeast lineages ». *FEMS Yeast Research* 20 (5): foaa040. <https://doi.org/10.1093/femsyr/foaa040>.
- Galagan, James E., Sarah E. Calvo, Christina Cuomo, Li-Jun Ma, Jennifer R. Wortman, Serafim Batzoglou, Su-In Lee, et al. 2005. « Sequencing of *Aspergillus Nidulans* and Comparative Analysis with *A. Fumigatus* and *A. Oryzae* ». *Nature* 438 (7071): 1105-15. <https://doi.org/10.1038/nature04341>.
- Galeta, Patrik, Martina Lázničková-Galetová, Mikhail Sablin, et Mietje Germonpré. 2021. « Morphological Evidence for Early Dog Domestication in the European Pleistocene: New Evidence from a Randomization Approach to Group Differences ». *The Anatomical Record* 304 (1): 42-62.  
<https://doi.org/10.1002/ar.24500>.
- Gallone, Brigida, Stijn Mertens, Jonathan L. Gordon, Steven Maere, Kevin J. Verstrepen, et Jan Steensels. 2018. « Origins, Evolution, Domestication and Diversity of *Saccharomyces Beer Yeasts* ». *Current Opinion in Biotechnology* 49 (février): 148-55. <https://doi.org/10.1016/j.copbio.2017.08.005>.
- Gallone, Brigida, Jan Steensels, Troels Prahl, Leah Soriaga, Veerle Saels, Beatriz Herrera-Malaver, Adriaan Merlevede, et al. 2016. « Domestication and Divergence of *Saccharomyces Cerevisiae Beer Yeasts* ». *Cell* 166 (6): 1397-1410.e16. <https://doi.org/10.1016/j.cell.2016.08.020>.
- Giannakou, Konstantina, Mark Cotterrell, et Daniela Delneri. 2020. « Genomic Adaptation of *Saccharomyces* Species to Industrial Environments ». *Frontiers in Genetics* 11. <https://www.frontiersin.org/article/10.3389/fgene.2020.00916>.
- Gibbons, John G., Leonidas Salichos, Jason C. Slot, David C. Rinker, Kriston L. McGary, Jonas G. King, Maren A. Klich, David L. Tabb, W. Hayes McDonald, et Antonis Rokas. 2012. « The Evolutionary Imprint of Domestication on Genome Variation and Function of the Filamentous Fungus *Aspergillus Oryzae* ». *Current Biology* 22 (15): 1403-9.  
<https://doi.org/10.1016/j.cub.2012.05.033>.
- Gillot, Guillaume, Jean-Luc Jany, Elisabeth Poirier, Marie-Bernadette Maillard, Stella Debaets, Anne Thierry, Emmanuel Coton, et Monika Coton. 2017. « Functional Diversity within the *Penicillium Roqueforti* Species ». *International Journal of Food Microbiology* 241 (janvier): 141-50.  
<https://doi.org/10.1016/j.ijfoodmicro.2016.10.001>.
- Giraud, Frédéric, Tatiana Giraud, Gabriela Aguilera, Elisabeth Fournier, Robert Samson, Corine Cruaud, Sandrine Lacoste, Jeanne Ropars, Aurélien Tellier, et Joëlle Dupont. 2010. « Microsatellite Loci to Recognize Species for the Cheese Starter and Contaminating Strains Associated with Cheese Manufacturing ». *International Journal of Food Microbiology* 137 (2-3): 204-13.  
<https://doi.org/10.1016/j.ijfoodmicro.2009.11.014>.
- Gladieux, Pierre, Sébastien Ravel, Adrien Rieux, Sandrine Cros-Arteil, Henri Adreit, Joëlle Milazzo, Maud Thierry, Elisabeth Fournier, Ryohei Terauchi, et Didier Tharreau. 2018. « Coexistence of Multiple Endemic and Pandemic Lineages

- of the Rice Blast Pathogen ». Édité par David Guttman. *MBio* 9 (2): e01806-17. <https://doi.org/10.1128/mBio.01806-17>.
- Gladieux, Pierre, Jeanne Ropars, Hélène Badouin, Antoine Branca, Gabriela Aguilera, Damien M. Vienne, Ricardo C. Rodríguez de la Vega, Sara Branco, et Tatiana Giraud. 2014. « Fungal Evolutionary Genomics Provides Insight into the Mechanisms of Adaptive Divergence in Eukaryotes ». *Molecular Ecology* 23 (4): 753-73. <https://doi.org/10.1111/mec.12631>.
- Gobbetti, Marco, Carlo G. Rizzello, Raffaella Di Cagno, et Maria De Angelis. 2014. « How the Sourdough May Affect the Functional Features of Leavened Baked Goods ». *Food Microbiology* 37 (février): 30-40. <https://doi.org/10.1016/j.fm.2013.04.012>.
- Groppi, Alexis, Shuo Liu, Amandine Cornille, Stéphane Decroocq, Quynh Trang Bui, David Tricon, Corinne Cruaud, et al. 2021. « Population Genomics of Apricots Unravels Domestication History and Adaptive Events ». *Nature Communications* 12 (1): 3956. <https://doi.org/10.1038/s41467-021-24283-6>.
- Gross, Briana L., et Kenneth M. Olsen. 2010. « Genetic perspectives on crop domestication ». *Trends in plant science* 15 (9): 529-37. <https://doi.org/10.1016/j.tplants.2010.05.008>.
- Harrouard, Jules, Chris Eberlein, Patricia Ballestra, Marguerite Dols-Lafargue, Isabelle Masneuf-Pomarede, Cécile Miot-Sertier, Joseph Schacherer, Warren Albertin, et Jeanne Ropars. 2022. « *Brettanomyces Bruxellensis*: Overview of the Genetic and Phenotypic Diversity of an Anthropized Yeast ». *Molecular Ecology*, avril, mec.16439. <https://doi.org/10.1111/mec.16439>.
- Hawksworth, David L., et Robert Lücking. 2017. « Fungal Diversity Revisited: 2.2 to 3.8 Million Species ». Édité par Joseph Heitman et Timothy Y. James. *Microbiology Spectrum* 5 (4): 5.4.10. <https://doi.org/10.1128/microbiolspec.FUNK-0052-2016>.
- Hermet, Antoine, Delphine Méheust, Jérôme Mounier, Georges Barbier, et Jean-Luc Jany. 2012. « Molecular Systematics in the Genus *Mucor* with Special Regards to Species Encountered in Cheese ». *Fungal Biology* 116 (6): 692-705. <https://doi.org/10.1016/j.funbio.2012.04.002>.
- Holt, Sylvester, Vaskar Mukherjee, Bart Lievens, Kevin J. Verstrepen, et Johan M. Thevelein. 2018. « Bioflavoring by Non-Conventional Yeasts in Sequential Beer Fermentations ». *Food Microbiology* 72 (juin): 55-66. <https://doi.org/10.1016/j.fm.2017.11.008>.
- Hufford, Matthew B, Xun Xu, Joost van Heerwaarden, Tanja Pyhäjärvi, Jer-Ming Chia, Reed A Cartwright, Robert J Elshire, et al. 2012. « Comparative Population Genomics of Maize Domestication and Improvement ». *Nature Genetics* 44 (7): 808-11. <https://doi.org/10.1038/ng.2309>.
- Irlinger, Françoise, et Jérôme Mounier. 2009. « Microbial Interactions in Cheese: Implications for Cheese Quality and Safety ». *Current Opinion in Biotechnology* 20 (2): 142-48. <https://doi.org/10.1016/j.copbio.2009.02.016>.
- Jacques, Noémie, Sandrine Mallet, Fatima Laaghouti, Colin R. Tinsley, et Serge Casaregola. 2017. « Specific Populations of the Yeast *Geotrichum Candidum* Revealed by Molecular Typing ». *Yeast* 34 (4): 165-78. <https://doi.org/10.1002/yea.3223>.
- Jollivet, Nathalie, Jean Chataud, Yves Vayssier, Maurice Bensoussan, et Jean-Marc Belin. 1994. « Production of Volatile Compounds in Model Milk and Cheese Media by Eight Strains of *Geotrichum Candidum* Link ». *Journal of Dairy Research* 61 (2): 241-48. <https://doi.org/10.1017/S0022029900028259>.

- Jones, Felicity C., Manfred G. Grabherr, Yingguang Frank Chan, Pamela Russell, Evan Mauceli, Jeremy Johnson, Ross Swofford, et al. 2012. « The Genomic Basis of Adaptive Evolution in Threespine Sticklebacks ». *Nature* 484 (7392): 55-61. <https://doi.org/10.1038/nature10944>.
- Kassam, Daud D., Dean C. Adams, Michio Hori, et Kosaku Yamaoka. 2003. « Morphometric Analysis on Ecomorphologically Equivalent Cichlid Species from Lakes Malawi and Tanganyika ». *Journal of Zoology* 260 (2): 153-57. <https://doi.org/10.1017/S0952836903003571>.
- Kawtharani, Hiba, Selma Pascale Snini, Sorphea Heang, Jalloul Bouajila, Patricia Taillandier, Florence Mathieu, et Sandra Beaufort. 2020. « Phenyllactic Acid Produced by *Geotrichum Candidum* Reduces *Fusarium Sporotrichioides* and *F. Langsethiae* Growth and T-2 Toxin Concentration ». *Toxins* 12 (4): 209. <https://doi.org/10.3390/toxins12040209>.
- Komatsuda, T., P. Maxim, N. Senthil, et Y. Mano. 2004. « High-Density AFLP Map of Nonbrittle Rachis 1 (Btr1) and 2 (Btr2) Genes in Barley (*Hordeum Vulgare* L.) ». *Theoretical and Applied Genetics* 109 (5): 986-95. <https://doi.org/10.1007/s00122-004-1710-0>.
- Koonin, Eugene V., et Yuri I. Wolf. 2009. « Is Evolution Darwinian or/and Lamarckian? ». *Biology Direct* 4 (1): 42. <https://doi.org/10.1186/1745-6150-4-42>.
- Lahue, Caitlin, Anne Madden, Robert Dunn, et Caiti Smukowski Heil. 2020. « History and Domestication of *Saccharomyces cerevisiae* in Bread Baking ». *Frontiers in Genetics* 11. <https://www.frontiersin.org/articles/10.3389/fgene.2020.584718>.
- Larson, Greger, Elinor K. Karlsson, Angela Perri, Matthew T. Webster, Simon Y. W. Ho, Joris Peters, Peter W. Stahl, et al. 2012. « Rethinking Dog Domestication by Integrating Genetics, Archeology, and Biogeography ». *Proceedings of the National Academy of Sciences* 109 (23): 8878-83. <https://doi.org/10.1073/pnas.1203005109>.
- Leducq, Jean-Baptiste, Lou Nielly-Thibault, Guillaume Charron, Chris Eberlein, Jukka-Pekka Verta, Pedram Samani, Kayla Sylvester, Chris Todd Hittinger, Graham Bell, et Christian R. Landry. 2016. « Speciation Driven by Hybridization and Chromosomal Plasticity in a Wild Yeast ». *Nature Microbiology* 1 (janvier): 15003. <https://doi.org/10.1038/nmicrobiol.2015.3>.
- Lee, Tracy Jiaye, Yu-Ching Liu, Wei-An Liu, Yu-Fei Lin, Hsin-Han Lee, Huei-Mien Ke, Jen-Pan Huang, et al. 2022. « Extensive Sampling of *Saccharomyces Cerevisiae* in Taiwan Reveals Ecology and Evolution of Predomesticated Lineages ». *Genome Research* 32 (5): 864-77. <https://doi.org/10.1101/gr.276286.121>.
- Legras, Jean-Luc, Virginie Galeote, Frédéric Bigey, Carole Camarasa, Souhir Marsit, Thibault Nidelet, Isabelle Sanchez, et al. 2018. « Adaptation of *S. Cerevisiae* to Fermented Food Environments Reveals Remarkable Genome Plasticity and the Footprints of Domestication ». Édité par Patricia Wittkopp. *Molecular Biology and Evolution* 35 (7): 1712-27. <https://doi.org/10.1093/molbev/msy066>.
- Libkind, Diego, Chris Todd Hittinger, Elisabete Valério, Carla Gonçalves, Jim Dover, Mark Johnston, Paula Gonçalves, et José Paulo Sampaio. 2011. « Microbe Domestication and the Identification of the Wild Genetic Stock of Lager-Brewing Yeast ». *Proceedings of the National Academy of Sciences* 108 (35): 14539-44. <https://doi.org/10.1073/pnas.1105430108>.

- Librado, Pablo, Naveed Khan, Antoine Fages, Mariya A. Kusliy, Tomasz Suchan, Laure Tonasso-Calvière, Stéphanie Schiavinato, et al. 2021. « The Origins and Spread of Domestic Horses from the Western Eurasian Steppes ». *Nature*, octobre, 1-7. <https://doi.org/10.1038/s41586-021-04018-9>.
- Liti, Gianni, David M. Carter, Alan M. Moses, Jonas Warringer, Leopold Parts, Stephen A. James, Robert P. Davey, et al. 2009. « Population Genomics of Domestic and Wild Yeasts ». *Nature* 458 (7236): 337-41. <https://doi.org/10.1038/nature07743>.
- Liu, Xuexue, Yanli Zhang, Wujun Liu, Yefang Li, Jianfei Pan, Yabin Pu, Jianlin Han, Ludovic Orlando, Yuehui Ma, et Lin Jiang. 2022. « A Single-Nucleotide Mutation within the TBX3 Enhancer Increased Body Size in Chinese Horses ». *Current Biology* 32 (2): 480-487.e6. <https://doi.org/10.1016/j.cub.2021.11.052>.
- Lynch, Kieran M., Stuart Wilkinson, Luk Daenen, et Elke K. Arendt. 2021. « An Update on Water Kefir: Microbiology, Composition and Production ». *International Journal of Food Microbiology* 345 (mai): 109128. <https://doi.org/10.1016/j.ijfoodmicro.2021.109128>.
- Mabry, Makenzie E, Sarah D Turner-Hissong, Evan Y Gallagher, Alex C McAlvay, Hong An, Patrick P Edger, Jonathan D Moore, et al. 2021. « The Evolutionary History of Wild, Domesticated, and Feral Brassica oleracea (Brassicaceae) ». *Molecular Biology and Evolution* 38 (10): 4419-34. <https://doi.org/10.1093/molbev/msab183>.
- Machida, Masayuki, Kiyoshi Asai, Motoaki Sano, Toshihiro Tanaka, Toshitaka Kumagai, Goro Terai, Ken-Ichi Kusumoto, et al. 2005. « Genome Sequencing and Analysis of *Aspergillus Oryzae* ». *Nature* 438 (7071): 1157-61. <https://doi.org/10.1038/nature04300>.
- Mahler, D. Luke, Travis Ingram, Liam J. Revell, et Jonathan B. Losos. 2013. « Exceptional Convergence on the Macroevolutionary Landscape in Island Lizard Radiations ». *Science (New York, N. Y.)* 341 (6143): 292-95. <https://doi.org/10.1126/science.1232392>.
- Maldonado, R. R. 2017. « A Review on Geotrichum Lipases: Production, Purification, Immobilization and Applications ». *Chemical and Biochemical Engineering Quarterly Journal* 30 (4): 439-54. <https://doi.org/10.15255/CABEQ.2016.907>.
- Malekinejad, H, J Aghazadeh-Attari, A Rezaabakhsh, M Sattari, et B Ghasemsoltani-Momtaz. 2015. « Neurotoxicity of Mycotoxins Produced in Vitro by *Penicillium Roqueforti* Isolated from Maize and Grass Silage ». *Human & Experimental Toxicology* 34 (10): 997-1005. <https://doi.org/10.1177/0960327114565493>.
- Martínez-Ainsworth, Natalia E., et Maud I. Tenaillon. 2016. « Superheroes and Masterminds of Plant Domestication ». *Comptes Rendus Biologies, Trajectories of genetics, 150 years after Mendel / Trajectoire de la génétique, 150 après Mendel* Guest Editors / Rédacteurs en chef invités : Bernard Dujon, Georges Pelletier, 339 (7): 268-73. <https://doi.org/10.1016/j.crv.2016.05.005>.
- Mattanovich, Diethard, Michael Sauer, et Brigitte Gasser. 2014. « Yeast Biotechnology: Teaching the Old Dog New Tricks », 5.
- McClure, Sarah B., Clayton Magill, Emil Podrug, Andrew M. T. Moore, Thomas K. Harper, Brendan J. Culleton, Douglas J. Kennett, et Katherine H. Freeman. 2018. « Fatty Acid Specific  $\Delta^{13}\text{C}$  Values Reveal Earliest Mediterranean Cheese Production 7,200 Years Ago ». Édité par Juan J. Llor. *PLOS ONE* 13 (9): e0202807. <https://doi.org/10.1371/journal.pone.0202807>.

- McSweeney, Paul L. H. 2004. « Biochemistry of Cheese Ripening ». *International Journal of Dairy Technology* 57 (2-3): 127-44.  
<https://doi.org/10.1111/j.1471-0307.2004.00147.x>.
- Molimard, P., I. Lesschaeve, I. Bouvier, L. Vassal, P. Schlich, S. Issanchou, et He Spinnler. 1994. « Bitterness and Nitrogen Fractions of Mold Ripened Cheese of Camembert Type - Impact of the Association of *Penicillium-Camemberti* with *Geotrichum-Candidum* ». *Lait* 74 (5): 361-74.
- Monerawela, Chandre, et Ursula Bond. 2018. « The Hybrid Genomes of *Saccharomyces Pastorianus* : A Current Perspective: The Genomes of Lager Yeasts ». *Yeast* 35 (1): 39-50. <https://doi.org/10.1002/yea.3250>.
- Morrow, Carl A., et James A. Fraser. 2013. « Ploidy Variation as an Adaptive Mechanism in Human Pathogenic Fungi ». *Seminars in Cell & Developmental Biology*, WASP/WAVE proteins: expanding members and functions & The role of ploidy variation on cellular adaptation, 24 (4): 339-46.  
<https://doi.org/10.1016/j.semcd.2013.01.008>.
- Neiman, M., A. D. Kay, et A. C. Krist. 2013. « Can Resource Costs of Polyploidy Provide an Advantage to Sex? ». *Heredity* 110 (2): 152-59.  
<https://doi.org/10.1038/hdy.2012.78>.
- Nielsen, Marianne S, Jens C Frisvad, et Per V Nielsen. 1998. « Protection by Fungal Starters against Growth and Secondary Metabolite Production of Fungal Spoilers of Cheese ». *International Journal of Food Microbiology* 42 (1): 91-99. [https://doi.org/10.1016/S0168-1605\(98\)00070-1](https://doi.org/10.1016/S0168-1605(98)00070-1).
- O'Brien, Martin, Kristian F. Nielsen, Pdraig O'Kiely, Patrick D. Forristal, Hubert T. Fuller, et Jens C. Frisvad. 2006. « Mycotoxins and Other Secondary Metabolites Produced in Vitro by *Penicillium Paneum* Frisvad and *Penicillium Roqueforti* Thom Isolated from Baled Grass Silage in Ireland ». *Journal of Agricultural and Food Chemistry* 54 (24): 9268-76.  
<https://doi.org/10.1021/jf0621018>.
- O'Quin, K. E., C. M. Hofmann, H. A. Hofmann, et K. L. Carleton. 2010. « Parallel Evolution of Opsin Gene Expression in African Cichlid Fishes ». *Molecular Biology and Evolution* 27 (12): 2839-54.  
<https://doi.org/10.1093/molbev/msq171>.
- Peng, Junhua H., Dongfa Sun, et Eviatar Nevo. 2011. « Domestication Evolution, Genetics and Genomics in Wheat ». *Molecular Breeding* 28 (3): 281.  
<https://doi.org/10.1007/s11032-011-9608-4>.
- Perkins, Vincent, Stéphanie Vignola, Marie-Hélène Lessard, Pier-Luc Plante, Jacques Corbeil, Eric Dugat-Bony, Michel Frenette, et Steve Labrie. 2020. « Phenotypic and Genetic Characterization of the Cheese Ripening Yeast *Geotrichum Candidum* ». *Frontiers in Microbiology* 11.  
<https://doi.org/10.3389/fmicb.2020.00737>.
- Peter, Jackson, Matteo De Chiara, Anne Friedrich, Jia-Xing Yue, David Pflieger, Anders Bergström, Anastasie Sigwalt, et al. 2018. « Genome Evolution across 1,011 *Saccharomyces Cerevisiae* Isolates ». *Nature* 556 (7701): 339-44.  
<https://doi.org/10.1038/s41586-018-0030-5>.
- Petersen, Jessica L., James R. Mickelson, Aaron K. Rendahl, Stephanie J. Valberg, Lisa S. Andersson, Jeanette Axelsson, Ernie Bailey, et al. 2013. « Genome-Wide Analysis Reveals Selection for Important Traits in Domestic Horse Breeds ». Édité par Joshua M. Akey. *PLoS Genetics* 9 (1): e1003211.  
<https://doi.org/10.1371/journal.pgen.1003211>.
- Pitt, J. I. 1979. « The Genus *Penicillium* and Its Teleomorphic States *Eupenicillium*

- and *Talaromyces*. » *The Genus Penicillium and Its Teleomorphic States Eupenicillium and Talaromyces*.  
<https://www.cabdirect.org/cabdirect/abstract/19801362607>.
- Plassais, Jocelyn, Bridgett M. vonHoldt, Heidi G. Parker, Alberto Carmagnini, Nicolas Dubos, Ilenia Papa, Kevin Bevant, et al. 2022. « Natural and Human-Driven Selection of a Single Non-Coding Body Size Variant in Ancient and Modern Canids ». *Current Biology* 32 (4): 889-897.e9.  
<https://doi.org/10.1016/j.cub.2021.12.036>.
- Purugganan, Michael D. 2019. « Evolutionary Insights into the Nature of Plant Domestication ». *Current Biology* 29 (14): R705-14.  
<https://doi.org/10.1016/j.cub.2019.05.053>.
- Purugganan, Michael D., et Dorian Q. Fuller. 2009. « The Nature of Selection during Plant Domestication ». *Nature* 457 (7231): 843-48.  
<https://doi.org/10.1038/nature07895>.
- Qanbari, Saber, Hubert Pausch, Sandra Jansen, Mehmet Somel, Tim M. Strom, Ruedi Fries, Rasmus Nielsen, et Henner Simianer. 2014. « Classic Selective Sweeps Revealed by Massive Sequencing in Cattle ». *PLoS Genetics* 10 (2): e1004148. <https://doi.org/10.1371/journal.pgen.1004148>.
- Rainieri, Sandra, Yukiko Kodama, Yoshinobu Kaneko, Kozaburo Mikata, Yoshihiro Nakao, et Toshihiko Ashikari. 2006. « Pure and Mixed Genetic Lines of *Saccharomyces Bayanus* and *Saccharomyces Pastorianus* and Their Contribution to the Lager Brewing Strain Genome ». *Applied and Environmental Microbiology* 72 (6): 3968-74.  
<https://doi.org/10.1128/AEM.02769-05>.
- Range, Friederike, et Sarah Marshall-Pescini. 2022. « Comparing Wolves and Dogs: Current Status and Implications for Human 'Self-Domestication' ». *Trends in Cognitive Sciences* 26 (4): 337-49. <https://doi.org/10.1016/j.tics.2022.01.003>.
- Roach, Michael J., et Anthony R. Borneman. 2020. « New Genome Assemblies Reveal Patterns of Domestication and Adaptation across *Brettanomyces* (Dekkera) Species ». *BMC Genomics* 21 (1): 194.  
<https://doi.org/10.1186/s12864-020-6595-z>.
- Rodríguez, Ariel, Tia Rusciano, Rickeisha Hamilton, Leandra Holmes, Deidra Jordan, et Katharina C. Wollenberg Valero. 2017. « Genomic and Phenotypic Signatures of Climate Adaptation in an Anolis Lizard ». *Ecology and Evolution* 7 (16): 6390-6403. <https://doi.org/10.1002/ece3.2985>.
- Roostita, R. 1996. « The Occurrence and Growth of Yeasts in Camembert and Blue-Veined Cheeses ». *International Journal of Food Microbiology* 28 (3): 393-404. [https://doi.org/10.1016/0168-1605\(95\)00018-6](https://doi.org/10.1016/0168-1605(95)00018-6).
- Ropars, Jeanne, Thibault Caron, Ying-Chu Lo, Bastien Bennetot, et Tatiana Giraud. 2020. « La domestication des champignons *Penicillium* du fromage ». *Comptes Rendus. Biologies* 343 (2): 155-76. <https://doi.org/10.5802/crbio.15>.
- Ropars, Jeanne, Corinne Cruaud, Sandrine Lacoste, et Joelle Dupont. 2012. « A Taxonomic and Ecological Overview of Cheese Fungi ». *International Journal of Food Microbiology* 155 (3): 199-210.  
<https://doi.org/10.1016/j.ijfoodmicro.2012.02.005>.
- Ropars, Jeanne, Estelle Didiot, Ricardo C. Rodríguez de la Vega, Bastien Bennetot, Monika Coton, Elisabeth Poirier, Emmanuel Coton, Alodie Snirc, Stéphanie Le Prieur, et Tatiana Giraud. 2020. « Domestication of the Emblematic White Cheese-Making Fungus *Penicillium Camemberti* and Its Diversification into Two Varieties ». *Current Biology*, septembre.



- <https://doi.org/10.1016/j.cub.2020.08.082>.
- Ropars, Jeanne, Ying-Chu Lo, Emilie Dumas, Alodie Snirc, Dominik Begerow, Tanja Rollnik, Sandrine Lacoste, Joëlle Dupont, Tatiana Giraud, et Manuela López-Villavicencio. 2016. « Fertility Depression among Cheese-making *Penicillium Roqueforti* Strains Suggests Degeneration during Domestication ». *Evolution* 70 (9): 2099-2109. <https://doi.org/10.1111/evo.13015>.
- Ropars, Jeanne, Manuela López-Villavicencio, Joëlle Dupont, Alodie Snirc, Guillaume Gillot, Monika Coton, Jean-Luc Jany, Emmanuel Coton, et Tatiana Giraud. 2014. « Induction of sexual reproduction and genetic diversity in the cheese fungus *Penicillium roqueforti* ». *Evolutionary Applications* 7 (4): 433-41. <https://doi.org/10.1111/eva.12140>.
- Ropars, Jeanne, Ricardo C. Rodriguez de la Vega, Manuela Lopez-Villavicencio, Jerome Gouzy, Erika Sallet, Emilie Dumas, Sandrine Lacoste, et al. 2015. « Adaptive Horizontal Gene Transfers between Multiple Cheese-Associated Fungi ». *Current Biology* 25 (19): 2562-69. <https://doi.org/10.1016/j.cub.2015.08.025>.
- Schwarzer, Julia, Bernhard Misof, Seraphin N. Ifuta, et Ulrich K. Schliewen. 2011. « Time and Origin of Cichlid Colonization of the Lower Congo Rapids ». *PLoS ONE* 6 (7): e22380. <https://doi.org/10.1371/journal.pone.0022380>.
- Scopel, Eduardo F C, James Hose, Douda Bensasson, et Audrey P Gasch. 2021. « Genetic variation in aneuploidy prevalence and tolerance across *Saccharomyces cerevisiae* lineages ». *Genetics* 217 (4): iyab015. <https://doi.org/10.1093/genetics/iyab015>.
- Stajich, Jason E., Mary L. Berbee, Meredith Blackwell, David S. Hibbett, Timothy Y. James, Joseph W. Spatafora, et John W. Taylor. 2009. « The Fungi ». *Current Biology: CB* 19 (18): R840-845. <https://doi.org/10.1016/j.cub.2009.07.004>.
- Steensels, Jan, Brigida Gallone, et Kevin J. Verstrepen. 2021. « Interspecific Hybridization as a Driver of Fungal Evolution and Adaptation ». *Nature Reviews Microbiology*, mars, 1-16. <https://doi.org/10.1038/s41579-021-00537-4>.
- Svartberg, Kenth, et Björn Forkman. 2002. « Personality Traits in the Domestic Dog (*Canis Familiaris*) ». *Applied Animal Behaviour Science* 79 (2): 133-55. [https://doi.org/10.1016/S0168-1591\(02\)00121-1](https://doi.org/10.1016/S0168-1591(02)00121-1).
- Tinsley, Colin R., Noémie Jacques, Marine Lucas, Cécile Grondin, Jean-Luc Legras, et Serge Casaregola. 2022. « Molecular Genetic Analysis with Microsatellite-like Loci Reveals Specific Dairy-Associated and Environmental Populations of the Yeast *Geotrichum Candidum* ». *Microorganisms* 10 (1): 103. <https://doi.org/10.3390/microorganisms10010103>.
- Tirosh, Itay, Adina Weinberger, Miri Carmi, et Naama Barkai. 2006. « A Genetic Signature of Interspecies Variations in Gene Expression ». *Nature Genetics* 38 (7): 830-34. <https://doi.org/10.1038/ng1819>.
- Tran, Thierry, Cosette Grandvalet, François Verdier, Antoine Martin, Hervé Alexandre, et Raphaëlle Tourdot-Maréchal. 2020. « Microbial Dynamics between Yeasts and Acetic Acid Bacteria in Kombucha: Impacts on the Chemical Composition of the Beverage ». *Foods* 9 (7): 963. <https://doi.org/10.3390/foods9070963>.
- Turner, George F., Ole Seehausen, Mairi E. Knight, Charlotte J. Allender, et Rosanna L. Robinson. 2008. « How Many Species of Cichlid Fishes Are There in African Lakes?: Cichlid Species Richness ». *Molecular Ecology* 10 (3): 793-806. <https://doi.org/10.1046/j.1365-294x.2001.01200.x>.

- Varela, Javier A., Martina Puricelli, Raúl A. Ortiz-Merino, Romina Giacomobono, Stephanie Braun-Galleani, Kenneth H. Wolfe, et John P. Morrissey. 2019. « Origin of Lactose Fermentation in *Kluyveromyces Lactis* by Interspecies Transfer of a Neo-Functionalized Gene Cluster during Domestication ». *Current Biology* 0 (0). <https://doi.org/10.1016/j.cub.2019.10.044>.
- Warmuth, Vera, Anders Eriksson, Mim A. Bower, Javier Cañon, Gus Cothran, Ottmar Distl, Marie-Louise Glowatzki-Mullis, et al. 2011. « European Domestic Horses Originated in Two Holocene Refugia ». Édité par Michael Hofreiter. *PLoS ONE* 6 (3): e18194. <https://doi.org/10.1371/journal.pone.0018194>.
- Wayne, Robert K., et Bridgett M. vonHoldt. 2012. « Evolutionary Genomics of Dog Domestication ». *Mammalian Genome* 23 (1): 3-18. <https://doi.org/10.1007/s00335-011-9386-7>.