



**Emna BEN SAÂD**

**Structure, diversité et dynamique  
des communautés microbiennes impliquées  
dans la fermentation du kombucha**

HAPLOTEAM | UMR 7156 | Université de Strasbourg | BioBrasseurs | 2022

ÉCOLE DOCTORALE ED 414

UMR 7156 : Génétique Moléculaire Génomique Microbiologique

**THÈSE** présentée par :

**Emna BEN SAÂD**

soutenue le : 16 décembre 2022

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Aspects moléculaires et cellulaires de la biologie

**Structure, diversité et dynamique des communautés  
microbiennes impliquées dans la fermentation du  
kombucha**

**THÈSE dirigée par :**

**Pr SCHACHERER Joseph**

Professeur, Université de Strasbourg

**THÈSE co-encadrée par :**

**Dr BLEYKASTEN Claudine**

Maître de Conférences, Université de Strasbourg

**RAPPORTEURS :**

**Dr ALBERTIN Warren**

Maître de Conférences, Université de Bordeaux

**Dr LEGRAS Jean-Luc**

Ingénieur de recherche, INRAE Montpellier

**AUTRES MEMBRES DU JURY :**

**Pr FAIRHEAD Cécile**

Professeur, Université Paris-Saclay

**Pr KEYSER Christine**

Professeur, Université de Strasbourg

**INVITES :**

**Dr WAGNER Renaud**

Ingénieur de recherche, Université de Strasbourg

**COUROT Olivier**

Co-fondateur BioBrasseurs, Strasbourg



# *Dédicaces*

*A mes adorables mères  
Mon cher père et mon frère trésor  
Pour l'amour qu'ils m'ont toujours porté, pour leur  
persévérance,  
Pour leur aide et encouragement permanents, y compris dans  
les moments de découragement et de doute,  
C'est à vous que je dois tout. Vous êtes dépensés pour moi.  
En reconnaissance de tous les sacrifices consentis par tous et  
chacun pour me permettre d'atteindre cette étape de ma vie,  
Que ce travail soit le témoignage de mon profond respect et  
ma gratitude,  
A la team Bio Brasseurs d'avoir cru en moi,  
A mes amis d'enfance, de ma vie studieuse, de ma vie  
professionnelle, qu'ils trouvent ici l'expression de mon  
dévouement pour leur soutien  
A toutes les belles rencontres tout au long de ma vie  
autonome à l'étranger,  
A ceux qui me sont chers, qui me l'ont été et à moi pour la  
personne que je suis devenue*

*Je dédie ce travail  
Mamani*



## **Résumé**

Elucider les interactions entre les espèces d'une niche écologique est une thématique majeure en biologie. Grâce aux avancées technologiques des méthodes de séquençage haut débit, cette thématique connaît un essor considérable dans le contexte des communautés microbiennes. Elles permettent d'identifier les taxons présents, de connaître leurs abondances et de mesurer l'impact des conditions environnementales, avec une très grande précision. Les communautés microbiennes appelées SCOBY (Symbiotic Culture of Bacteria and Yeast) ont représenté les modèles d'étude de ma thèse. Ces communautés de levures et de bactéries, dispersées entre une phase liquide et un biofilm de cellulose bactérienne, permettent d'élaborer le kombucha, une boisson fermentée ancestrale, pétillante et acidulée. Dans un premier temps, la caractérisation par métabarcoding d'une collection de 18 SCOBY d'origines variées a constitué une ressource pour décrire la diversité taxonomique de ces communautés. Dans un second temps, des séries temporelles, au laboratoire et en brasserie, ont permis de caractériser leur stabilité au cours des fermentations successives.

**Mots-clé :** fermentation, kombucha, métagénomique, diversité taxonomique, SCOBY, microbiologie.

## **Abstract**

Elucidating the interactions between the species of an ecological niche is a major theme in biology. Thanks to the recent advances in high-throughput sequencing methods, this theme is experiencing considerable growth in the context of microbial communities. They make it possible to identify the taxa present, to assess their abundance, and to measure the impact of environmental conditions, with great precision. The microbial communities called SCOBY (Symbiotic Culture of Bacteria and Yeast) represented the models of my thesis. These communities of yeasts and bacteria, dispersed between a liquid phase and a biofilm of bacterial cellulose, are involved in kombucha production, an ancestral fermented soft drink. First, the characterization by metabarcoding of a collection of 18 SCOBYs of various origins constituted a resource to describe the taxonomic diversity of these communities. Secondly, the study of time series, in the laboratory and in the brewery, made it possible to characterize their stability during successive fermentations.

**Keywords:** fermentation, kombucha, metagenomics, taxonomic diversity, SCOBY, microbiology.



## Remerciements

Cette thèse a été réalisée à Université de Strasbourg, dans le cadre d'une convention CIFRE, établie entre le laboratoire de Génétique Moléculaire, Génomique et Microbiologie, UMR7156, sous la direction du Pr. Joseph Schacherer et l'encadrement du Dr Claudine Bleykasten, dans l'équipe Variation intra-spécifique et évolution des génomes, et l'entreprise BioBrasseurs. Au terme de ce travail, je remercie tous ceux qui ont participé de près ou de loin à l'élaboration de mon projet de thèse dans les meilleures conditions.

Tout d'abord, j'aimerais remercier très sincèrement, les membres de mon comité de thèse Dr Warren Albertin, Dr Jean-Luc Legras, Pr Cécile Fairhead, Pr Christine Keyser, Dr Renaud Wagner, pour avoir bien accepté d'évaluer mon travail. Je souhaite par ailleurs remercier plus particulièrement Renaud Wagner d'avoir aussi participé à mon comité de mi-thèse et de m'avoir donné l'accès à l'Ultraturax. Merci à Warren Albertin pour Mama Kombucha (SCOPY MK). Merci également à Olivier Courot pour son investissement dans ce travail et de m'avoir donné l'opportunité de concrétiser mon idée de projet.

On dit toujours que le trajet est aussi important que la destination, C'est pour ça que je ne remercierai jamais assez Joseph. Merci d'avoir cru en moi et en mon projet appliqué et pour ta rigueur scientifique. Je tiens à remercier particulièrement Claudine, tu m'as encadrée avec patience durant ces 44 mois. Merci pour ton enthousiasme, depuis les premiers échanges en amont de la concrétisation du projet. Merci de m'avoir poussé, de ton écoute, des discussions sans fin et de ta disponibilité permanente. C'est à vous deux que j'exprime mes vifs remerciements pour vos conseils prodigués, votre patience et votre persévérance.

Olivier et Frédérique, merci de m'avoir fait confiance, il y a maintenant 6 ans en me proposant d'intégrer l'équipe BioBrasseurs, je n'aurais jamais pu imaginer l'aventure sans vous avoir rencontrés. C'est bien le destin qui nous a réunis autour du nœud de l'infini de KYO Kombucha.

Je remercie Olivier Rué, de la plateforme Migale INRAE, d'avoir pris le temps de répondre à tous mes mails, afin de mettre à disposition la banque D2 sur l'outil FROGS. Je tiens à remercier Dr Delphine Sicard d'avoir participé à mon comité

de mi-thèse et de m'avoir mis en contact avec Mr Sylvain Santoni de l'UMR AGAP à Montpellier, que je remercie pour l'échange et pour la qualité des données de séquençage générées.

Ma gratitude s'adresse également à l'équipe Haploteam, pour leur accueil chaleureux ainsi que leur disponibilité et les échanges constructifs lors des lab-meetings : Anne merci d'avoir les réponses à toutes mes questions, Fabien et Elie pour vos aides à débogger les codes dans R, qui m'ont pris des heures de recherche, en quelques secondes.

Merci à tous les membres actuels de l'équipe Claudia, Jing, Marion, Andreas, Victor, Isabelle et Gauthier et les anciens de l'équipe Elodie, Sabrina, Téo, Jean-Seb, Omar, Emilien, Arthur. Je vous remercie tous de votre soutien, pour votre bonne humeur et pour les moments agréables gravés dans la mémoire, au labo et dehors avec les activités que j'ai pu pratiquer avec vous. Je suis reconnaissante et j'ai autant de souvenirs de chacun d'entre vous. Je ne pourrais pas imaginer une ambiance de travail aussi propice dans une autre équipe. Et je n'aurai jamais imaginer courir 5 km à l'Ekiden après 6 mois de mon arrivée au labo et la pression de ne pas faire perdre mon équipe, moi qui n'ai jamais été fan des activités sportives, je me suis trouvée au terrain de foot, à l'escalade, en randonnée. Tout ça m'a permis de sortir de ma zone de confort et de tenter de nouvelles choses. Merci à vous tous et bonne continuation à vous tous.

Abhishek, Jan and Sam thanks for being part of the team that made a good environment for each of us to progress. It was nice to meet you, I wish you luck.

J'aimerais bien évidemment remercier l'ensemble des mes amis qui m'ont soutenu d'une manière ou d'une autre Amina, Nouha, Rayenne, Djamila, Pénélope, Marwen qui ont été toujours à mes côtés à me motiver et m'aider à garder un bon moral.

Je remercie toutes les nouvelles rencontres que j'ai fait depuis mon arrivée sur Strasbourg et les membres de la team BioBrasseurs Leila, Nicolas, Eric, Manon....

Enfin, mon plus grand merci revient tout naturellement à ma famille : Mes deux mamans, mon père et mon frère à qui je dois la personne que je suis et qu'ils soient assurés de ma sincère gratitude. Merci pour votre confiance.

# Sommaire

<b>Introduction.....</b>	<b>1</b>
<b>1. Les communautés microbiennes .....</b>	<b>3</b>
1.1. La diversité des communautés microbiennes à travers les aliments fermentés.....	3
1.2. Diversité des communautés microbiennes selon le processus d'initiation de la fermentation .....	6
1.3. Le kombucha sous toutes les coutures .....	10
1.3.1. Histoire du kombucha .....	11
1.3.2. Déroulement d'une fermentation de kombucha .....	12
1.3.3. À quels taxons appartiennent les levures et les bactéries impliqués dans la production de kombucha ? .....	18
1.3.4. Conditions et paramètres de fermentation.....	26
<b>2. Les contributions des méthodes de séquençage haut débit à la compréhension des communautés microbiennes impliquées dans les processus fermentaires.....</b>	<b>31</b>
2.1. Identification directe par observations au microscope .....	32
2.3. Caractérisation par des méthodes métagénomiques .....	35
2.3.2. Les méthodes métagénomiques non ciblées : principe et évolution du séquençage shotgun appliqué aux communautés microbiennes.....	48
<b>Histoire, contexte et vue d'ensemble du projet de thèse .....</b>	<b>51</b>
<b>Chapitre 1 : Mise au point des méthodes du projet .....</b>	<b>52</b>
<b>1. Cultures et fermentations du kombucha au laboratoire.....</b>	<b>57</b>
1.1 Établir un protocole de fermentation standard .....	57
1.2 Établir des collections de prélèvements représentatifs.....	61
<b>2. Mise au point des différentes étapes du protocole de métagénomique ciblé par séquençage haut débit (ou métabarcoding) .....</b>	<b>63</b>
2.1 Le traitement des données de séquençage.....	64
2.2 Extraction de l'ADN total des cellules d'un prélèvement.....	67
2.3 Génération des bibliothèques d'amplicons.....	71
2.4 Comparaison des données de séquençage générées par les deux plateformes .....	80

<b>Chapitre 2 :Exploration de la diversité taxonomique de SCOBY .....</b>	<b>E85</b>
Introduction.....	<b>86</b>
Résultats.....	<b>87</b>
Les starters utilisés pour générer les SCOBY .....	87
Génération d'une collection de SCOBY comparables .....	89
Isolement et caractérisation d'une collection de souches de levures et de bactéries	93
Diversity in the microbial communities involved in kombucha fermentations performed with SCOBY from different origins.....	<b>97</b>
Introduction.....	<b>97</b>
Results.....	<b>100</b>
Discussion.....	<b>114</b>
Material and methods.....	<b>116</b>
<b>Chapitre3 : Dynamique des communautés microbiennes .....</b>	<b>125</b>
<b>1 Introduction.....</b>	<b>126</b>
<b>2 Résultats.....</b>	<b>129</b>
2.1 Composition microbienne des SCOBY des sites de BioBrasseurs .....	129
2.1.1 Schéma expérimental et collection des prélèvements .....	129
2.1.2 Résultats des compositions microbiennes .....	133
2.1.3. Discussion .....	135
2.2 Effet des back-slopping successifs à l'échelle laboratoire .....	137
2.2.1 Premier essai : suivi de 20 lignées de fermentation parallèles .....	138
2.2.2 Deuxième essai : suivi de 11 lignées de fermentation parallèles .....	145
2.2.3 Discussion .....	150
2.3 Effets de différentes conditions de conservation du starter.....	151
2.3.1 Description du dispositif expérimental et prélèvements .....	151
2.3.2 Résultats des compositions microbiennes .....	155
2.3.3 Discussion .....	159
<b>3. Conclusions.....</b>	<b>161</b>
<b>Conclusion générale et perspectives .....</b>	<b>163</b>

**Références..... 173**

**Annexes ..... 189**

Publication en cours de rédaction ..... 191

Communications scientifiques ..... 192

Activités chez le producteur de kombucha BioBrasseurs ..... 193



# Introduction

## Préambule

A l'origine des travaux que je vais présenter dans cette thèse, se trouve l'entreprise Bio Brasseurs, productrice du kombucha à Strasbourg, qui a souhaité orienter sa démarche stratégique vers une meilleure connaissance des communautés microbiennes. Ainsi, un dispositif CIFRE a pu être monté, dans le cadre d'une collaboration de recherche avec l'équipe Variation intra-spécifique et évolution des génomes dirigée par le Pr Joseph SCHACHERER, au sein du laboratoire de Génétique Moléculaire, Génomique et Microbiologie, UMR7156. Dans ce dispositif, toute entreprise de droit français peut financer le travail d'un doctorant pour le placer au cœur d'une collaboration de recherche avec un laboratoire public. J'ai pu être intégrée dans ce dispositif en tant que doctorante, grâce à ma connaissance passée du kombucha, développée lors de mon projet de fin d'études et lors de mon implication dans le démarrage des activités de production de la brasserie.





L'alimentation pour l'Homme est une fonction vitale qui a évolué au cours du temps. L'exploration des ressources disponibles dans la nature était la première finalité avec la cueillette, la chasse et la pêche. L'utilisation de ces produits à l'état brut s'est complétée des processus de transformation, dont les fermentations qui permettent une diversification des aliments. Parmi ces aliments, Le kombucha est une boisson obtenue par fermentation d'une infusion de thé sucrée. Les communautés de micro-organismes, qui effectuent les transformations par fermentation, incluent des espèces pouvant appartenir à différents domaines du vivant et constituent des modèles intéressants pour l'étude des interactions entre espèces. La compréhension des communautés impliquées dans les processus fermentaires est en plein essor avec le développement des méthodes moléculaires appliquées à des approches métagénomiques. Cette introduction va décrire, dans une première partie, les spécificités du kombucha et de sa fermentation par rapport aux grandes familles de boissons fermentées. La deuxième partie sera consacrée aux méthodes d'études des communautés microbiennes avec en particulier l'apport des méthodes les plus récentes basées sur les techniques de séquençage dites à 'haut débit'.

## **1. Les communautés microbiennes**

### **1.1. La diversité des communautés microbiennes à travers les aliments fermentés**

Avant de se consacrer aux communautés de micro-organismes du kombucha, il est intéressant d'avoir un survol de l'ensemble des aliments fermentés. Pour illustrer tous les aspects de cette diversité, une large gamme de produits fermentés ont été groupés en fonction des familles de produit dans une table périodique (Figure 1, Gänzle, 2022). Vingt familles de produits ont été identifiées en fonction du produit fini obtenu, avec une grande part d'entre elles correspondant aux boissons fermentées telles que les bières, les vins, les produits laitiers, les

vinaigres, le kombucha, le café et diverses autres boissons moins connues. Ces boissons se distinguent donc bien sûr par les micro-organismes qui ont permis de les élaborer et par leur présence dans le liquide directement ou adhérent à un support solide, ou dans les deux, comme dans le cas du kombucha, du vinaigre ou des kéfirs. Différents types de fermentations ou de transformation du substrat peuvent se dérouler et plusieurs cycles fermentaires peuvent être nécessaires, comme pour le kombucha qui résulte de la fermentation alcoolique et de celle dite acétique. Une caractéristique importante renseignées par cette table est la présence ou non de micro-organismes vivants dans le produit consommé alors que ces produits sont souvent consommés pour leurs propriétés probiotiques réelles ou supposées. Les produits cuits (comme le pain) ou distillés (comme la téquila) ne contiennent plus les micro-organismes présents lors de la fermentation.

**Figure 1. Classification des aliments fermentés selon un principe de table périodique.**

Ce tableau permet de visualiser les informations sur l'origine géographique, les ingrédients principaux, le pH, la durée de fermentation, les métabolites majeurs, les micro-organismes principaux et le besoin (nom souligné) ou non de starter. Le kombucha est encadré. (Gänzle, 2022) (image haute résolution à ce lien : <https://drive.google.com/drive/folders/10cH7XaNqGQAKdL3VHhd1tzWXIGaTdfkV>)

# Periodic Table of Fermented Foods

Michael Gänzle

Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB, Canada

1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		17		18																																																																																																																																																																					
Wine and cider		Feder-wasser		White wine		Red wine		Champagne		Light beer		Dark beer		Tubers and roots		Beverages from cereals or tubers		Cereal porridges		Bread (wheat and other)		Soy and beans		Soy-based condiments		Vegetable products		Dairy products		Acidified products		Rennet coagulated cheese		Surface ripened cheese		Fish		Summer sausage																																																																																																																																																																	
1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		17		18																																																																																																																																																																					
1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		17		18																																																																																																																																																																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200

Lactic acid bacteria		Other bacteria (Gram positive)		Gram-negative bacteria	
1		2		3	
1	2	3	4	5	6

Acknowledgements: Dr. Jim McClure and Dr. Ying Yu are acknowledged for helpful suggestions. Copyright: This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. The exact file used to generate the table was adapted from: [https://doi.org/10.1007/978-1-4939-9888-5\\_1](https://doi.org/10.1007/978-1-4939-9888-5_1). This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

## 1.2. Diversité des communautés microbiennes selon le processus d'initiation de la fermentation

Pour l'étude des communautés microbiennes impliquées dans l'élaboration des aliments fermentés, il est important de prendre en compte l'origine des micro-organismes impliqués à savoir s'ils proviennent de la flore indigène du substrat initial ou s'ils ont été ajoutés. Cette distinction reflète la manière dont les communautés microbiennes se constituent puis sont propagées et permet de comprendre les variations en compositions microbiennes observées pour le même produit.

### 1.2.1. Communautés constituées lors de fermentations spontanées

Lorsqu'on parle de fermentation spontanée, aucun ajout de micro-organisme n'est effectué et seule la microflore indigène et/ou celle de l'air environnant contribue à l'élaboration de la composition microbienne (Callanan et al., 2021).

Comme exemples, on peut citer les communautés qui se forment lors de la fermentation des graines de cacao et celle des bières de style lambic. On peut aussi citer le processus de lacto-fermentation pour obtenir de la choucroute, des pickles ou du kimchi. Il faut souligner que dans ces derniers exemples, des composés sont ajoutés en début de fermentation pour inhiber les activités non désirées de micro-organismes présents dans la flore indigène. Ainsi, pour les végétaux lactofermentés, une teneur élevée en sel (2%) ralentit l'activité des levures et des champignons mais pas celle, désirée, des bactéries lactiques. La fermentation développe ensuite un environnement non propice aux développements microbiens ultérieurs de contaminants exogènes.

Dans le cas des fermentations spontanées, les fluctuations saisonnières et les différences de qualité du substrat initial contribuent aux variations de composition des communautés microbiennes et par conséquent aux propriétés de l'aliment

fermenté. Ainsi, les compositions microbiennes du kimchi produit par la fermentation de choux préalablement conservés en chambre froide, différent de celles de kimchi produit avec des choux fraîchement récoltés.

### 1.2.2. Communautés constituées lors de l'utilisation de starters

Pour certains types de fermentation ou pour améliorer la maîtrise de certains procédés spontanés, l'ajout intentionnel de micro-organismes, sous forme de starter, est pratiqué pour initier ou conduire la fermentation (Callanan et al., 2021 ; Min et al., 2022) Un starter est défini comme une préparation ou un matériel contenant soit un micro-organisme particulier soit un grand nombre de micro-organismes variés qui est ajouté pour initier, accélérer ou améliorer le déroulement d'une fermentation (Corsetti et al., 2012). De nombreux aspects caractérisent les starters :

- un starter peut être utilisé soit en complément de la microflore indigène ou comme seule source de micro-organismes suite à l'élimination de la flore indigène, généralement par chauffage,
- il existe des starters primaires ou secondaires, selon qu'ils sont ajoutés au début de la fermentation pour l'initier ou la stimuler, ou en cours de fermentation afin d'affiner les arômes du produit final (Parente et al., 2017).

Nous allons décrire la classification des starters en se basant sur la façon dont ils sont élaborés (Parente et al., 2017) car cet aspect conditionne le statut de leur composition microbienne qui peut être indéfinie ou définie.

#### 1.2.2.1. Les starters naturels obtenus par fermentation spontanée

Certains starters sont eux-mêmes obtenus par fermentation spontanée. Dans l'exemple des levains du pain, les starters sont obtenus à partir de la fermentation des mêmes types de substrats que la pâte de pain au levain, à savoir de la farine et de l'eau. Il y a cependant des différences avec la fermentation de la pâte à pain :

certaines farines, comme celle de seigle, sont plus appropriées pour l'élaboration d'un starter et du sucre est souvent ajouté pour favoriser la fermentation. La maturation d'un starter de levain dure plusieurs jours et il est ensuite entretenu en ajoutant de la farine et du sucre à une fraction du starter mature, un processus appelé back-slopping. Les micro-organismes impliqués proviennent de la farine mais aussi de l'atelier où le starter est élaboré, la composition de ces starters naturels est donc indéfinie car le nombre de souches microbiennes et leurs proportions sont variables d'un atelier à l'autre et selon de degré de maturation (von Gastrow et al., 2022). Ces starters sont ensuite utilisés pour ensemercer les divers mélanges de farines utilisées pour la fermentation des pâtes au levain donnant différentes sortes de pain. De nouveaux starters peuvent être générés à tout moment en recourant à une fermentation spontanée.

#### 1.2.2.2. Les starters naturels dépendants du produit final de fermentation

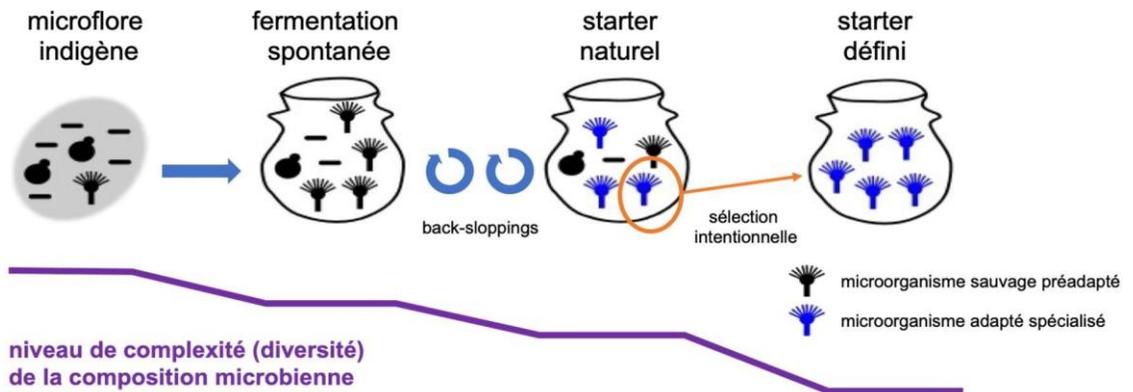
Pour les produits fermentés tels que le kombucha, les kéfirs et le Gwell, un gros lait fermenté en Bretagne, il est impossible de recourir à une fermentation spontanée. De plus, une de leur particularité est que le produit final de la fermentation (ou les produits) est également l'inoculum de la fermentation suivante : il faut impérativement disposer du ou des produits de la fermentation pour démarrer une nouvelle fermentation. Si le produit/starter est perdu ou altéré, il est nécessaire de s'en procurer auprès de producteurs amateurs ou professionnels. Le procédé de propagation est également appelé back-slopping et la composition de ces starters naturels est indéfinie.

#### 1.2.2.3. Les starters commerciaux définis

Leur composition est connue, aussi bien pour le nombre de souches présentes que pour leurs proportions. Chaque souche ayant été cultivée individuellement et préalablement caractérisée, sélectionnée, voir améliorée (mais non OGM) (Derkx et al., 2014). Ces starters sont utilisés pour les mêmes types de fermentations que

celles conduites avec des starters indéfinis, leur intérêt réside dans l'utilisation de souches qui présentent des propriétés fonctionnelles et un potentiel technologique connu. Ces souches ont été sélectionnées soit pour leur capacité à fonctionner sans être affecté par des variations dans la composition de la matrice de base de l'aliment, soit lié à un apport technologique conférant à l'aliment des propriétés particulières. Pour un effet optimal, les souches utilisées doivent être adaptées à chaque type de fermentation (Aldrete-Tapia et al., 2020), ce sont donc des starters très spécialisés.

De manière générale, si on compare la composition d'une microflore indigène avec celle d'un produit fermenté spontanément par cette microflore et celle des starters naturels qui peuvent en dériver, on peut observer une diminution continue de la complexité de la composition (Gibbons and Rinker, 2015). Cette diminution de la complexité (ou de la diversité) se mesure par le nombre de taxons présents et par leurs proportions similaires. La diversité diminue parce que des taxons de la microflore indigène disparaissent, parce qu'ils n'étaient pas pré-adaptés aux conditions du milieu de fermentation, et que d'autres deviennent majoritaires. Cette diminution de la diversité est encore plus importante pour les starters qui sont entretenus par back-slopping sur de longues périodes. Elle peut s'accompagner d'une évolution des taxons présents, par différents mécanismes comme la sélection d'événements de mutation, d'hybridation ou de transfert horizontal de gènes qui leurs confèrent un avantage sélectif dans les conditions de fermentation. Ce processus est apparenté à une forme de domestication de ces communautés microbiennes. La domestication est poussée à son extrême pour les starters composés d'un assemblage définis de souches pures (Figure 2).



**Figure 2.** Diminution de la complexité des compositions des communautés microbiennes au fur et à mesure de la spécialisation des starters (adapté de (Gibbons and Rinker, 2015))

### 1.3. Le kombucha sous toutes les coutures

Le kombucha est une boisson obtenue par fermentation microbienne d'une infusion de thé sucrée. L'ensemble des micro-organismes impliqués porte le nom de SCOBY (Symbiotic Culture of Bacteria and Yeast) (Soares et al., 2021). La boisson est légèrement sucrée, d'où sa faible teneur en calories, légèrement effervescente et sa saveur fruitée et acidulée. Elle est rafraîchissante et constitue une alternative saine aux boissons de type soda. Sa composition comprend divers sucres dont le saccharose, le glucose et le fructose, avec une teneur globale de l'ordre de 3% (v/v). Elle comprend également des acides organiques comme l'acide acétique, gluconique et glucuronique, des vitamines hydrosolubles, des acides aminés, des amines biogènes, des purines, des pigments, des lipides, des protéines, des enzymes hydrolytiques, de l'éthanol, du dioxyde de carbone, des polyphénols, de l'acide D-saccharique-1,4-lactone, des minéraux (manganèse, fer, nickel, cuivre, zinc, plomb, cobalt, chrome et cadmium), des anions (fluorure, chlorure, bromure, iodure, nitrate, phosphate et sulfate) et des produits métaboliques des levures et des bactéries (Jayabalan et al., 2014). Le kombucha contient aussi de l'éthanol, généralement à des teneurs inférieures à 1.2 % ce qui lui permet d'être accepté comme boisson non alcoolisée (Code de la santé

publique, 2016) mais il existe des procédés permettant de produire des 'hard kombucha' pouvant titrer jusqu'à 5% d'éthanol. Lorsqu'elle n'est pas pasteurisée, la boisson contient également des bactéries et des levures vivantes. C'est une boisson décrite comme fonctionnelle car de nombreuses études lui confèrent des propriétés biologiques intéressantes pour les organismes qui la consomment (Jayabalan et al., 2014 ; Diez-Ozaeta and Astiazaran, 2022).

### 1.3.1. Histoire du kombucha

L'origine géographique du kombucha reste indéterminé bien que liée à une région ou un pays d'Asie de l'est : Japon, Mandchourie, Tibet, Corée et Russie sont des prétendants (Petruzzello, 2021).

La mythologie du premier kombucha rapporte l'histoire millénaire d'un phénomène spontané : une infusion sucrée par du miel est oubliée pendant quelques jours et finit par avoir une couche qui flotte à la surface. Les mouches de fruit attirées par le sucre auraient contaminé le liquide par des bactéries acétiques et lactiques dont l'activité auraient donné le premier kombucha (Tadeusz, 2020).

Il est également dit qu'un médecin nommé "Kombu" a introduit le champignon du thé, "cha" au Japon vers 414 après J.-C. depuis la Corée, afin de soigner les troubles digestifs de l'empereur Inkyo (Jarrell et al., 2000).

Le thé Kombucha a été reconnu au Japon à la fin du 19e siècle et était appelé "thé champignon" ou "thé Kvass". Plus tard, avec l'extension des routes commerciales, le "Thé Kvass" se serait répandu en Inde et en Russie par les routes commerciales. Le kombucha s'est répandu en Russie dans les années 1800 et a été reconnu comme une médecine populaire efficace dans de nombreuses communautés rurales. La Russie est donc probablement le berceau du kombucha moderne.

Entre (1913-1995), le kombucha s'est diffusé et a commencé à être une pratique de fermentation au sein du foyer. Durant le XIXe siècle, le kombucha a été

introduit en Europe (Suisse, Danemark, Pologne, Allemagne, Hongrie, Italie, Espagne, France et en Afrique du Nord (les pays gouvernés par la France) (Blanc, 1996 ; Frank et al., 2008). Le kombucha était très populaire en Russie et en Europe jusqu'à la seconde guerre mondiale. En Italie, le kombucha est devenu la boisson favorite des élites de la société (kombucha kamp, n.d.) Des études sur les bienfaits de la consommation régulière du kombucha ont été tenues en Suisse durant les années 1960, puis en Russie, suite à la catastrophe nucléaire de Tchernobyl en 1986, où le kombucha est présenté comme une boisson procurant une meilleure résistance aux polluants (Soares et al., 2021). Le kombucha a été brassé à la maison dans des bocaux à petits volumes, considéré comme une sorte de médecine traditionnelle. Le consortium microbien a été partagé entre les membres des familles et les amis. Une des méthodes adoptées a été le partage de l'inoculum en quatre : un quart est gardé, et le reste est offert à des membres de la communauté (Tadeusz, 2021). Ce qui explique la diffusion rapide du kombucha maison.

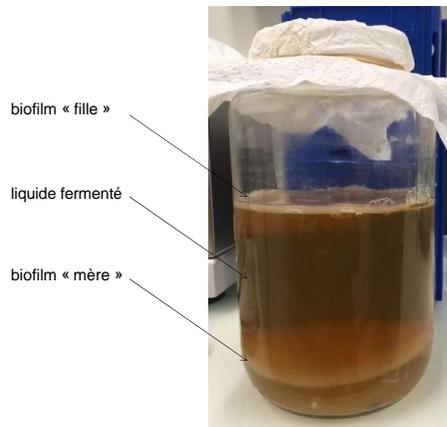
À partir de 1995, le kombucha s'est diffusé en tant que boisson commercialisée. C'est au début des années 1990 que des marques allemandes se sont lancées dans la production et la commercialisation du kombucha.

En parallèle aux États-Unis, la première brasserie commercialisant cette boisson s'est établie en Californie. À cette échelle, le kombucha est devenu une boisson légère et fonctionnelle qui se veut plus agréable que médicinale et une bonne alternative aux sodas pour les consommateurs de boissons non alcoolisées (Tadeusz, 2021).

### 1.3.2. Déroulement d'une fermentation de kombucha

#### 1.3.2.1. Une fermentation dans deux fractions : la phase liquide et le biofilm

Une caractéristique essentielle de la fermentation de kombucha est que l'activité des micro-organismes du SCOBY se déroule dans deux phases (Figure 3) : la fraction liquide issue de l'infusion de thé d'une part et une fraction solide ou biofilm macroscopique d'autre part, qui est formé de cellulose synthétisée par certaines espèces bactériennes du SCOBY (Soares et al., 2021).



**Figure 3. :** Fermentation domestique du kombucha au bout de 14 jours (1.2 litres)

En fin de fermentation, la fraction liquide est utilisée comme boisson mais aussi comme une partie du starter pour initier la fermentation suivante. Elle permet une acidification de la nouvelle infusion de thé sucrée et la propagation d'une partie de la communauté microbienne avec des concentrations qui atteignent  $1.10^8$  CFU (Colony Forming Unit)/ml pour les levures et  $1.10^9$  CFU/ml pour les bactéries (Coton et al., 2017 ; Jayabalan et al., 2014).

Le biofilm est utilisé comme l'autre partie du starter pour reconstituer un SCOBY complet. En effet, il s'agit d'une couche de cellulose translucide qui flotte à l'interface liquide-air et qui se constitue au cours de la fermentation. Son épaisseur finale est d'environ 1 cm, soit 50 g, pour un volume d'un litre de fermentation. Dans les cuves de fermentation, il peut atteindre plusieurs kilogrammes d'où le qualificatif de macroscopique. Il contient des bactéries et des levures avec des concentrations en CFU/g dix fois plus élevées que dans la fraction liquide (Coton

et al., 2017 ; Gramza-Michałowska et al., 2016). Le biofilm utilisé comme partie du starter s'appelle le biofilm “mère”. Pendant la fermentation, ce biofilm “mère” descend progressivement au fond du liquide et la nouvelle couche de biofilm qui se forme à la surface est appelée “fille”. Pour la fermentation suivante, c’est généralement le biofilm “fille” qui est utilisé comme partie du starter lors du processus traditionnel de back-slopping (et de fait, il devient “mère”). Le biofilm “mère” peut également être réutilisé et donner une deuxième “fille”. C’est souvent par ce moyen que les producteurs amateurs partagent leur starter. Selon le livre décrivant les bonnes pratiques d’élaboration du kombucha artisanal (Crum and LaGory, 2016) des observations relatives à l’aspect du biofilm sont décrites : sa couleur varie du blanc au brun clair, il résiste à la déchirure lorsqu'on le presse entre le pouce et l'index.

La fermentation de kombucha peut être effectuée dans le seul but de produire ce biofilm de cellulose (Kiangkitiwan et al., 2022) dont de nombreuses applications comme biopolymère dans les secteurs médicaux, cosmétiques et textiles, notamment, sont à l’étude (Mishra et al., 2022). En effet, la cellulose bactérienne présente une structure en nanofibrilles ainsi qu’une pureté et des propriétés mécaniques plus intéressantes que la cellulose d’origine végétale.

#### 1.3.2.2. Évolution des compositions biochimiques et microbiologiques au cours du processus

La fermentation principale du kombucha, initiée selon le mode traditionnel de back-slopping, dure généralement 15 jours. La composition biochimique du liquide évolue au cours du temps, avec la diminution de la concentration en saccharose en parallèle de l’augmentation transitoire de celles du glucose et du fructose. Leur consommation est associée à la formation d’éthanol dont la concentration augmente pour atteindre son maximum vers le huitième jour. La

concentration des acides organiques, comme l'acide acétique, augmente continuellement et le pH diminue passant d'une valeur initiale de 5,5 à 3. En parallèle, le biofilm-fille s'épaissit à l'interface air-liquide (Malbaša et al., 2008). Les concentrations en micro-organismes dans le liquide augmentent d'un facteur 10 en cours de fermentation avec des variations selon les taxons dépendant de la composition microbienne initiale (Coton et al., 2017).

Il est à noter qu'il n'y a pas de réglementations pour encadrer la production de kombucha commercialisé. Certaines boissons vendues comme kombucha ne sont pas produites en utilisant la méthode traditionnelle de back-slopping mais par refermentation courte, après dilution du liquide d'une fermentation précédente dans une infusion de thé sucré, processus appelé 'one-day' kombucha (<https://www.mannanova.com/manna-k-fr>). Par conséquent, ces boissons peuvent présenter des teneurs en sucres, en différents composés et des compositions en micro-organismes très variables selon le fabricant.

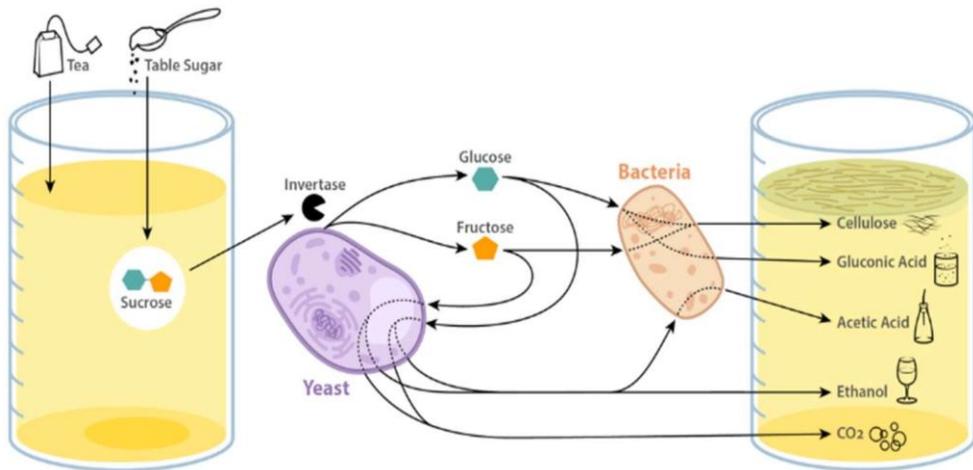
#### 1.3.2.3. Interactions entre micro-organismes

Jusqu'à présent, aucune étude n'a clairement démontré quelles sont les voies métaboliques impliquées dans cette évolution des compositions en métabolites principaux et en micro-organismes au cours de la fermentation de kombucha. La version la plus citée dans la littérature décrit l'utilisation du saccharose par les levures, qui est hydrolysé en glucose et fructose grâce l'invertase qu'elles produisent. Ces études décrivent que les bactéries présentes seraient incapables d'utiliser le saccharose ce qui rend leur activité dépendante de la présence des levures (Tran et al., 2020a ; Villarreal-Soto et al., 2018). Cependant, l'acidification causée par l'ajout du liquide d'une fermentation précédente peut entraîner une hydrolyse du saccharose en sucres simples qui pourrait être suffisante pour initier l'activité des bactéries, ce qui expliquerait que nous avons

observé la formation d'un léger biofilm de cellulose dans une infusion de thé sucréeensemencée avec une culture des seules bactéries (observations non publiées). Les levures utilisent ensuite une partie du fructose et du glucose pour produire de l'éthanol. Les bactéries utilisent les sucres simples en présence d'oxygène, pour produire de la cellulose et des acides organiques et utilisent l'éthanol pour produire de l'acide acétique (Dufresne and Farnworth, 2000). Les interactions entre micro-organismes, illustrées par la Figure 4, ne sont pas encore clairement définies, comme le montre bien les différentes dénominations : le terme SCOBY, qui suggère des interactions de type symbiotique, est largement utilisé, mais le terme KMC (Kombucha Mutualistic Consortium) est également utilisé par certains auteurs (Góes-Neto et al., 2021).

Deux études de (Tran et al., 2022a, 2020b) réalisées sur des combinaisons d'espèces de levures et de bactéries en mono- et co-cultures, suggèrent que les levures stimulent la formation du biofilm. Les hypothèses émises pour expliquer cette stimulation supposent un effet structural soutenu par les formes cellulaires filamenteuses ou pseudo-hyphales observées pour les levures, ainsi qu'un effet métabolique avec apport de nutriments aux bactéries, suite à l'autolyse possible des levures. Ces observations sont certainement dépendantes des espèces présentes.

**Figure 4.** Les interactions métaboliques principales qui seraient impliquées dans la composition finale du kombucha (May et al., 2019)



#### 1.3.2.4. Aromatisation et deuxième fermentation optionnelles

À la fin de la fermentation principale, les biofilms cellulosiques mère et fille sont retirés et une filtration grossière peut avoir lieu afin d'éliminer les éventuelles particules de cellulose. Le liquide généré est transvasé dans des bouteilles en verre étanches. Il existe des étapes supplémentaires, appliquées dans les productions artisanales et industrielles qui visent à améliorer la qualité sensorielle de la boisson. Une pratique courante consiste à prolonger la fermentation en absence d'oxygène. Il s'agit d'une fermentation secondaire réalisée sans l'ajout d'un inoculum, réalisée en bouteille capsulée ayant comme objectif la production d'un maximum de pétillant mais qui s'accompagne d'une augmentation de la teneur en éthanol. L'autre pratique est une étape d'aromatisation durant laquelle des composés aromatiques tels que des fruits, des herbes et des huiles essentielles sont ajoutés à la boisson. L'incorporation totale des arômes dans le bouillon prend entre 2 à 5 jours selon la puissance du composé (Crum and LaGory, 2016). Dans la production industrielle où une limite de teneur en alcool est définie, afin de ne pas être considérée comme boisson alcoolisée, l'aromatisation peut se dérouler sans passer par une étape de refermentation en anaérobiose. Il suffit d'ajouter, en

cuve et à froid, les jus ou les fruits ou les substances aromatiques et d'ajouter du gaz carbonique avant l'étape d'embouteillage afin d'augmenter le pétillant. Ces étapes impactent les compositions en métabolites et les compositions en micro-organismes, notamment si les composés pour l'aromatization sont porteurs d'une microflore indigène.

L'aromatization chez BioBrasseurs se fait par ajout de jus de fruits ou racines pressés à froid (gingembre, curcuma, jus de pomme, d'orange de pêche et de framboise) ou par infusion à froid de boutons de rose, de mauve, d'hibiscus, de café ou de cannelle et d'anis.

### 1.3.3. À quels taxons appartiennent les levures et les bactéries impliqués dans la production de kombucha ?

Le nom du SCOBY réfère à la présence d'une multitude de micro-organismes cohabitant ensemble. L'osmotolérance, la résistance à l'éthanol, à l'acétate et une température de multiplication variant de 20°C à 30°C représentent les caractéristiques rassemblant la communauté active dans un SCOBY. Dans cette partie, nous allons détailler les différents taxons qui ont été détectés dans les études précédemment publiées sur le sujet.

#### 1.3.3.1. Les levures

Les levures détectées dans différents kombucha couvrent un large spectre taxonomique avec des espèces qui appartiennent à la sous-division des Saccharomycotina et à celle des Taphrinomycotina.

#### 1.3.3.1.1. Levures de la famille des Pichiaceae (Saccharomycotina)

Les levures de la famille des Pichiaceae sont systématiquement détectées dans les SCOBY analysés. Les levures mise en évidence appartiennent à deux genres

\*Les levures du genre *Brettanomyces*

Les levures du genre *Brettanomyces* ont été découvertes vers 1903 dans des vieilles bières anglaises (Verachtert and Derdelinckx, 2014). Du point de vue métabolique, le genre *Brettanomyces* consomme le sucre et accumule l'éthanol dans les conditions aérobiques. Il a aussi la capacité de produire une sorte de biofilm ou pseudo hyphes. En absence d'oxygène, la fermentation alcoolique est inhibée et le sucre est converti en acétate (Lebleux et al., 2020). *Brettanomyces* semble présenter une capacité à s'intégrer à différents milieux fermentaires.

Les différentes espèces de *Brettanomyces* qui ont été isolées du kombucha et détectées appartiennent principalement à deux espèces : *B. bruxellensis* et *B. anomalus* (Coton et al., 2017 ; Grassi et al., 2022 ; Marsh et al., 2014 ; Reva et al., 2015). *B. bruxellensis*, détecté dans une large gamme de produits fermentés, est étudié surtout car il cause des altérations organoleptiques dans le vin . Les phénols volatils produits à partir des acides hydroxycinnamiques présents dans la pellicule des raisins donnent une odeur désagréable de cuir ou de sueur de cheval aux vins (Suárez et al., 2007). Ces mêmes composés aromatiques, essentiellement l'éthyl phénol, sont appréciés dans d'autres boissons telles que la bière de style lambic et le kombucha auxquels ils procurent une odeur typique. Le séquençage du génome entier de nombreux isolats a montré une variabilité au niveau de la composition des génomes avec différents niveaux de ploïdie (diploïdes, autotriploïdes et allotriploïdes) et différents événements de polyploïdisation (Eberlein et al., 2021 ; Gounot et al., 2020).

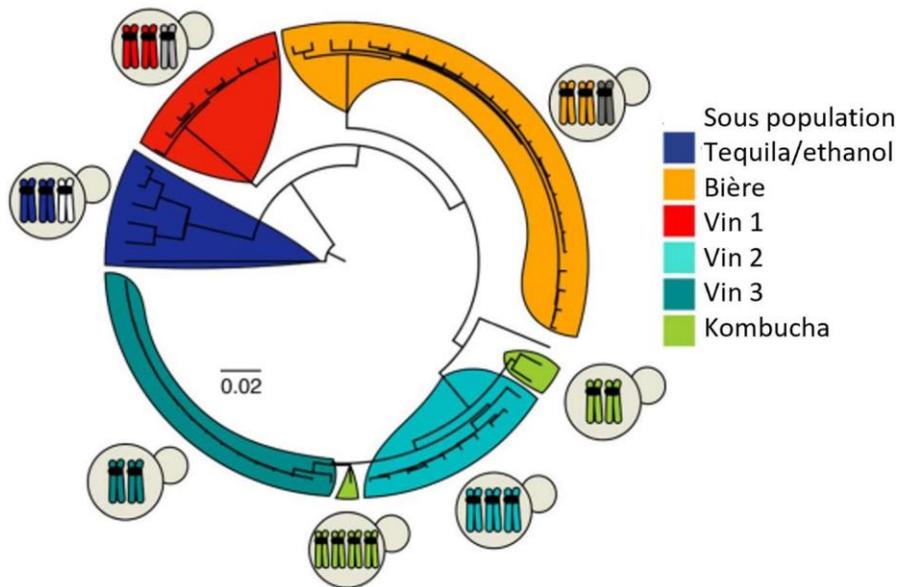
L'ensemble des 71 isolats caractérisés de cette espèce se répartissent en six sous-populations qui correspondent à différents milieux anthropiques nommées à partir

de rapports précédents (Avramova et al., 2018 ; Gounot et al., 2020) vin, bière, tequila/bioéthanol, kombucha illustrés par la (Figure 5). Il semble que les divers niveaux de ploïdie permettent de répondre à une adaptation aux exigences du milieu fermentaire (Eberlein et al., 2021). Il a été souligné que le niveau de ploïdie est conservé au sein des sous-populations génétiquement divergentes et chaque sous-population est caractérisée par un événement de polyploïdisation qui lui est propre.

La structure de la population est principalement expliquée par le niveau de ploïdie (~47% de la variance), suivi par l'origine géographique (~5% de la variance) et finalement par l'origine écologique (~6% de la variance) (Avramova et al., 2018). Cependant, en considérant les isolats non viticoles, l'origine géographique explique ~55% de la variance totale indiquant que les isolats du vin sont fortement disséminés à travers le monde par rapport aux souches isolées d'autres substrats. Quant aux souches de *B. bruxellensis* isolé du kombucha, il a été mis en évidence en séquençant le génome entier de 3 souches que deux appartenaient à une même sous population diploïde et une à une sous population tétraploïde (Eberlein et al., 2021).

\*Les levures du genre *Pichia*

Le genre *Pichia* est également souvent représenté dans différentes boissons fermentées et il a la particularité de former une couche à la surface de vin (Comuzzo and Battistutta, 2019).



**Figure 5.** Diversité génétique au sein des espèces de *Brettanomyces bruxellensis* : Les six sous-populations sont représentées par différentes couleurs sur l'arbre qui a été construite à partir des courtes lectures générées par un séquençage illumina du génome entier des 71 isolats de *B. bruxellensis* alignées sur le génome de référence *B. bruxellensis* (Fournier et al., 2017) et les 24 313 variants génétiques répartis uniformément sur le génome (Eberlein et al., 2021). Le niveau de ploïdie de chaque population a été schématisé, trois chromosomes représentent les triploïdes et deux chromosomes représentent les groupes diploïdes, Les chromosomes gris clair, gris foncé et blancs représentent les génomes indépendants haploïdes acquis d'espèces sœurs inconnues.

C'est peut-être cette caractéristique qui explique sa présence sur un support cellulosique dans une fermentation de type kombucha. C'est majoritairement l'espèce *Pichia membranefaciens* qui a été détectée (Coton et al., 2017) en plus d'autres espèces minoritaires comme *P. occidentalis*. Les levures de cette famille présentes dans les SCOBY appartiennent également aux espèces *Kregervanrija fluxuum* et *Kregervanrija pseudodelftensis*, qui ont été définies à partir d'un même taxon anciennement nommé *Pichia fluxuum*, après séquençage des régions ITS1 et ITS2 de l'ADNr (Kurtzman, 2006).

#### 1.3.3.1.2. Levures de la famille des Saccharomycetaceae (Saccharomycotina)

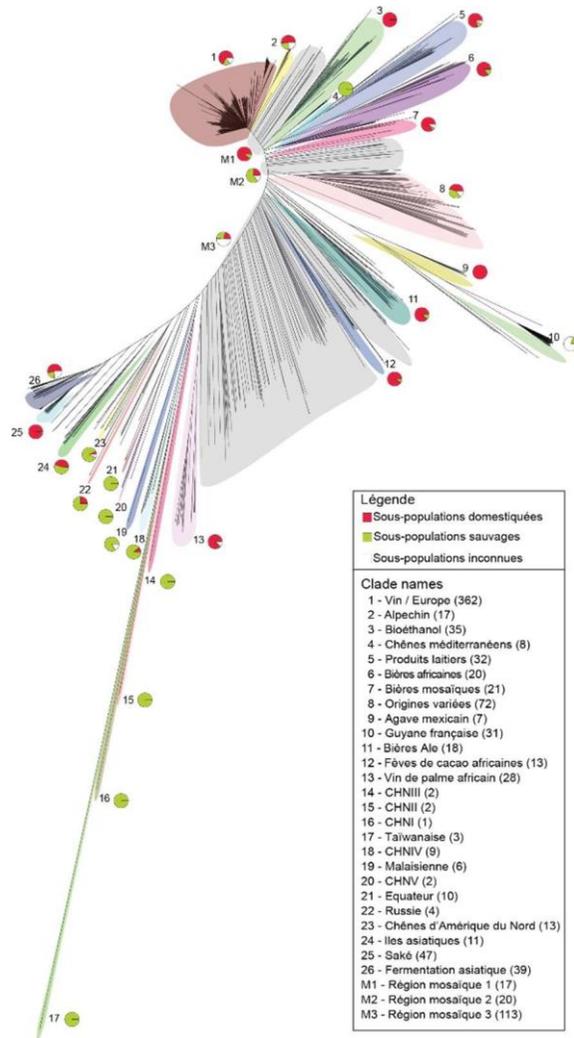
Dans la famille des Saccharomycetaceae, trois genres sont détectés dans les SCOBY : le genre *Saccharomyces*, le genre *Zygosaccharomyces* et le genre *Candida*.

##### \*Les levures du genre *Saccharomyces*

L'espèce modèle *Saccharomyces cerevisiae* est régulièrement détectée dans les SCOBY (Liu et al., 1996). Cette espèce est utilisée pour de nombreux processus de fermentation allant de la fermentation du vin à celle de la bière en passant par le pain et les fromages. Pour cette espèce, la séquence génomique de plus de 2000 isolats est disponible, dont plus de 1000 ont été séquencés et analysés dans mon équipe d'accueil (Peter et al., 2018). Cette étude a révélé que ces isolats sont répartis dans 26 sous-populations qui se distinguent soit par leur origine géographique, dans le cas des isolats de la nature, soit par leur origine écologique (Figure 6). Il serait intéressant de savoir si les isolats de *S. cerevisiae* isolés de kombucha sont regroupés dans l'une de ces sous-populations ou s'ils sont distribués dans plusieurs sous-populations.

##### \*Les levures du genre *Zygosaccharomyces*

Le genre *Zygosaccharomyces* a également été identifié dans plusieurs études (Marsh et al., 2014 ; Sievers et al., 1995), et majoritairement dans la dernière étude avec une abondance relativement importante supérieure à 95% dans certains échantillons (Marsh et al., 2014). Différentes espèces ont été caractérisées comme *Z. rouxii*, *Z. bailii* et *Z. lentus* (Jayabalan et al., 2008 ; Liu et al., 1996 ; Steels et al., 1999). L'espèce *Z. lentus* en particulier est caractérisée pour son osmotolérance, sa capacité à croître à des valeurs basses de pH. Elle est considérée comme un contaminant des jus de fruits (Steels et al., 1999) contrairement au kombucha où l'environnement acide semble propice pour son développement.



**Figure 6.** Arbre phylogénétique, de type Neighbour Joining, basé sur la diversité nucléotidique entre 1011 isolats naturels de *Saccharomyces cerevisiae*. Plusieurs sous-populations domestiquées ou sauvages sont retrouvées et réparties au sein de 26 clades spécifiques et 3 régions mosaïques. Une matrice des distances établie entre les 1 544 489 sites bialléliques des 1011 souches a été utilisée pour construire cet arbre. Figure adaptée de Peter et al., 2018.

#### \*Les levures du genre *Candida*

Le genre *Candida* a été détecté dans un grand nombre de boissons à base de kombucha et majoritairement dans l'étude de (Chakravorty et al., 2016).

Différentes espèces ont été isolées *C. obutsa*, *C. famata*, *C. guilliermondii*, *C. colliculosa*, *C. kefir*, *C. krusei*, et *C. stellata* (Ramadai and Abulreesh, 2010 ; Teoh et al., 2004) à partir de différentes origines géographiques : Mexique, Arabie Saoudite, Australie et Turquie. Dans l'optique d'une reclassification des espèces anamorphiques dans les genres typés téléomorphes (Englezos et al., 2017), le genre *Starmerella* a été décrit pour accueillir l'état sexuel de *Candida* (Lachance, 2011).

D'autres espèces de Saccharomycetaceae ont été détectées dans du kombucha, tels que *Torulaspora delbrueckii*, *Kluyveromyces africanus*, *Hanseniaspora valbyensis* (Coton et al., 2017 ; Jayabalan et al., 2014 ; Teoh et al., 2004). Cette dernière est commune à la fermentation du cidre. *H. valbyensis* produit davantage d'acétate d'éthyle et d'acétate de phényl éthyl que *S. cerevisiae*, à l'origine respectivement, d'arômes fruités et de rose (Xu et al., 2006).

#### 1.3.3.1.3. Levures de la famille des Schizosaccharomycetaceae (Taphrinomycotina)

Elles sont représentées par l'espèce *Schizosaccharomyces pombe* qui a été détectée dans quelques études (Reiss, 1994 ; Teoh et al., 2004). Cette espèce est caractérisée par un fort pouvoir fermentaire et sa libération de quantités importantes de polysaccharides (Comitini et al., 2011).

### 1.3.3.2. Les bactéries

Les bactéries présentes dans les SCOBY de kombucha sont des Protéobactéries de deux groupes : les Firmicutes et les Alphaprotéobactéries. La présence des bactéries dans un consortium de kombucha est indispensable afin de produire les acides organiques et procurer à la boisson le goût acidulé caractéristique.

#### 1.3.3.2.1. Les bactéries du groupe des Alphaprotéobactéries.

Pour la fermentation du kombucha, il s'agit de bactéries dites acétiques, qui sont des bactéries à Gram négatif, de la famille des Acetobacteraceae. Le genre type de la famille, où dix-sept genres sont signalés, est *Acetobacter*. Quant au genre *Gluconobacter*, il est présent dans deux groupes taxonomiques selon leur préférence métabolique : l'un a la capacité d'oxydation de l'éthanol plus que le glucose et oxyde l'acide acétique en dioxyde de carbone et en eau. Le second groupe a la capacité d'oxydation du glucose plutôt que l'éthanol et il n'oxyde pas l'acide acétique. Les souches du genre *Neokomagataea* étaient phylogénétiquement liées à celles du genre *Gluconobacter*. Les souches de ce genre se distinguent du genre *Gluconobacter* car leur croissance est réprimée en présence de 0,35 % d'acide acétique (v/v) (Yamada et al., 2012).

Sur le plan phylogénétique, le genre *Gluconacetobacter* se compose de deux groupes : le groupe *Gluconacetobacter liquefaciens* et le groupe *Gluconacetobacter xylinus*. Pour ce dernier groupe, le genre *Komagataeibacter* a été proposé (Yamada et al., 2012) sur la base de l'analyse de la séquence du gène de l'ARNr 16S et des caractérisations morphologiques, physiologiques et écologiques. Onze espèces du genre *Gluconacetobacter* ont été transférées dans le nouveau genre *Komagataeibacter* (Matsushita et al., 2016).

#### 1.3.3.2.2. Les bactéries du groupe des Firmicutes

Dans les SCOBY, on peut trouver des bactéries appelées bactéries lactiques. Ces bactéries lactiques sont présentes majoritairement par les deux familles : les Lactobacillaceae et les Leuconostocaceae, qui sont des bactéries à Gram positif. Elles sont généralement considérées comme des aérobies facultatifs avec un métabolisme fermentatif (Ricciardi et al., 2014). Elles peuvent être classées en fonction de leur voie fermentaire, homofermentaire où elles ne forment que le lactate à partir du glucose. La deuxième voie est hétérofermentaire obligatoire ou facultative où elles produisent de l'acétate et l'éthanol en plus du lactate. Le métabolisme des hexoses est différent pour chaque voie fermentaire et génèrent différents métabolites intermédiaires. Afin de réaliser son métabolisme, les bactéries hétérofermentaires ont diverses préférences par rapport au sucre comme substrat sous forme d'hexoses fructose / maltose ou de disaccharides tel que le saccharose (Gänzle, 2022).

Au niveau de la famille, le métabolisme homo ou hétérofermentaire des bactéries lactiques est conservé (Zheng et al., 2020). En se basant sur les données de séquences du génome entier, 23 nouveaux genres issus de *Lactobacillus* ont été classés en *Paralactobacillus*. Cette reclassification a touché l'espèce *Lactobacillus delbrueckii* qui a été divisée en 25 groupes. Le genre *Oenococcus* a été détecté dans du kombucha, il fait partie de la famille des Leuconostocaceae, et suit un métabolisme hétérofermentaire (Coton et al., 2017).

#### 1.3.4. Conditions et paramètres de fermentation

Les recettes, les volumes et les procédés de fermentation du kombucha sont variables entre les producteurs amateurs ou pour la commercialisation, mais également dans la littérature scientifique. Une vue d'ensemble de ces différents paramètres était nécessaire pour définir les conditions de nos études. Ils sont

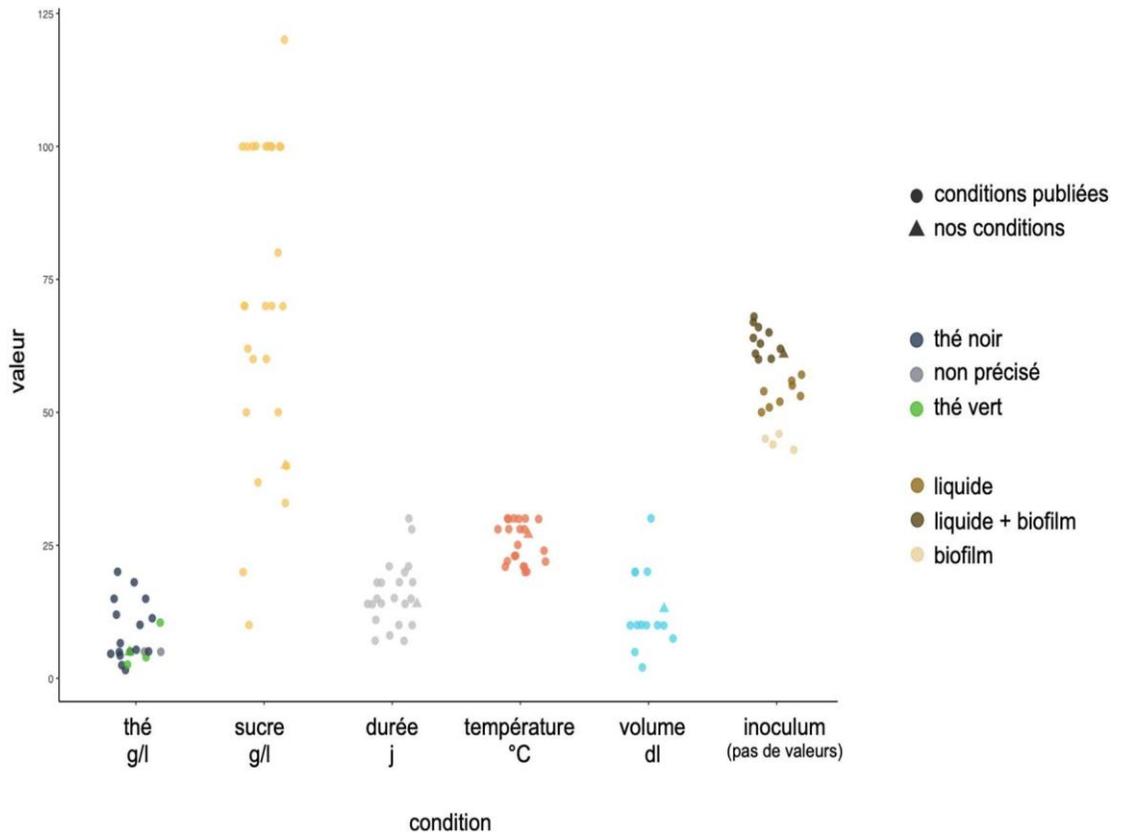
décrits de manière synthétique ci-dessous, en comparaison avec les paramètres que nous avons utilisés pour nos études et qui seront détaillés au chapitre suivant.

#### 1.3.4.1. Récipient de fermentation, volume et géométrie : du bocal à la cuve des brasseurs

Les récipients utilisés pour la fermentation doivent résister aux acides organiques (Crum and LaGory, 2016). À l'échelle domestique, la fermentation se déroule dans un bocal en verre d'une contenance de 1 à 2 litres et chez les producteurs, le volume des cuves de fermentation en acier inoxydable varie de 100 à 4000 litres (Arikan et al., 2020 ; Teoh et al., 2004). Pour les études en laboratoire, le volume des fermentations étudiées varie entre 200 ml et 3 litres (Figure 7). Dans tous les cas, une large ouverture est nécessaire pour manipuler facilement le biofilm et pour un échange d'air suffisant avec l'environnement. L'ouverture des récipients et des cuves est juste recouverte d'un tissu de coton propre ou d'un papier hygiénique pour une protection contre les poussières et les insectes (Dutta and Paul, 2019).

Une étude menée en mesurant la vitesse de fermentation du kombucha dans différents volumes a pointé l'importance de la surface interfaciale spécifique qui permet de s'assurer du bon déroulement de la cinétique d'acidification bactérienne et du dégagement suffisant de dioxyde de carbone généré par les levures. Elle a été calculée, dans une étude simulant différentes dimensions de réacteurs lors de l'augmentation en échelle d'une fermentation de kombucha. La formule tient en compte la surface du liquide par rapport au volume occupé par le liquide (Cvetković et al., 2008).

En prenant en considération la géométrie des bocaux (cf 2.2.1) et les données dans l'étude de (Cvetković et al., 2008), le volume du liquide dans notre étude est de 1,3 L environ, légèrement supérieur à la médiane.



**Figure 7** : Vue d'ensemble des différents paramètres appliqués aux fermentations de kombucha de différentes études (données de (Laavanya et al., 2021)). Les valeurs ou caractéristiques utilisées dans nos études sont également indiquées (triangles).

#### 1.3.4.2. Les substrats de fermentation

Le kombucha est traditionnellement produit à partir d'une infusion de feuilles de thé (*Camellia sinensis*). Le thé vert ou le thé noir (dont les feuilles ont subi une plus forte oxydation) sont utilisés. Dans le cadre des études de la fermentation du kombucha, c'est le thé noir qui a été le plus utilisé (Figure 7). Il a été décrit que le type de thé influence la fermentation à cause des variations dans la composition qui peuvent modifier la prédominance des bactéries dans le kombucha. Le thé

noir contient plus d'azote et de purine que le thé vert, les tanins supplémentaires du thé vert donnent un goût légèrement amer (Coton et al., 2017 ; Drewnowski and Gomez-Carneros, 2000). Les teneurs en thé noir ou vert utilisées dans différentes études varient entre 2,5 g et 20 g par litre d'infusion (Figure 7). Il a également été décrit que lorsque la concentration de feuilles de thé dépasse 6 g/L, elle inhibe la croissance des bactéries acétiques et diminue donc la production de cellulose (Gaggia et al., 2018). Dans nos études, nous utilisons les mêmes conditions que BioBrasseurs à savoir du thé vert à raison de 5 g/l. Pour avoir une infusion de thé, une température d'infusion entre 70 et 95 °C est recommandée afin de libérer les arômes et augmenter l'activité antioxydante du thé (Dufresne and Farnworth, 2000).

D'autres types de thé ou de végétaux infusés peuvent également être utilisés : le thé blanc, Puer, le thé vert japonais, le thé au jasmin, le thé oolong, le thé de mûrier, les thés à la sauge, au thym et à la menthe, l'echinacea (Cvetković et al., 2008 ; Leal et al., 2018 ; Saimaiti et al., 2022). Des jus de fruits ont été ajoutés pour augmenter les effets antioxydants du kombucha. Des liquides non conventionnels ont également été testés, comme le coca cola, le vin rouge, le vin blanc, le vinaigre, l'extrait de topinambour, le lait, le petit-lait sucré frais, le petit-lait sucré reconstitué (Barakat et al., 2022 ; Emiljanowicz and Malinowska-Pańczyk, 2020). Ces substrats alternatifs sont utilisés pour produire des boissons mais aussi des biofilms de nanocellulose bactérienne (Gupte et al., 2021).

Le sucre principalement utilisé est le saccharose dont la concentration varie entre 10 et 120 g par litre selon les études (Figure 7). Le saccharose issu de la canne est souvent utilisé mais d'autres sources et d'autres sucres sont également testés comme le miel ou la mélasse de betterave à sucre (Malbaša et al., 2008).

#### 1.3.4.3. Inoculum pour la fermentation du kombucha

La fermentation typique du kombucha nécessite une étape d'inoculation avec un starter naturel issu des produits d'une fermentation précédente. Les produits de fin de fermentation sont constitués par le biofilm et la fraction liquide qui est utilisée pour la boisson. On peut distinguer trois types de starters naturels : (i) la méthode traditionnelle et la plus pratiquée (Figure 7) utilise les deux fractions, liquide et solide d'une même fermentation de kombucha (Dufresne and Farnworth, 2000 ; Jayabalan et al., 2008 ; Manuel et al., 2014). Les proportions sont généralement de 10 % v/v pour la fraction liquide et de 2% w/v pour le biofilm. (ii) La fraction solide seule, constituée du biofilm cellulosique peut être utilisée, avec des proportions de 2 à 7 % w/v. (iii) La dernière forme d'inoculum peut être constituée de la fraction liquide seule avec des proportions de 10 à 16 % (v/v).

En laboratoire, quelques études récentes ont montré la possibilité d'élaborer des consortia synthétiques composés d'un mélange de quelques espèces de levures et bactéries isolées de la même culture de kombucha pour effectuer des fermentations de type kombucha (Savary et al., 2021 ; Tran et al., 2020a).

#### 1.3.4.3. Durée et conditions de température

La fermentation du kombucha a été pratiquée à l'échelle maison depuis des milliers d'années et prend généralement 7 à 17 jours à température ambiante. En utilisant les données collectées (Figure 7), la médiane décrivant la tendance centrale est de 15 jours (Laavanya et al., 2021). Durant l'élaboration du kombucha, une température constante tout au long du process est recommandée. Elle est définie entre 20 et 30°C afin d'assurer les activités des différents micro-organismes (Figure 7). Dans le cadre de notre étude, nous avons opté pour une

durée de 14 jours et une température constante de 27°C, suivant bien la tendance des autres études tout en s'adaptant aux paramètres utilisés chez BioBrasseurs.

## **2. Les contributions des méthodes de séquençage haut débit à la compréhension des communautés microbiennes impliquées dans les processus fermentaires**

L'augmentation des connaissances sur les communautés microbiennes est liée à l'évolution de nombreuses méthodes permettant des analyses à faible, moyen ou haut débit. Je vais donc présenter ci-après, un inventaire de ces méthodes, en orientant cette présentation autour de leur évolution dans le temps et mettant l'accent sur les méthodes de séquençage à haut débit puisqu'elles ont été mises en œuvre dans les différentes études de ce travail de thèse. Les exemples cités seront dans les contextes des communautés environnementales et des microbiotes humains mais aussi dans le contexte des processus fermentaires et en particulier en relation avec la fermentation du kombucha.

Ces méthodes se distinguent par une combinaison de caractéristiques à différents niveaux. On peut distinguer les méthodes dépendantes de la culture, des méthodes indépendantes de la culture. Ces deux catégories peuvent se décliner en méthodes qui visent à la détection d'un taxon particulier ou la détection exhaustive des taxons présents, ainsi qu'en méthodes qualitatives et/ou quantitatives. Il en résulte des niveaux de précision des résultats très variables, de la simple détection d'un taxon précis à sa quantification dans une communauté jusqu'à la description des fonctions (gènes fonctionnels) présents dans une niche écologique. Bien sûr, le terme de taxon lui-même peut représenter des niveaux divers de la taxonomie, du phylum à la souche, selon les méthodes mises en œuvre. Enfin, on peut également caractériser ces méthodes selon le nombre d'échantillons qui peuvent être traités

simultanément : on parle de faible, moyen et haut débit selon la complexité des équipements et des traitements de données nécessaires aux analyses.

## 2.1. Identification directe par observations au microscope

La méthode la plus directe et certainement la méthode historique, qui permet d'explorer la diversité microbienne d'un prélèvement, est l'observation des cellules au microscope. À partir d'un prélèvement *in situ*, dans un environnement naturel ou un milieu de fermentation, ces observations microscopiques simples (état frais) ou après fixation et coloration permettent une première évaluation de la complexité d'une communauté microbienne. Une identification morphologique et une numération peuvent être effectuées. Elles permettent de distinguer différentes catégories de protistes, bactéries, levures et moisissures mais leur résolution pour l'identification des rangs taxonomique est très faible.

## 2.2. Caractérisation par des méthodes dépendantes de la culture

### 2.2.1. Isolement

Les méthodes dites "dépendantes de la culture" impliquent de pouvoir isoler les différents micro-organismes d'un prélèvement et les faire sur des milieux de culture, synthétiques ou complexes, qui répondent à leurs besoins nutritionnels. Ces cultures microbiennes peuvent avoir un dénombrement de la flore totale comme objectif ou bien elles peuvent viser l'identification d'un micro-organisme particulier. Les isolements nécessitent soit l'utilisation de milieux de culture très sélectifs permettant la multiplication de taxons précis de micro-organismes et l'identification de leur(s) rang(s) taxonomique(s), comme par exemple la gélose Chapman ou MSA (Mannitol Salt Agar) qui est un milieu d'isolement sélectif utilisé pour la recherche des bactéries du genre *Staphylococcus* (Gebremeskel et al., 2022) soit l'utilisation de milieux pas ou peu sélectifs et l'identification est

effectuée dans un second temps, par des méthodes dont certaines vont être décrites ci-dessous.

Une des limitations principales de ces méthodes est le fait que, lorsque les cellules sont isolées de leur communauté, de nombreux micro-organismes sont considérés comme non-cultivables. Les raisons de cette incapacité à se multiplier sont multiples. Soit il n'a pas été possible de reconstituer un milieu nutritif ou des conditions leur permettant de se multiplier à l'état isolé, soit les cellules présentent une perte de la cultivabilité car elles sont dans un état métabolique particulier appelé VBNC (Viable But Non Culturable) ou dans une forme dormante, qui sont peu ou pas propices à initier leur multiplication à l'état isolé sur un milieu standard (Li et al., 2014). Leur multiplication peut aussi être dépendante de la présence ou du contact avec d'autres cellules ou micro-organismes, conditions difficiles à reproduire sur un milieu d'étude.

Pour donner une idée de l'importance de cette limitation, on peut citer le chiffre suivant (Lloyd et al., 2018) : la part de diversité des micro-organismes non cultivés, également appelée la 'matière noire microbiologique', représenterait jusqu'à 81 % des taxons identifiés à partir des bases de données, pour les seules bactéries. Il faut cependant noter que cette étude ne précise pas si ces taxons ne sont effectivement pas cultivables ou s'ils n'ont pas été mis en évidence en utilisant des méthodes dépendantes de la culture. Dans le contexte de l'exploration d'une communauté microbienne de composition pas ou peu connue, utiliser uniquement des approches dépendantes de la culture pose donc le problème de ne pas détecter certains taxons ou de ne pas être en mesure d'estimer justement leur biomasse relative.

### 2.2.2. Caractérisations par des méthodes dépendantes de la culture

Plusieurs méthodes de caractérisation peuvent être elles aussi dépendantes de la culture. Elles permettent de déterminer des caractéristiques physiologiques et biochimiques dans des milieux de culture appropriés. Les profils typiques de ces caractéristiques permettent l'identification des rangs taxonomiques. Les milieux utilisés permettent de tester la multiplication dans différentes conditions de température, de pH, de disponibilité de nutriments, l'utilisation de certaines sources de carbone, par exemple. La caractérisation biochimique est liée à l'assimilation ou la réduction de composés organiques et minéraux. L'identification des isolats est effectuée grâce à des isolats de référence ou des clés taxonomiques, comme celle des levures basée sur les travaux de Kurtzman et al (Kurtzman et al., 2011). L'utilisation systèmes de test commercialisés, miniaturisés et automatisables (galeries API<sup>®</sup>, Appareils et Procédés d'Identification ; Omnilog<sup>™</sup>) permettent d'augmenter le débit de traitement pour l'identification d'un certain nombre d'espèces (quelques centaines).

Ces méthodes d'identification présentent une limitation liée au très faible débit, en regard du travail laborieux et du coût élevé, qui impliquent qu'un nombre d'échantillons peut être traité et/ou un nombre restreint d'espèces peut être identifié et quantifié, elles sont donc peu appropriées à la détermination exhaustive des compositions microbiennes des communautés.

### 2.2.3. Identification par des méthodes moléculaires

On citera dans cette partie les deux principales méthodes moléculaires appliquées sur le matériel génétique ou protéique extraits à partir de micro-organismes isolés et donc nécessairement cultivables. Ces deux méthodes sont moins laborieuses que les méthodes dépendantes de la culture mais elles restent limitées à un nombre

modéré d'isolats puisque la caractérisation nécessite de manipuler individuellement chaque isolat.

#### 2.2.3.1. Par spectrométrie de masse de type MALDI-TOF (Matrix Assisted Laser Desorption Ionization-Time Of Flight)

Cette méthode est applicable aux micro-organismes isolés de nombreuses espèces de bactéries et de levures et moisissures. Pour identifier l'espèce d'un micro-organisme par MALDI-TOF/MS, il faut obtenir un profil (ou empreinte) réalisé à partir de ses protéines totales, qui sont ionisées à l'aide d'un rayon laser et séparées dans un champ électrique selon des temps de vol dont la durée dépend essentiellement de leur masse. Chaque espèce a sa propre empreinte "signature" et l'identification est effectuée par comparaison du profil d'empreinte de masse spectrale obtenu, avec ceux des bases de données (Barreau et al., 2013). Ces banques doivent être exhaustives et ce qui n'est pas toujours le cas si un nouvel environnement est exploré pour lesquels peu de spectres de référence sont disponibles.

#### 2.2.3.2. Par séquençage d'amplicons de régions caractéristiques

Pour identifier l'espèce d'un micro-organisme isolé, il est possible d'amplifier par PCR et de séquencer une région du génome qui sert de marqueur taxonomique (voir ci-dessous pour plus amples détails). La méthode de séquençage de type Sanger est appliquée.

### 2.3. Caractérisation par des méthodes métagénomiques

Les méthodes métagénomiques désignent des méthodes d'étude qui permettent d'explorer la biodiversité d'un environnement donné directement à partir du matériel génétique extrait des communautés microbiennes prélevées *in situ* dans leur environnement (naturel ou milieu de fermentation), sans passer par les étapes

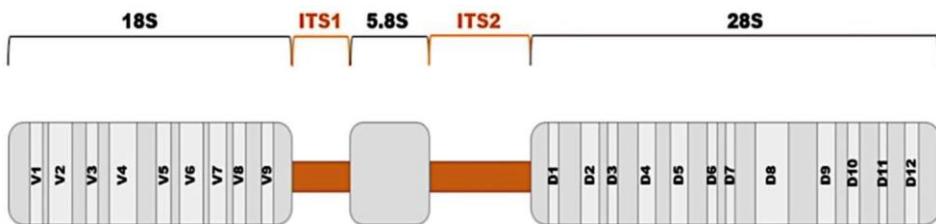
d'isolement et de culture. Il existe deux types d'approches : (i) la caractérisation ou le séquençage d'un ou plusieurs gènes particuliers, appelés marqueurs ou barcodes, communs à plusieurs espèces, on parle alors de métagénomique ciblée ou métabarcoding. (ii) Les méthodes non ciblées avec le séquençage des génomes de tous les organismes présents dans le milieu à analyser, sans distinction des espèces (bactéries, virus, champignons...) : on parle alors de métagénomique shotgun. Nous verrons comment les méthodes historiques ont évolué en méthodes à haut débit grâce aux avancées des stratégies de séquençage dites de nouvelle génération.

### 2.3.1. Méthodes ciblées

Ces méthodes permettent d'identifier les rangs taxonomiques à partir des caractéristiques d'une région particulière du génome, appelée marqueur ou 'barcode'. Idéalement, la région qui peut servir de marqueur doit être commune aux individus du plus haut rang taxonomique possible, d'une façon à assurer une couverture maximale de taxons avec un seul marqueur. Cette région doit également être suffisamment variable entre les taxons de rang inférieurs (idéalement entre les espèces du même genre voir entre les souches de la même espèce) pour lui conférer un pouvoir discriminant. Un bon marqueur est également informatif à l'échelle phylogénétique : par exemple, les différences entre espèces du même genre doivent être inférieures aux différences entre espèces de genres différents. Ceci permet l'assignation des taxons inconnus au taxon d'ordre supérieur : par exemple si un OTU détecté correspond à une espèce qui n'est pas (encore) répertoriée dans les banques de données, il est possible de lui assigner un genre ou au moins une famille.

### 2.3.1.1. Exemples de marqueurs utilisés

La région des gènes codant pour les ARN ribosomiques (ARNr) est le marqueur le plus utilisé. En effet, ces gènes sont universels dans le monde vivant et présentent l'avantage d'être présents en plusieurs copies ce qui augmente la sensibilité de leur détection. Ils sont organisés en opéron dont le nombre de copie est variable : entre une et 15 copies chez les bactéries (Espejo and Plaza, 2018) alors que chez les champignons, les copies sont organisées en tandem et leur nombre peut atteindre plusieurs centaines (Sipiczki, 2022). Les deux plus grands gènes de l'opéron codent (i) l'ARN de la petite sous-unité du ribosome (ARNr 16S chez les bactéries et 18S chez les eucaryotes), (ii) l'ARN de la grande sous-unité du ribosome (ARNr 23S chez les bactéries et 25-28S chez les eucaryotes). Ces gènes sont séparés par un ou deux segments intergéniques ITS (Internal Transcribed Spacer) de longueur et de séquence très variables. Lorsqu'il y a deux ITS, ils sont séparés par le gène de l'ARNr 5.8S chez les eucaryotes et par un gène d'ARNt et/ou le gène de l'ARNr 5S chez les procaryotes (Figure 1.8). Les gènes des grands ARNr contiennent une succession des domaines hautement conservés séparés par des domaines hypervariables : de V1 à V9 et de D1 à D12 pour les ARNr des petites et grandes sous-unités, respectivement (Figure 8).



**Figure 8.** Structure typique d'un opéron eucaryote codant les ARN ribosomiques 18S, 28S et 5.8S, avec la position des segments hypervariables V1 à V9 et D1 à D12 (adapté de D'Andrea et al., 2021)

Les gènes entiers peuvent servir de marqueur mais le plus souvent le marqueur utilisé se limite à un ou deux segments hypervariables consécutifs, pour des raisons de limitation de taille de la méthode de détection. Pour les levures, c'est le segment d'environ 650 pb contenant les régions D1 et D2 du gène de l'ARNr 25S qui le plus informatif à l'échelle phylogénétique (Kurtzman and Robnett, 1998). Pour les bactéries et les archées, les études utilisent différentes régions hypervariables du gène de l'ARNr 16S. Leur choix dépend du type d'environnement analysé car le pouvoir de résolution des différents segments est variable selon les espèces. Par exemple le segment V1-V2-V3 a été utilisé dans le projet conduit par The Human Microbiome Project Consortium (Huttenhower et al., 2012), le segment V1-V2 a été également utilisé pour des études de la constitution bactérienne du microbiote intestinal (Pop et al., 2014). Le segment V6 a été utilisé pour les bactéries et archées marines (Huber et al., 2007) et tout récemment, les segments V7-V8 ont été utilisés pour les bactéries du kombucha (Grassi et al., 2022). Cependant, pour l'identification des bactéries présentes dans les processus de fermentation, ce sont souvent les segments V2-V3, V3-V4 et V4 qui sont utilisés (Bukin et al., 2019 ; Liu et al., 2020 ; Niraj and Shardendu, 2020 ; Walters et al., 2015). Enfin, le segment V9 de l'ARN 18S eucaryote a aussi été utilisé comme marqueur (Amaral-Zettler et al., 2009).

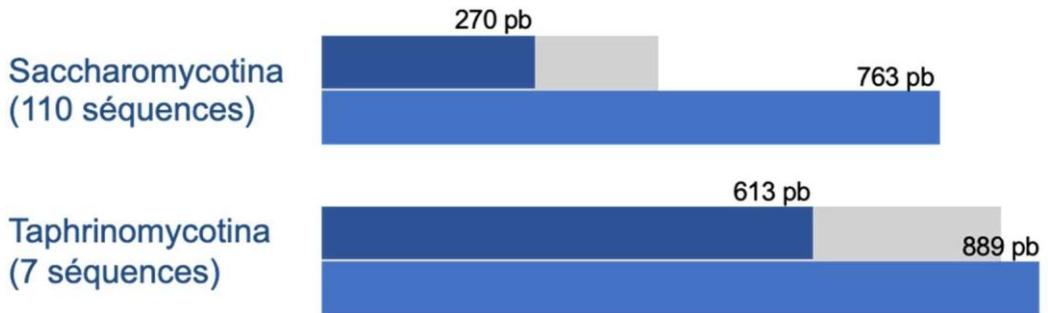
Les régions ITS sont également utilisées comme marqueur, en particulier chez les champignons pour lesquels elles permettent d'identifier l'espèce (Seifert, 2009). La région ITS complète peut être utilisée (segment ITS1-ARNr 5.8S-ITS2) ou les segments ITS2 ou ITS1 peuvent être utilisés seuls (Blaalid et al., 2013).

Il a été mis en évidence que les amorces amplifiant le segment ITS1 n'étaient pas assez consensuelles. ITS1 était une cible préférentielle grâce aux courtes séquences d'amplicons (Bokulich et Mills 2013).

Contrairement aux segments des gènes d'ARNr, la longueur des segments ITS est très variable selon l'espèce (Figure 9), ce qui est à la fois un avantage pour l'identification mais aussi un inconvénient, notamment pour les méthodes basées

sur une amplification par PCR (Nagy et al., 2012).

Chez les levures, un marqueur de longueur constante, le segment D1-D2 (660 pb) du gène d'ARNr 25S est également très utilisé (Fell et al., 2000 ; Kurtzman et al., 2011 ; Vu et al., 2016).



**Figure 9.** Tailles minimales (rectangles bleu foncé) et maximales (rectangles bleus), en pb, de la région ITS complète, répertoriées chez 110 espèces de la sous-division des *Saccharomycotina* et 7 espèces de la sous-division des *Taphrinomycotina* (d'après Seifert, 2009). Les rectangles gris indiquent les tailles moyennes.

D'autres régions non ribosomiques peuvent servir de cible et être utilisées comme marqueurs moléculaires, nous pouvons citer certains marqueurs qui ont été utilisé au sein des phyla des procaryotes et des eucaryotes, comme les gènes de la sous-unité B de l'ADN gyrase (*gyrB*), la sous-unité B de l'ARN polymérase (*rpoB*), du facteur d'élongation TU (*tuf*), de l'ADN recombinase (*recA*), de la sous-unité D de la dinitrogénase alpha (*nifD*) (Banasiewicz et al., 2021 ; Ćurčić et al., 2021 ; Guzman et al., 2022 ; Ogier et al., 2019 ; Zampieri et al., 2021).

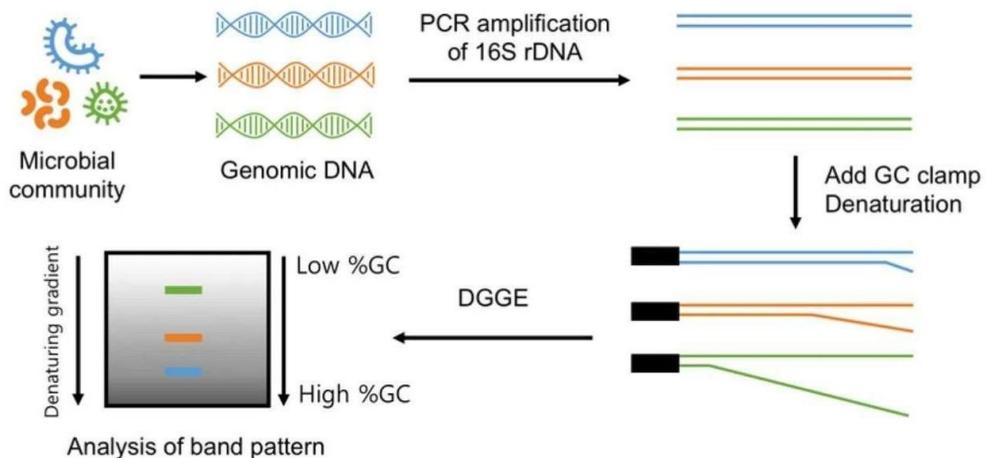
Au-delà des micro-organismes, de l'ADN végétal ou animal peut être détecté dans les études métagénomiques, quand ces organismes sont les supports ou les hôtes de micro-organismes de la communauté. Chez les végétaux, le recours à l'ADN chloroplastique pour différencier les taxons à l'échelle espèce est intéressant du fait qu'ils présentent un nombre de copies par cellule important. Différents marqueurs de cet organelle sont largement utilisés, tels que *ndhF*, *rbcL* ou *trnK*

(Zhang et al., 2013).

Si l'ADN est très dégradé comme dans les aliments très transformés ou les prélèvements environnementaux de permafrost, un marqueur très court de 130 pb, représenté par la boucle P6 de l'intron du gène chloroplastique d'ARNt trnL (UAA) peut être utilisé (Taberlet et al., 2007). Le marqueur de référence pour la caractérisation des espèces animales est quant à lui un marqueur de l'ADN mitochondrial représenté par un segment du gène de la cytochrome oxydase 1, COI (Hebert et al., 2003).

### 2.3.1.2. Méthode de caractérisation par PCR-DGGE (Denaturing Gradient Gel Electrophoresis)

Cette méthode d'empreinte génétique est basée sur la migration électrophorétique différentielle d'amplicons d'ADN simple brin de même taille mais de séquence variable, en conditions dénaturantes, illustrée par la Figure 10. Utilisée initialement pour la détection de mutations (Cariello et al., 1991 ; Tartary et al., 1993), elle s'est révélée intéressante en microbiologie.



**Figure 10.** Diagramme schématisé du principe de la méthode de caractérisation par PCR-DGGE (Sharma et al., 2020)

Si elle est appliquée à des fragments de PCR obtenus à partir des marqueurs cités plus haut, les profils obtenus permettent d'identifier l'espèce d'un micro-organisme isolé par rapport au profil d'une souche de référence ou d'observer les variations entre communautés microbiennes.

Dans ce cas, les profils montrent plusieurs bandes caractéristiques des taxons présents et l'intensité de celles-ci reflète l'abondance relative des taxons. L'ADN des bandes peut également être élué et séquencé pour confirmer l'identification taxonomique. Cette méthode a encore été mise en œuvre très récemment pour comparer des communautés de kombucha, en utilisant le marqueur V7V8 pour révéler les bactéries et le marqueur D1D2 pour les levures (Grassi et al., 2022). Elle reste cependant une méthode laborieuse et peu précise.

Parmi les diverses méthodes basées sur un profil moléculaire, on peut encore citer la méthode basée sur les polymorphismes de conformation simple brin (SSCP), la méthode basée sur les polymorphismes de l'espacement intergénique ribosomique (RISA), et sa version automatisée (ARISA) et les méthodes basées sur les polymorphismes de longueur des fragments de restriction (RFLP) (Johnston-Monje, D., et Lopez Mejia, J., 2020).

#### 2.3.1.3. Banques de clonages d'amplicons de PCR sur les marqueurs taxonomiques

Le séquençage du gène de l'ARNr 16S ou du segment D1-D2 de l'ARNr 25S, amplifiés par PCR à partir d'isolats (de bactérie ou de levure, respectivement) est généralement suffisant pour faire une identification au niveau de l'espèce, grâce à la comparaison de la séquence obtenue avec celles des banques de données de séquençage. Il a été possible de mettre en œuvre cette approche dans le cadre de communautés microbiennes avant de disposer des méthodes de séquençage à haut-débit. Pour ceci, il était nécessaire de passer par une étape de clonage

aléatoire, avec ligation dans un vecteur plasmidique afin de pouvoir séquencer individuellement quelques dizaines de clones des banques ainsi constituées. Cette méthode est certes laborieuse mais elle a permis de révéler l'existence et l'importance de micro-organismes non cultivés (Taylor et al., 2008).

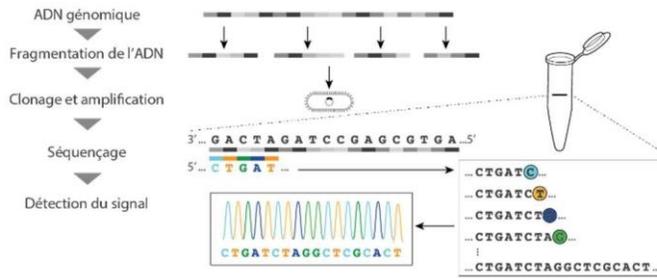
On peut considérer cette méthode comme l'ancêtre des méthodes actuelles, pour lesquelles les nouvelles techniques de séquençage permettent notamment de s'affranchir de l'étape de clonage.

Ces méthodes de caractérisation, bien qu'elles soient à faible débit, ont permis de compenser le travail laborieux de culture, pour travailler directement sur le matériel génétique. Toutefois, la distinction entre la fraction vivante et morte n'est plus possible. Il en résulte une estimation approximative, semi-quantitative de la communauté microbienne représentée en proportions relatives.

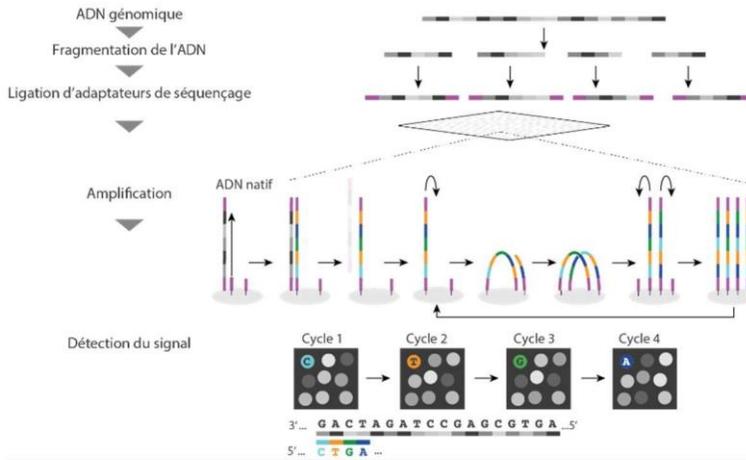
Dans les produits alimentaires, où différents constituants naturels sont présents dans la matrice, l'efficacité de l'extraction d'ADN et la pureté de l'ADN extrait sont critiques à cause de composés pouvant inhiber la réaction d'amplification par PCR.

**Figure 11.** Évolution et principe des stratégies de séquençage:(A.) Le séquençage de 1ère génération de type Sanger. (B.) La technologie de séquençage Illumina (C) Le séquençage de type Oxford Nanopore Figure adaptée de Shendure et al. (Shendure et al., 2017)

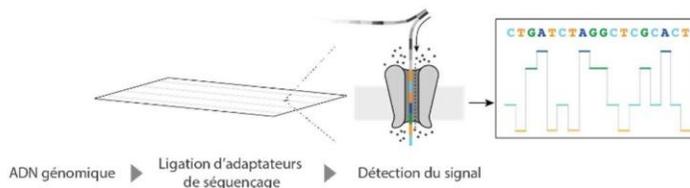
**A. Le séquençage de 1<sup>ère</sup> génération (Sanger)**



**B. Le séquençage de 2<sup>ème</sup> génération (NGS)**



**C. Le séquençage de 3<sup>ème</sup> génération (Oxford Nanopore)**



**2.3.1.4. Caractérisations de communautés microbiennes par des méthodes de séquençage à haut débit ciblé**

Par les méthodes à haut débit, on considérera celles dites de deuxième et de troisième génération. Le principe commun qui les distingue de celles dites de première génération, est la possibilité de séquencer en parallèle des millions de molécules d'ADN (voire d'ARN) individuelles. Les technologies utilisées et les signaux mesurés sont très différents comme illustrée dans les exemples de la Figure 11 mais ne seront pas détaillés ici. Ce qui distingue les méthodes de

deuxième génération de celles de troisième génération, est la longueur de lecture des séquences générées.

Ces méthodes sont donc des méthodes à haut débit qui permettent de générer des quantités de données massives avec une diminution importante des manipulations, des délais et surtout une diminution des coûts incomparable à ceux des méthodes de première génération, si on les rapportait à la même quantité de données (Rothberg and Leamon, 2008). Ces méthodes ont non seulement révolutionné les études basées sur le séquençage de génomes (et par là, la connaissance de la diversité des micro-organismes), mais également les études métagénomiques, comme on va le voir dans les paragraphes suivants.

Le principe de cibler par PCR et caractériser un marqueur spécifique à partir de l'ADN extrait de l'ensemble des cellules de prélèvement environnementaux, comme celui d'une région du gène de l'ARN 16S, est la base des différentes techniques basées sur les empreintes moléculaires (voir PCR-DGGE), et dans les techniques de clonage moléculaire décrites précédemment.

Cette même base est reprise pour les méthodes de caractérisation à haut débit ciblées. Lorsque la première méthode de séquençage à haut débit a été mise au point (la méthode de pyroséquençage 454, (Liu et al., 2012)), elle a été appliquée presque immédiatement pour séquencer directement et séparément les molécules de ces mélanges d'amplicons, sans étape supplémentaire après amplification par PCR (Huber et al., 2007 ; Sogin et al., 2006). Plusieurs milliers de molécules d'amplicons ont ainsi pu être séquencées par échantillon de manière simultanée conduisant à une augmentation d'un facteur 100 à 1000. La quantité de données générées a permis d'augmenter la précision des analyses du contenu en taxons et de détecter les taxons rares alors que les méthodes à faible et moyen débit basées sur le séquençage de quelques dizaines de clones ne détectent que les taxons majoritaires. L'ensemble des amplicons correspondant à ces taxons rares représente plus de 30% des amplicons générés. Ces études effectuées avec comme

marqueur le segment V6 du gène de l'ARN 16S, ont aussi montré que même avec des marqueurs de petite taille (120 pb pour le segment V6) il est possible de différencier les taxons de manière satisfaisante, au moins à l'échelle du genre, pour les bactéries.

À l'heure actuelle, cette méthode est principalement mise en œuvre avec la technologie de séquençage Illumina, avec des applications dans tous les domaines de la microbiologie : environnemental, médical et alimentaire. Cette méthode a été publiée en même temps que celle du pyroséquençage (Margulies et al., 2005) cependant, la très courte taille des lectures a été un frein à l'utilisation de cette technologie pour l'étude de l'ADN environnemental et des amplicons générés sur les marqueurs. Son évolution qui permet des lectures paires de 200 à 300 pb, son coût très inférieur au pyroséquençage et son plus faible taux d'erreur de séquençage (~ 0.1%) en font la méthode de choix à l'heure actuelle. Des études ciblées utilisent également les méthodes développées pour le séquençage à lectures longues comme la technologie PacBio (Callahan et al., 2019) et Oxford Nanopore (Matsuo et al., 2021), avec un développement de logiciels d'analyse dédiés à remédier aux erreurs de lecture NanoCLUST, par exemple, (Rodríguez-Pérez et al., 2021). Dans cette application, elles permettent de séquencer des amplicons recouvrant toute la séquence de l'ARN 16S, augmentant le pouvoir de résolution pour une affiliation plus précise des séquences à l'échelle espèce. Enfin, le protocole Loopseq de fragmentation de longs amplicons par introduction de code-barres permet la couverture de toute la séquence de l'ARN 16S tout en ayant recours à une plateforme de séquençage Illumina, les lectures étant assemblées au cours du traitement des données (Jeong et al., 2021).

### **2.3.1.5. Les limites des méthodes indépendantes de la culture et des méthodes ciblées**

Plusieurs étapes sont limitantes pour ces méthodes de caractérisation effectuées directement sur le matériel génétique et non sur des cellules mises en culture. La plupart des biais observés sont décrits par (Ercolini, 2004).

Ces méthodes ne permettent pas de distinguer les signaux des cellules vivantes et des cellules mortes, ce qui entraîne une part d'imprécision dans la mesure des proportions relatives des taxons.

L'efficacité de l'extraction d'ADN et la pureté de l'ADN extrait sont critiques pour plusieurs raisons :

- dans les produits alimentaires où différents constituants naturels sont présents dans la matrice, des composés peuvent interférer avec l'extraction d'ADN et également inhiber la réaction d'amplification par PCR, comme les polyphénols et les polysaccharides (Nagori and Purohit, 2012).
- les taxons présents peuvent avoir des architectures cellulaires qui les rendent diversement sensibles aux méthodes d'extraction, ce qui cause également une part d'imprécision dans la mesure des proportions relatives des taxons, si l'ADN est extrait avec des rendements variables selon les taxons.

Pour les méthodes ciblées, une étape de PCR est nécessaire et elle est également la source de nombreux biais comme l'amplification préférentielle du marqueur pour certains taxons et la formation d'amplicons chimériques.

Enfin, les marqueurs courts, représentés par un seul ou deux segments hypervariables du gène de l'ARN 16S bactérien, sont limités pour l'assignation taxonomique. La résolution se fait au niveau genre d'où la mise au point des méthodes longues lectures ou amplicons fragmentés citées plus haut.

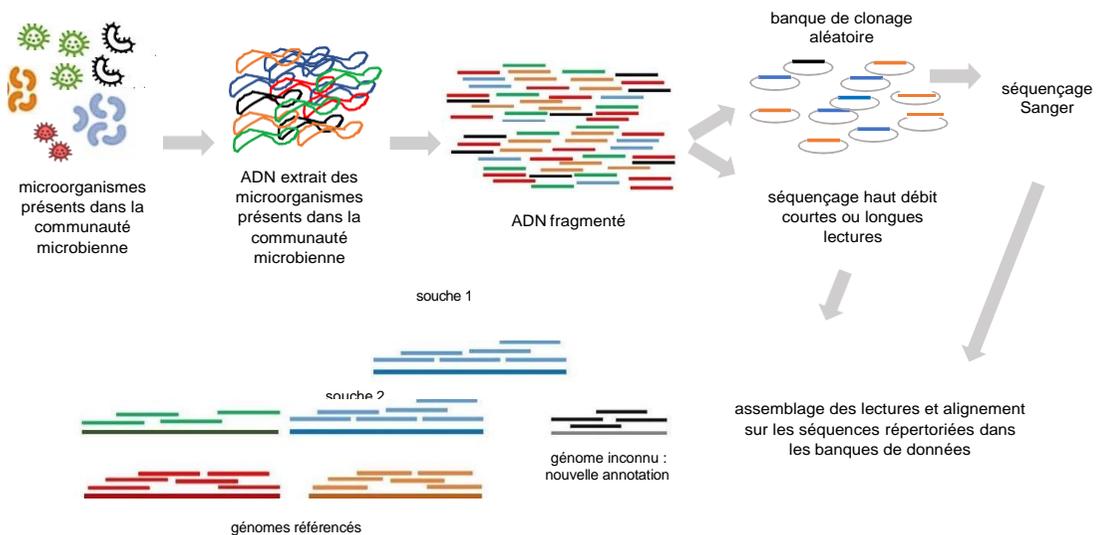
Les nombreuses copies en tandem des opérons d'ARN ribosomiques (ou rDNA array) peuvent présenter des différences de séquences qui perturbent l'assignation taxonomique, si ces différences entre les copies d'un même génome sont plus importantes que les différences entre copies d'espèces différentes. Ce problème est rencontré dans plusieurs espèces de levures (Sipiczki, 2022), notamment pour les espèces des genre *Zygosaccharomyces* et *Pichia*, présentes dans les SCOBY de kombucha. Chez les espèces concernées, les différents mécanismes permettant l'évolution concertée de ces copies (par exemple la conversion génique) et par conséquent leur homogénéité de séquence, semblent peu efficaces (Dakal et al., 2016).

Comme le nombre de copies en tandem des opérons d'ARN ribosomiques est variable d'une espèce à l'autre, voire d'une souche à l'autre (Espejo and Plaza, 2018 ; Peter et al., 2018) des corrections devraient être appliquées pour l'estimation des proportions de taxons. Cependant les informations sur le nombre de copies n'est pas toujours connu, par conséquent les proportions estimées restent à considérer comme des proportions d'amplicon correspondant à un taxon donné plutôt que l'abondance réelle des cellules de ce taxon.

Pour terminer cette liste, un inconvénient majeur des méthodes ciblées est la nécessité d'utiliser des marqueurs différents (donc des bibliothèques d'amplicons indépendantes) pour explorer des communautés contenant des taxons appartenant à plusieurs domaines du vivant, comme celles du kombucha. Par conséquent, il n'est pas possible d'estimer directement les proportions relatives des levures par rapport aux bactéries alors que ce paramètre est important pour comprendre leurs interactions d'où l'intérêt d'utiliser les méthodes non ciblées de type shotgun.

### 2.3.2. Les méthodes métagénomiques non ciblées : principe et évolution du séquençage shotgun appliqué aux communautés microbiennes

Les premières approches se basent comme précédemment sur le séquençage de type Sanger de banques de clonage aléatoire mais celles-ci ne sont plus basées sur des amplicons ciblant des marqueurs taxonomiques mais sur des fragments d'ADN total extrait des cellules d'un échantillon (Figure 12) (Schmeisser et al., 2003). En complément des approches ciblées, ces études ont permis de découvrir des nouveaux gènes (Krause et al., 2006), d'assembler des génomes d'espèces non cultivées et de reconstruire leur histoire évolutive (Venter et al., 2004). Ces études ont également initié l'exploration des voies métaboliques présentes sur un site à partir de son contenu en gènes (Gill et al., 2006 ; Kapardar et al., 2010; Pelletier et al., 2008).



**Figure 12.** Principe de la méthode métagénomique shotgun, appliquée en modes bas ou haut débit (adapté de <https://www.france-genomique.org/expertises-technologiques/metagenomique/>)

La première application d'une méthode de séquençage haut débit à une approche métagénomique shotgun a été réalisée en 2006, avec la méthode de pyroséquençage, sur des prélèvements miniers (Edwards et al., 2006). Cette étude

a comparé les résultats des méthodes de séquençage de type Sanger sur banques classiques avec ceux du séquençage haut débit, aboutissant à des compositions microbiennes similaires pour un coût 10 à 30 fois moins élevé. Elle a aussi montré que son application dans un contexte de métagénomique comparative permet de mettre en évidence que les différences de compositions microbiennes observées sont associées à des enrichissements en groupes de gènes caractéristiques, de voies métaboliques particulières des environnements explorés.

Un apport inestimable de ces méthodes non ciblées est l'exploration facilitée des taxons viraux qui coexistent avec les taxons microbiens (Angly et al., 2006; Edwards and Rohwer, 2005) et qui ont un rôle important dans la régulation des populations microbiennes y compris celles des fermentations alimentaires (Ledormand et al., 2021).

Dans la perspective de comprendre les interactions entre micro-organismes, ces approches métagénomiques donnent une vision de communautés de gènes, plutôt que de communautés de taxons et permet d'appréhender dans toute sa complexité la diversité génétique contenue dans les populations microbiennes (Wilmes et al., 2009). Les progrès technologiques devront permettre de contourner certaines limitations comme la nécessité d'extraire des quantités d'ADN plus importantes que pour les approches ciblées, la détection des taxons rares et le traitement des éventuelles contaminations par de l'ADN végétal et/ou animal.

De manière plus générale, les méthodes de séquençage haut débit, appliquées aux approches métagénomiques ciblées ou non ciblées, permettent de multiplier le nombre d'échantillons et d'atteindre une précision sans précédent. Elles ont ouvert de larges perspectives dans la discipline de l'écologie microbienne, pour l'exploration comparative de la dynamique des communautés microbiennes en réponse à des variations environnementales, tout en atteignant des niveaux taxonomiques très fins à l'échelle de la souche.



# **Histoire, contexte et vue d'ensemble du projet de thèse**

Mon intérêt pour la fermentation kombucha a commencé en 2014 lors de mon projet de fin d'études pour l'obtention du diplôme national d'ingénieur de l'INSAT (Institut National des Sciences Appliquées et de Technologie de Tunis). Durant ce projet, je me suis intéressée à la maîtrise de la fermentation du kombucha à partir du thé noir dans des fermenteurs de 1 et 10 litres. J'ai pu reconstituer un SCOBY à partir de cultures pures d'une souche de *Brettanomyces bruxellensis* et quatre souches de *Komagataeibacter xylinus*, puis maîtriser sa fermentation dans des volumes de 1 litre à 10 litres dans des conditions industrielles. À la fin du projet, j'ai conservé le SCOBY comme c'est la tradition chez les amateurs de kombucha : garder et entretenir son propre SCOBY et le partager.

Brasser une boisson non alcoolisée de caractère, atypique, a été l'objectif des fondateurs de BioBrasseurs. Le kombucha a répondu à leurs attentes : une boisson à base de thé, peu sucrée qui peut être brassée à l'artisanal mais dans des grands volumes. En 2016, je les ai rejoints ... avec mon SCOBY à l'origine du starter de BioBrasseurs. Ce SCOBY a subi de nombreuses étapes : une conservation de deux années au frigo, des back-sloppings à Tunis, un passage à la fermentation du thé vert, une augmentation en échelle à Strasbourg qui a abouti produire en cuves de 1000 litres puis de nombreux back-sloppings en production et la division du SCOBY pour alimenter plusieurs lignes de production parallèles. Le mode de propagation du biofilm a changé : les biofilms mère et fille ne se détachent plus. Durant ces différentes transitions, les manipulations ont été réalisées dans des conditions propres mais non stériles. Aucun incident de contamination n'a été détecté jusqu'à présent.

C'est à ce stade que BioBrasseurs commence à s'interroger quant à l'évolution de la composition microbienne : les différentes conditions subies par le SCOBY ont certainement changé sa composition de départ, ce qui soulève la problématique de sa stabilité dans le temps. Afin de développer le volet recherche et

développement de l'entreprise, avec le souhait de collaborer avec un laboratoire de recherche intéressé par les communautés microbiennes, une convention CIFRE a été établie avec l'équipe de Joseph Schacherer, de l'UMR7156 - GMGM.

L'objectif de mon projet de thèse a consisté à tirer profit des avancées technologiques en métagénomique et des ressources disponibles afin de caractériser la communauté microbienne impliquée dans la fermentation du kombucha. Pour ce faire, une approche indépendante de la culture, pour déterminer la composition microbiologique par métagénomique ciblée (ou métabarcoding), a été adoptée. Quant aux ressources biologiques disponibles, BioBrasseurs a mis à disposition ses SCOBY, substrats et une collection de kombucha commercialisées non pasteurisées.

Au laboratoire, le premier challenge a été de reproduire la fermentation de kombucha à une échelle 'domestique', pour des volumes néanmoins supérieurs à un litre comme cela est recommandé (Crum and LaGory, 2016). Dans le **premier chapitre** de ma thèse, nous décrivons les différentes mises au point de l'ensemble des méthodes à mettre en œuvre. D'une part nous avons adapté des conditions laboratoires, qui sont plus contrôlées en termes de paramètres de fermentation par rapport aux conditions domestiques ou industrielles ; d'autre part, nous avons mis en place les protocoles d'obtention des librairies d'amplicons pour le séquençage à haut débit et d'analyse des données. Nous avons notamment développé un nouveau marqueur de métabarcoding pour caractériser les espèces de levures, qui repose sur le segment D2 du gène de l'ARNr 26S. Une fois en place, ces méthodes ont été utilisées dans tous les axes de mon projet.

Deux axes principaux ont été développés autour de la variabilité dans la composition microbiologique des communautés impliquées. Le premier axe, traité dans le **deuxième chapitre**, repose sur l'utilisation d'une collection de starters et a consisté à étudier la diversité taxonomique des SCOBY qui

s'assemblent spontanément à partir de ces starters. Au total, une collection de 18 SCOBY a été rassemblée et étudiée. Contrairement aux travaux publiés sur les SCOBY de kombucha, qui prennent peu en compte l'historique fermentaire des échantillons étudiés, nos SCOBY se sont assemblés dans les mêmes conditions de fermentation. Nos résultats sont donc dépendants de la composition initiale du starter et de l'adaptation des taxons présents aux conditions que nous avons fixées. Nous avons établi les profils taxonomiques fongiques et bactériens pour 75 échantillons issus des différentes fractions de liquide de fermentation et de biofilm de ces SCOBY qui ont été propagés par back-slopping sur trois cycles consécutifs de fermentation. Cette exploration a permis la définition du microbiote core, qui correspond aux taxons présents dans tous les SCOBY et du microbiote accessoire. Ces profils ont aussi permis de mettre en évidence des fluctuations des compositions au cours des trois cycles, affectant plus particulièrement les abondances du microbiote accessoire.

L'évolution des SCOBY de kombucha dans le temps n'a pas encore été étudiée. Chez les brasseurs et les producteurs amateurs, étant donné les conditions d'hygiène non stériles, la probabilité de colonisation des SCOBY par des taxons de l'environnement est certainement courante. De plus, la stabilité à long terme de la composition du microbiote core n'est pas connue. Mon **troisième chapitre** explore ainsi cette évolution potentielle pouvant inquiéter le partenaire industriel qui se demande même un "âge opérationnel de SCOBY" peut avoir lieu. Nous avons traité cette thématique avec une exploration de trois types de séries temporelles. La première série temporelle a été conduite à l'échelle de production dans l'environnement de la brasserie : les compositions microbiennes de brassins successifs de lignées de production parallèles ont été analysées durant une année. La deuxième série temporelle a été conduite au laboratoire avec un design expérimental permettant de propager en parallèle des réplicats biologiques d'un SCOBY, au cours de nombreux back-sloppings. Avec des manipulations en

conditions stériles, l'effet de 'l'âge' des SCOBY était le paramètre à tester sur la composition taxonomique. La troisième série temporelle a également été conduite au laboratoire, avec des starters issus d'un même SCOBY mais soumis au préalable à des conditions de stockage différentes. Les analyses de l'ensemble de ces séries a montré une variation importante entre les SCOBY et entre les fermentations successives. Elles ont également permis d'identifier les impacts d'événements environnementaux sur les compositions microbiennes. Elles constituent une base pour développer des études permettant de définir des profils de variation afin de mieux comprendre les interactions entre les taxons des SCOBY.

# **Chapitre 1 : Mise au point des méthodes du projet**

## **Introduction**

Explorer la stabilité et l'évolution des communautés microbiennes dans l'environnement du kombucha a nécessité la mise en place de protocoles et leur optimisation afin d'assurer, d'une part, le bon déroulement de fermentations de type kombucha dans les conditions de laboratoire, et d'autre part, l'obtention des résultats représentatifs de séquençage haut débit qui permettent de détecter les effets de conditions différentes sur la composition de ces communautés. C'étaient des aspects techniques alors inédits au laboratoire, le premier objectif de ma thèse a donc été la mise en place de protocoles de référence pour les travaux de ma thèse et pour des études ultérieures des communautés microbiennes. C'est pourquoi un chapitre est consacré aux tests effectués dans le cadre de cette mise en place.

### **1. Cultures et fermentations du kombucha au laboratoire**

#### 1.1 Établir un protocole de fermentation standard

La première étape commune à tous les axes de ce projet de thèse a consisté à établir notre protocole standard pour conduire des fermentations de type kombucha au laboratoire. Notre choix a été d'utiliser la méthode traditionnelle de back-slopping qui utilise comme starter naturel les deux produits d'une fermentation précédente à savoir la boisson et le biofilm-fille (Figure 1.1). Ce protocole devait tenir compte des paramètres décrits dans l'introduction mais aussi des contraintes de volume d'infusion et de manipulation des essais. Nous avons opté pour une fermentation à l'échelle de celles des producteurs amateurs qui fermentent du kombucha dans leur foyer. Les données décrites dans la littérature (Laavanya et al., 2021) récapitulant les différents paramètres affectant

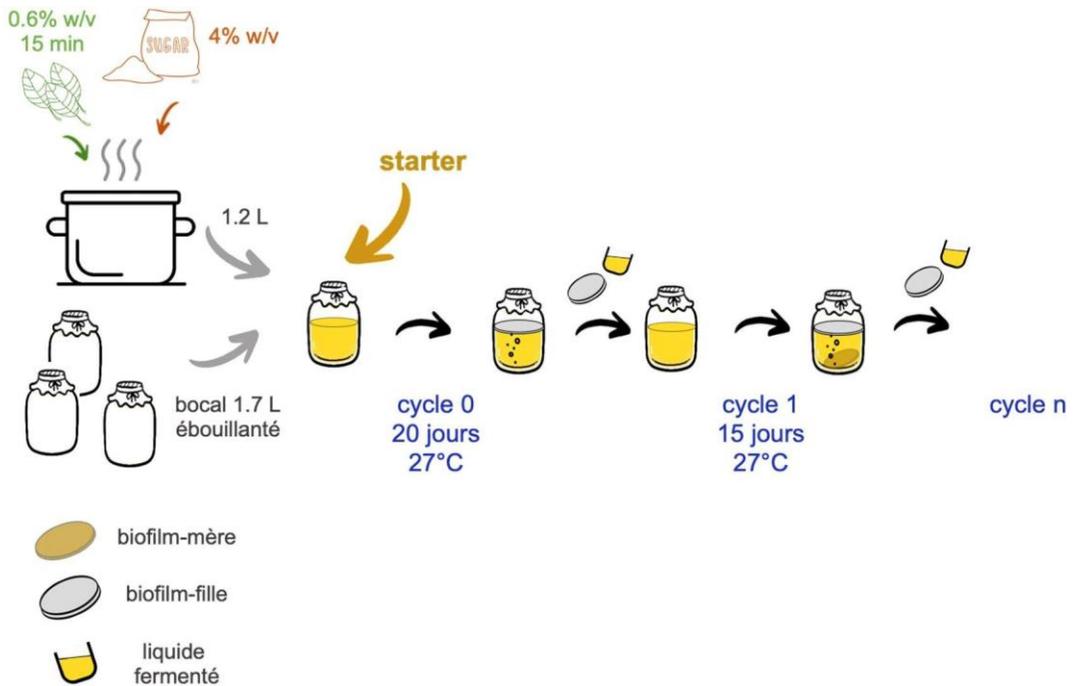
la fermentation de type kombucha et les conditions utilisées chez BioBrasseurs nous ont servis comme référence afin de définir notre protocole de back-slopping.

Les bocaux que nous utilisons ont un diamètre de 11,1 cm de diamètre et un volume total de 1,7 litres. En se basant sur l'étude de (Cvetković et al., 2008), le volume de liquide optimal est de 1,3 litres. Le volume des bocaux utilisé a satisfait la contrainte de gérer l'espace afin de pouvoir placer simultanément jusqu'à 30 fermentations en parallèle dans des incubateurs standards. La quantité de thé infusé et la concentration en sucre de canne utilisée correspondent aux conditions utilisées en routine dans la production chez BioBrasseurs. La quantité de thé vert équivalent à 0.5% (w/v) est infusée pendant 15 minutes. La concentration de sucre de canne est de 4% (w/v). L'eau utilisée est un mélange composé de 90% d'eau filtrée ultrapure (station Millipore) et 10% d'eau potable du robinet. Ce mélange permet de limiter la dureté de l'eau de robinet tout en assurant une présence minimale de minéraux. En effet, il a été montré que des teneurs en calcium élevées interfèrent avec l'extraction des acides organiques et polyphénols du thé lors de l'infusion (Mossion et al., 2008). Afin de limiter au maximum les contaminations, nous avons stérilisé les bocaux recouverts de gaze par passage au four Pasteur (2h à 200°C).

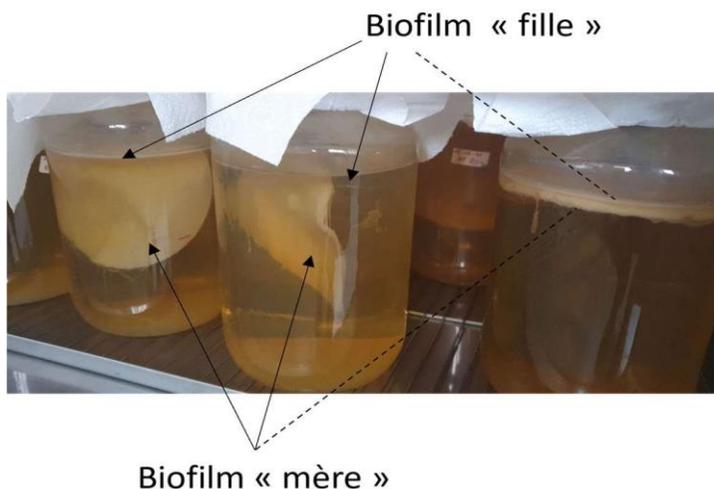
Les starters utilisés pour les premiers essais étaient représentés soit par des SCOBY commerciaux, soit par du liquide de fin de fermentation et du biofilm provenant des cuves de BioBrasseurs. Pour considérer la réussite de la fermentation de type kombucha, il est nécessaire d'observer la formation d'un disque de biofilm-fille en surface, régulier, translucide et atteignant une épaisseur d'environ 5 mm au bout de 2 semaines (Dutta and Gachhui, 2007). Ces paramètres pris en compte, nos premiers essais de fermentations n'ont pas été satisfaisants. L'absence de formation d'un biofilm-fille a pointé la problématique d'une inhibition de l'activité des micro-organismes et le traitement des récipients

utilisés pouvaient être à l'origine de cette inhibition. En effet, les amateurs de kombucha qui en produisent chez eux n'ont généralement pas de stérilisateur mais ils insistent sur l'importance du nettoyage des récipients et ustensiles utilisés (<https://www.kefir-kombucha.com>). Nous avons donc modifié le traitement et la stérilisation des bocaux de manière à imiter des conditions domestiques. Ainsi, les bocaux sont rincés deux fois à l'eau déminéralisée à la sortie de la machine à laver pour éviter la moindre trace de détergent. Nous avons évité le passage au four Pasteur qui peut s'accompagner d'émissions de composés lors du chauffage des papiers indicateurs de stérilisation et du brunissement de la gaze. Juste avant d'être remplis d'infusion et d'être inoculés, les bocaux sont traités par deux rinçages à l'eau bouillante (Jayabalan et al. 2014), égouttés à l'envers puis recouverts d'un simple papier absorbant propre. Les essais effectués en suivant ces traitements ont été satisfaisants et reproductibles sur plusieurs cycles successifs de fermentation (Figure 1.1). Pour évaluer le risque de contamination, des suivis réguliers d'incubation d'infusion non inoculée ont été effectués. Un seul cas de contamination de ces contrôles a été observé, avec le développement visible de moisissures, pour la dizaine d'essais effectués au cours des trois années, et les fermentations inoculées en parallèle de ce cas n'avaient pas montré une telle contamination.

Au cours des cycles de fermentation, les bocaux sont incubés dans des conditions statiques, à une température constante de 27°C, pendant une durée de 14 jours, durant lesquels aucune manipulation n'est effectuée afin de ne perturber ni le métabolisme de la communauté microbienne, notamment la formation du biofilm de cellulose, ni la distribution spatiale des nutriments, de l'oxygène et des produits de fermentation. Après 14 jours, l'épaisseur du biofilm est d'environ 5 mm et l'odeur est fruitée et légèrement acidulée.



**Figure 1.1.** Initiation d'une lignée de fermentation. L'infusion sucrée et refroidie est ensemencée avec un starter qui peut prendre différentes formes (SCOBY commercial, rétentat Biobrassieurs...), au bout de 20 jours, si un biofilm cellulosique s'est formé, les back-slopping traditionnels successifs peuvent être effectués.



**Figure 1.2.** Observation des bocaux de fermentation à l'échelle laboratoire au septième jour de fermentation. À ce stade, le biofilm mère commence à s'alourdir et tomber au fond du bocal et seul le biofilm mère reste en surface et continue de s'épaissir. Dans le bocal de droite, le biofilm mère flotte encore sous le biofilm fille.

Une légère effervescence peut être observée et un dosage de composés comme le saccharose consommé, l'acide acétique ou l'éthanol produits peuvent confirmer l'aboutissement de la fermentation. Après cette durée, la fermentation peut être arrêtée et le back-slopping suivant peut être enclenché (Figure 1.1). À partir d'une fermentation terminée, un back-slopping de routine consiste à inoculer 1,2 litres d'infusion sucrée fraîche, avec 50 ml du liquide d'une fermentation précédente et le biofilm-fille.

## 1.2 Établir des collections de prélèvements représentatifs

Du fait de la fermentation en conditions statiques et de la nature solide d'une partie des produits de fermentation, la problématique de réaliser des prélèvements comparables entre les essais s'est posée.

Pour les essais en laboratoire, comme les prélèvements sont effectués en fin de fermentation, il suffit de mélanger soigneusement la fraction liquide avec une pipette stérile. Pour les biofilms mère et fille (Figure 1.2), une bande de 2 cm découpée sur le bord des disques sur toute leur épaisseur permet d'avoir des échantillons représentatifs. En effet, une étude a montré que la composition microbienne diffère sur l'épaisseur du biofilm-fille, de la face supérieure à la face inférieure, mais elle ne diffère pas en différents points d'une même surface (Coton et al., 2017).

Notre projet nécessitait également d'avoir des échantillons comparables issus des cuves de fermentation de BioBrasseurs. Il n'était cependant pas possible d'homogénéiser le volume des cuves de 1000 à 4000 litres. Quant aux biofilms, leur aspect est très différent de ceux des essais en laboratoire. Ils pèsent entre 10 à 30 kg et les biofilms mère et fille ne sont pas séparés. Ils sont donc très hétérogènes, cela est visible avec des zones claires correspondant à de la cellulose fraîchement synthétisée et des zones brunes correspondant à de la cellulose plus ancienne.

Cette hétérogénéité des phases liquides et solides des cuves de production s'est également manifestée lors des essais de fermentation au laboratoire initiés avec les starters issus de ces cuves. Les aspects de ces fermentations étaient très variables dans la vitesse de formation du biofilm-fille. Une solution à cette problématique a été trouvée à l'occasion d'une évolution dans le procédé de fabrication chez BioBrasseurs. Pour répondre à certains critères de qualité de la boisson commercialisée, une étape de filtration grossière du liquide des cuves a été instaurée avant de procéder aux étapes d'aromatisation et d'embouteillage. Le rétentat de cette filtration est une suspension dense de l'ensemble des micro-organismes de la cuve et de fines particules de biofilm. Il est considéré comme un sous-produit de fabrication qui s'est révélé très intéressant dans notre projet afin de constituer des prélèvements représentatifs. Nous avons développé plusieurs applications pour utiliser ces rétentats de filtration :

- les cellules des rétentats sont récupérées pour déterminer la composition microbienne par métabarcoding et suivre ainsi l'évolution de la production chez BioBrasseurs,

- le rétentat est utilisé comme starter pour établir ce que nous avons dénommé une 'lignée de fermentation' (Figure 1.1). En effet, lorsque du rétentat est inoculé à raison de 5% (v/v) à une infusion de thé sucrée, un biofilm se constitue au bout de 3 semaines. C'est une durée habituelle pour le développement d'un biofilm à partir d'une suspension microbienne. Ce biofilm et le liquide de cette fermentation initiale permettent ensuite d'initier les fermentations successives par back-slopping traditionnel.

Un intérêt particulier de cette application est de pouvoir effectuer des répliquats biologiques. En effet, la suspension microbienne d'un rétentat est facile à homogénéiser et à répartir pour constituer des lignées de fermentations parallèles *a priori* parfaitement identiques.

- le rétentat peut être traité dans plusieurs conditions avant de servir de starter pour initier des lignées de fermentation. Les effets de ces conditions sur la

composition microbienne des fermentations ultérieures pourront donc être comparés de manière rigoureuse.

Les résultats de ces applications sont décrits dans le troisième chapitre de la thèse.

## **2. Mise au point des différentes étapes du protocole de métagénomique ciblé par séquençage haut débit (ou métabarcoding).**

Une fois les échantillons prélevés, l'analyse des communautés microbiennes réalisée par métabarcoding implique plusieurs étapes, dans l'ordre chronologique :

- la récupération des cellules du prélèvement et leur conservation
- l'extraction d'ADN génomique des cellules,
- la génération de bibliothèques d'amplicons de PCR avec des amorces qui ciblent le ou les marqueurs taxonomiques choisis selon les taxons dont on souhaite détecter la présence et estimer les proportions,
- le 'multiplexage maison' : si on cible plusieurs marqueurs taxonomiques pour un même échantillon, il est possible de mélanger des bibliothèques des différents marqueurs et de traiter ce mélange dans le même run de séquençage,
- le séquençage des bibliothèques par une plateforme de séquençage,
- le traitement des lectures de séquençage

Les protocoles de certaines de ces étapes ont nécessité des mises au point qui sont décrites ci-dessous. Avant de les décrire, je présenterai d'abord l'étape de traitement des lectures de séquençage car la validation de certains aspects des protocoles a nécessité d'analyser les séquences obtenues dans les différentes conditions.

## 2.1 Le traitement des données de séquençage

C'est l'étape finale du processus expérimental, juste avant de pouvoir comparer les compositions microbiennes de différents échantillons. Elle permet d'obtenir les abondances des lectures de séquençage affiliées à un taxon. Ces données constituent des "proxy" pour estimer l'abondance **relative** des cellules de ce taxon dans le prélèvement.

### 2.1.1 L'outils de traitement FROGS

Les bibliothèques d'amplicons sont séquencées en lectures paires, par la technologie Illumina Miseq, aboutissant à des lectures de 250 pb. La longueur de nos amplicons varie de 330 à 460 pb, selon le marqueur ciblé (voir § 2.2.3.2), ce qui garantit un chevauchement de 40 à 170 pb, des deux lectures d'un même amplicon (appelées les lectures R1 et R2) et permet leur assemblage. Ces lectures sont archivées dans deux fichiers paires sous format FASTQ.

Le pipeline appelé FROGS, "Find Rapidly OTU by Galaxy Solutions, supporté par la plateforme en ligne Galaxy (Escudie et al., 2018) est conçu pour analyser ce type de données dans le contexte de communautés microbiennes. Nous avons utilisé cette solution pour traiter les séquences reçues des plateformes de séquençage.

Une fois les fichiers de lectures chargés, la première étape consiste à renseigner la séquence des amorces qui ont ciblé le marqueur que l'on veut analyser. Cette étape permet de ne traiter que les lectures de ce marqueur, dans le cas où l'on avait mélangé des bibliothèques. Les lectures R1 et R2 sont ensuite assemblées, avec une tolérance de mésappariement de 10% dans la région chevauchante et les séquences produites peuvent être filtrées selon leur longueur. Les séquences retenues sont ensuite regroupées en 'clusters' avec l'outil SWARM (Mahé et al., 2015) selon une dissimilarité séquentielle fixe en s'appuyant sur un seuil de liaison locale.

La divergence maximale entre les séquences d'un cluster est de 3%. Certains clusters sont en fait des artéfacts formés lors de l'étape d'amplification PCR. En effet, lors d'une extension incomplète pendant un cycle de PCR, un brin partiellement étendu peut s'hybrider, au cours des cycles suivants, à une matrice d'une séquence différente mais similaire. Ce phénomène aboutit à des chimères qui correspondent à des clusters regroupant généralement de petits nombres de séquences. Sur FROGS, c'est l'outil VSEARCH (Rognes et al., 2016) qui est utilisé pour détecter et éliminer ces chimères avec une stratégie qui consiste à aligner les séquences des clusters peu abondants contre les séquences des clusters abondants.

Les clusters retenus sont considérés comme des Unités Taxonomiques Opérationnelles (OTU). Chaque OTU est ensuite affilié à un taxon. La séquence consensus de chaque OTU est comparée, par BLAST contre une banque de données contenant un ensemble de séquences correspondant au marqueur étudié où à chaque séquence est attribuée une assignation taxonomique contenant sept rangs taxonomiques. Par exemple, dans la banque Silva 138 (<https://www.arb-silva.de/documentation/release-138/>), qui est la banque la plus exhaustive pour le gène de l'ADNr 16S (Quast et al., 2013), les sept rangs taxonomiques correspondent au domaine, au phylum, à la classe, à l'ordre, à la famille, au genre et à l'espèce. Cette banque est l'une des banques utilisées par FROGS. Pour une assignation taxonomique plus précise, des paramètres relatifs aux taux d'identité minimal, au taux de couverture, ou à une longueur d'alignement peuvent être pris en compte. Si la séquence de l'OTU est identique à celle de deux taxons ou plus dans la banque, l'affiliation est attribuée au rang taxonomique supérieur et la mention 'multi-affiliation' est ajoutée pour le(s) rang(s) taxonomique(s) inférieur(s). Par exemple, le marqueur V3-V4 permet une affiliation jusqu'au niveau genre, au mieux, mais pas au niveau espèce car la diversité du segment V3-V4 n'est pas suffisante pour affilier l'espèce de manière incontestable.

Les séquences consensus des OTU peuvent être récupérées et traitées manuellement si on souhaite affiner l'assignation taxonomique. En effet, la qualité de l'assignation est très dépendante de la qualité de la banque, en particulier si elle est à jour pour les séquences répertoriées et pour les dénominations les plus récentes des rangs taxonomiques.

### 2.1.2 Outils de visualisation des données et indices de comparaison

Une fois la composition en OTU validée pour un ensemble d'échantillons, le package Phyloseq (McMurdie and Holmes, 2013) est très utile pour les représenter et les comparer. Il est par exemple possible, pour un ensemble d'échantillons présentant des nombres de lectures variables, de ramener les compositions au même nombre de lectures totales (généralement celui de l'échantillon ayant le plus faible nombre de lectures) avec une fonction de raréfaction. Il est aussi possible "d'agglomérer" les OTU qui correspondent à un même rang taxonomique. En effet, au moment du clustering des séquences, des séquences qui correspondent au même taxon se retrouvent dans des OTU différents. Cela permet de comparer les échantillons pour travailler, par exemple, à l'échelle espèce ou à l'échelle genre, ce qui a plus de sens au niveau biologique. Enfin, il est possible de calculer des indices de diversité qui sont utilisés pour décrire la biodiversité d'un échantillon et d'estimer la distance entre les compositions de plusieurs échantillons. Dans nos études, les indices et valeurs utilisés sont les suivants :

- la richesse (notée R) correspond simplement au nombre de taxons observés,
- l'indice de Gini-Simpson (Guisu and Guisu, 2010) est l'un des indices qui permet d'évaluer la diversité d'une communauté. Il est compris entre 0 et 1. Plus il est élevé, plus la diversité est importante. Il dépend du nombre R de taxons présents et de leurs proportions relatives  $p_i$ . Par exemple, la diversité d'une communauté A est plus élevée s'il y a peu de taxons mais qu'ils sont dans des

proportions équivalentes que celle d'une communauté B avec plus de taxons dont un taxon majoritaire. En d'autres termes, la valeur de cet indice indique que si on tire deux individus au hasard, il y a plus de chances qu'ils appartiennent à des taxons différents dans la communauté A que dans la communauté B.

La formule de calcul est la suivante :  $GS = 1 - \sum_{i=1}^R p_i^2$

- la dissimilarité de Bray-Curtis (Somerfield, 2008) est calculée entre les échantillons deux à deux. Elle correspond à la somme des rapports des valeurs absolues des différences de quantités de chaque taxon sur la quantité totale de ce taxon dans les deux échantillons. La valeur est comprise entre 0 (deux échantillons identiques) et 1 (deux échantillons complètement dissimilaires).

La formule de calcul est la suivante :  $BC_{ij} = \sum \frac{|n_{ik} - n_{jk}|}{(n_{ik} + n_{jk})}$

## 2.2 Extraction de l'ADN total des cellules d'un prélèvement

Les cellules d'un prélèvement de la phase liquide sont récupérées par centrifugation à 10,000 g pendant 10 minutes. Le culot cellulaire peut être conservé à -20°C jusqu'au moment de l'extraction d'ADN. Pour une fermentation en bocaux, les cellules d'un volume de 50 ml de la phase liquide sont récupérées. Pour les rétentats de filtration, les cellules d'un volume de 5 ml sont récupérées. Le protocole d'extraction est crucial pour obtenir un ADN débarrassé des potentiels inhibiteurs de l'amplification PCR (comme les polyphénols, naturellement présents dans le kombucha). Notre protocole est basé sur l'utilisation du kit DNeasy PowerSoil Pro Kit de Qiagen, qui est très utilisé pour les échantillons alimentaires (Evans et al., 2018). Nous avons appliqué des ajustements au protocole du fournisseur pour optimiser nos préparations d'ADN

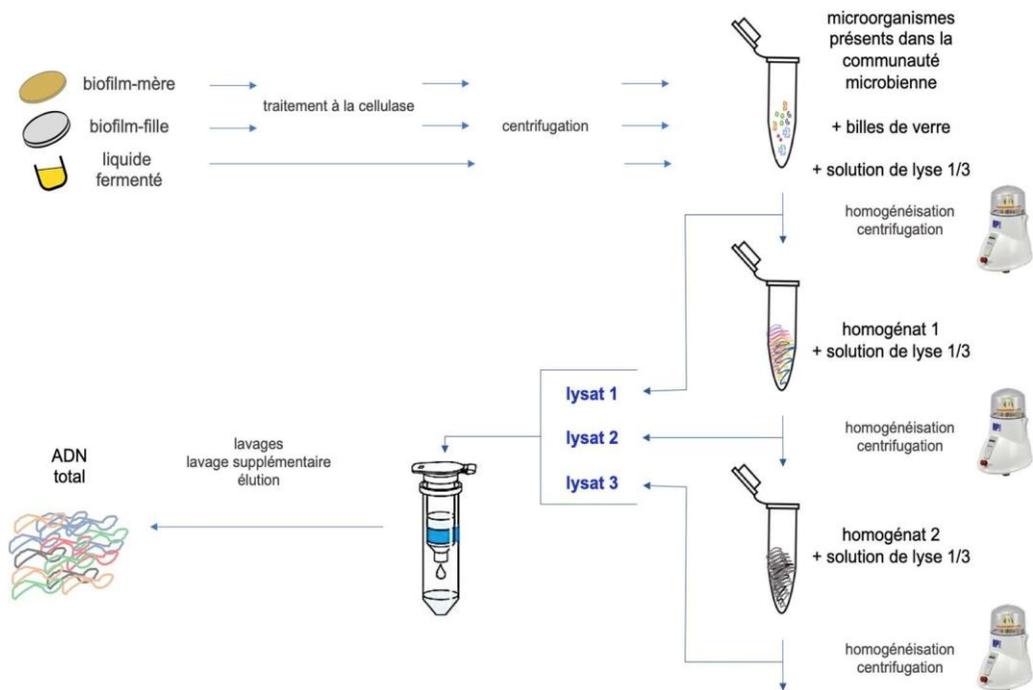
qui ont été définis à partir de nombreux exemples de protocoles de la littérature et d'échanges avec Monika Coton qui a travaillé sur les SCOBY de kombucha.

### 2.2.1 Lyse des cellules

Cette étape du protocole est critique car il est nécessaire que l'ADN extrait soit représentatif de toutes les espèces présentes et de leur proportion relative, la lyse des cellules doit donc être complète. Dans le cas du kombucha, comme cela a été décrit dans l'introduction, les espèces présentes peuvent avoir des niveaux très variables de résistance à la lyse mécanique selon qu'il s'agisse de levures, de bactéries à Gram positif ou de bactéries à Gram négatif. Cependant, si on applique une lyse mécanique suffisamment intense pour lyser les cellules les plus résistantes, l'ADN extrait des cellules les plus fragiles risque d'être trop fragmenté et ces espèces seraient sous représentées lors des détections ultérieures. Des protocoles de lyse enzymatique sont également décrits pour les levures et les bactéries mais ils ne sont pas forcément efficaces sur toutes les espèces. Enfin, les seuls effets des détergents sont insuffisants pour la lyse de cellules à paroi.

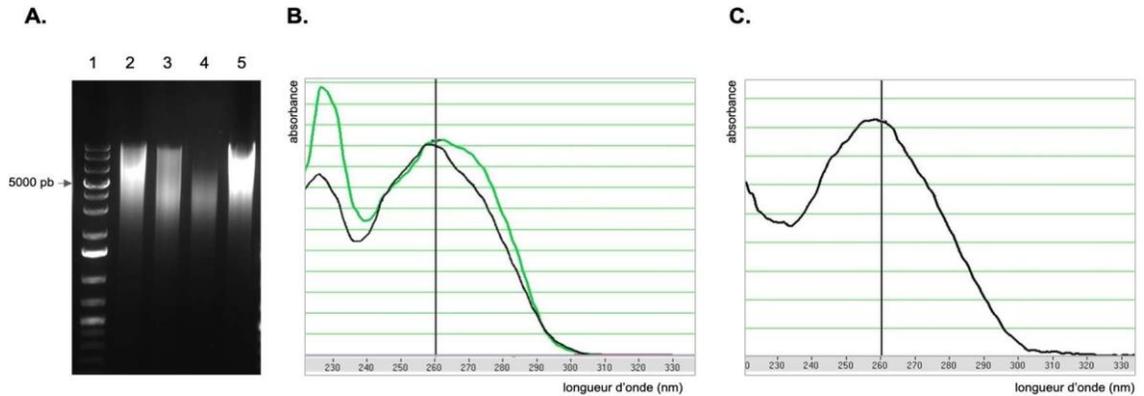
Nous avons élaboré un protocole inédit qui consiste en plusieurs étapes séquentielles de traitement mécanique permettant de pousser la lyse des cellules tout en préservant la qualité de l'ADN extrait (Figure 1.3). Les traitements mécaniques sont effectués avec un homogénéisateur de tissus FastPrep (MP Biomedicals) et des billes de verre de diamètre 425-600  $\mu\text{m}$  (Sigma Aldrich). L'observation de l'état des cellules après chaque traitement mécanique, visualisées au microscope dans des états frais et la qualité de l'ADN extrait (Figure 1.4 A) ont permis de définir un nombre de trois cycles de traitement mécanique. Le protocole de lyse a donc été adapté, en fractionnant en trois parties la quantité de solution de lyse prévue par le protocole du fournisseur. Les cellules totales ont été soumises à un premier traitement mécanique avec une part de la

solution de lyse puis le lysat a été séparé par centrifugation des débris et des cellules non lysées et finalement récupéré et réservé. De la même manière, les cellules non lysées ont été soumises à un deuxième traitement mécanique avec une nouvelle part de la solution de lyse et le lysat de cette étape a été récupéré avant de soumettre les cellules restantes à un troisième traitement mécanique avec la dernière part de la solution de lyse.



**Figure 1.3.** Grandes étapes de l'extraction d'ADN des cellules des différentes fractions d'un SCOBY de kombucha.

Les trois lysats ont ensuite été regroupés avant de poursuivre la purification de l'ADN extrait selon le protocole du fournisseur, avec une nouvelle modification comme suit : avant l'étape finale d'élution de l'ADN fixé sur une colonne de silice, nous avons effectué un lavage supplémentaire avec la dernière solution de lavage du kit.



**Figure 1.4.** Analyses de la qualité de l'ADN extrait et purifié. (A) Comparaison des profils de l'ADN des trois lysats séquentiels successifs (pistes 2, 3, 4) et rassemblés (piste 5), après électrophorèse en gel d'agarose 0.8%, avec le marqueur de taille GeneRuler 1 kb Plus de Thermo Scientific (piste 1). (B) Spectre d'absorption d'une préparation d'ADN purifié selon le protocole du fournisseur (ligne verte), effectué au spectrophotomètre Nanodrop 2000 (C) Spectre d'absorption d'une préparation d'ADN purifié avec un lavage supplémentaire avant l'étape d'élution, au spectrophotomètre Nanodrop 2000.

Ce lavage supplémentaire a permis d'éliminer des composés présentant un pic d'absorption à 230 nm que nous supposons être de la guanidine présente dans des traitements de la colonne et qu'un seul lavage ne suffit pas à éliminer (Figure 1.4 B et C). Ce protocole modifié est devenu le protocole standard pour tous les échantillons analysés pendant la thèse, il permet d'obtenir en routine 300 ng d'ADN par échantillon avec une majorité de fragments de taille supérieure à 5 kb et exempt de contaminants qui pourraient interférer avec les étapes de formation des bibliothèques de séquençage.

### 2.2.2 Cas particulier des cellules du biofilm de cellulose

Il est à noter que la libération des cellules du compartiment biofilm composé de cellulose bactérienne, a nécessité un traitement préalable à l'extraction d'ADN. Parmi les différents protocoles décrits dans la littérature, nous avons opté pour

un traitement enzymatique à la cellulase qui permet de digérer complètement la cellulose (700 U pour 2 g de biofilm) (Figure 1.3).

Il est à noter que quelques biofilms très épais et très denses ont nécessité une durée plus longue de digestion. Il existe également des protocoles de broyage mécanique avec instruments qui ne sont pas disponibles au laboratoire, comme un disperseur Ultraturrax (Ika) ou un broyeur homogénéisateur Stomacher (Seward).

Grâce à Renaud Wagner (ESBS), nous avons pu comparer l'ADN extrait de cellules de biofilm récupérées après traitement avec un Ultraturrax à celui de cellules de biofilm récupérées par digestion enzymatique. Les quantités d'ADN extraites à partir de mêmes quantités de biofilm étant comparables, nous avons donc décidé de traiter tous nos échantillons par la digestion à la cellulase. Les cellules de 2 g de biofilm libérées après traitement à la cellulase sont récupérées par centrifugation à 10,000 g pendant 10 minutes. Le culot cellulaire peut être conservé à -20°C jusqu'au moment de l'extraction d'ADN.

**Au total, pendant la thèse, les extractions d'ADN de 213 échantillons ont été effectuées dont 173 de cellules de la phase liquide et 40 de cellules de biofilm.**

## 2.3 Génération des bibliothèques d'amplicons

### 2.3.1 Amorces de PCR.

La préparation des bibliothèques d'amplicons se fait en deux étapes : une première amplification de la région correspondant au marqueur taxonomique est effectuée au laboratoire (PCR1). Les amorces utilisées pour la PCR1 sont étendues en 5' et contiennent une séquence servant d'adaptateur pour une deuxième amplification par PCR effectuée par la plateforme de séquençage (PCR2). Lors de la PCR2, les index uniques de multiplexage et les index de séquençage Illumina sont ajoutés.

Les échantillons traités pendant la thèse ont été séquencés par deux plateformes de séquençage. La première plateforme sollicitée était la plateforme Genewiz et son service Amplicon-EZ (<https://www.genewiz.com/en-GB/Public/Services/Next-Generation-Sequencing/Amplicon-Sequencing-Services>). Les avantages de cette plateforme sont la possibilité d'envoi d'échantillons individuels, un temps de séquençage estimé court (2 semaines, normalement), un nombre minimal de séquences garanti à 50000 par échantillon. Les contraintes sont : la limite de taille maximale de l'amplicon, fixée à 550 pb, la préconisation de traiter un amplicon unique et de taille constante par essai. La nécessité de fournir des amplicons purifiés, avec une quantité minimale d'ADN de 500 ng et une concentration de 20 ng/μl est également contraignante lorsque les échantillons sont nombreux. Lors de l'avancement du projet, le nombre des échantillons est devenu important, et nous avons pu bénéficier des services de la plateforme INRAE de séquençage Miseq à Montpellier (UMR AGAP, Amélioration Génétique et Adaptation Des Plantes). Les échantillons d'une même série sont traités dans le même run de séquençage. Il n'y a pas de quantité d'ADN minimale exigée, la seule exigence est relative à la vérification sur gel d'électrophorèse du bon déroulement de l'amplification des régions cibles. La purification de la PCR1 est effectuée par la plateforme. Les amorces qui ont été utilisées afin de générer les amplicons sont décrites dans la Table 1.1 et les adaptateurs ajoutés en 5' des amorces sont décrits dans la Table 1.2. Les deux plateformes ont des stratégies différentes pour éviter la saturation du signal lors du séquençage d'amplicons. La stratégie de la plateforme de Montpellier implique d'ajouter dans nos amorces pour la PCR1 des positions dégénérées entre la séquence ciblée par les amorces de la PCR2 et la séquence ciblant le marqueur taxonomique.

Marqueur	Taxons ciblés	Orientation	Amorce originale	Référence	Séquence adaptée
V3V4	bactéries et archées	sens	Pro341F	Takahashi et al, 2014	(adaptateur1*)-CCTACGGGNBGCASCAG
V3V4	bactéries et archées	anti-sens	Pro805R	Takahashi et al, 2014	(adaptateur2*)-GACTACNVGGGTATCTAATCC
ITS1	levures et champignons	sens	ITS1	Bellemain et al, 2010	(adaptateur1*)-CTTGTCATTTAGAGGAAGTAA
ITS1	levures et champignons	anti-sens	ITS2	Bellemain et al, 2010	(adaptateur2*)-GCTGCGTTCTTCATCGATGC
D2	levures et champignons	sens	NL-D2	cette étude	(adaptateur1*)-AGAGACCGATAGCRAACAAGTA
D2	levures et champignons	anti-sens	NL-4	O'Donnell & Cigelnik, 1997	(adaptateur2*)-GGTCCGTGTTCAAGACGG

**Table 1.1.** les séquences cibles des amorces utilisées pour la PCR1, effectuée au laboratoire, selon le marqueur ciblé. La séquence de chaque amorce contient en 5', une séquence servant d'adaptateur pour la PCR2, effectuée par la plateforme de séquençage.

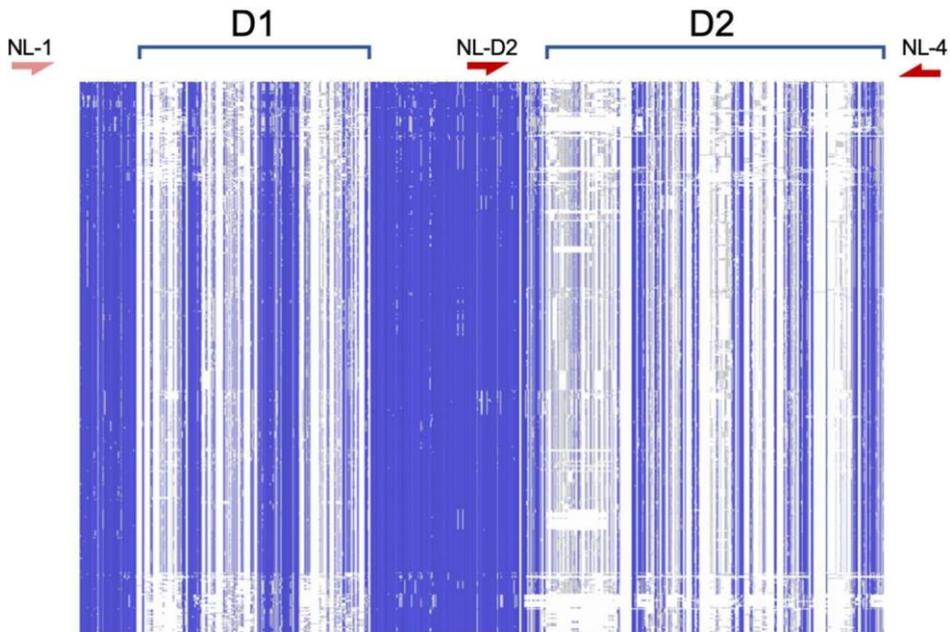
Adaptation	Séquence	Orientation	Plateforme
adaptateur 1-g	ACACTCTTTCCCTACACGACGCTCTCCGATCT	sens	Genewiz
adaptateur 2-g	GACTGGAGTTCAGACGTGTGCTCTCCGATCT	anti-sens	Genewiz
adaptateur 1-a	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNN	sens	Montpellier
adaptateur 1-b	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNDV	sens	Montpellier
adaptateur 1-c	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNNRM	sens	Montpellier
adaptateur 2-a	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNN	anti-sens	Montpellier
adaptateur 2-b	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNNH	anti-sens	Montpellier
adaptateur 2-c	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNNNW	anti-sens	Montpellier

**Table 1.2.** les séquences des adaptateurs ajoutés en 5' des amorces de la Table 1. Pour effectuer la PCR1 envoyée à la plateforme de Montpellier, il est nécessaire de préparer un mélange équimolaire de trois amorces sens et trois amorces antisens comportant chacune l'un des adaptateurs 1-a, 1-b, 1-c ou 2-a, 2-b, 2-c, respectivement. Les séquences dégénérées et les longueurs variables des adaptateurs évitent la saturation du signal lors du séquençage

### 2.3.2 Choix et validation des marqueurs taxonomiques

Comme cela a été décrit dans l'introduction, plusieurs marqueurs ciblant l'opéron codant les ARN ribosomiques peuvent être utilisés. Voici comment les marqueurs que nous avons utilisés ont été validés.

Dans les études de métagénomique ciblée en séquençage haut débit, les taxons de levures sont détectés dans des bibliothèques d'amplicons ciblant les marqueurs ITS1 ou ITS2 (Blaalid et al., 2013).



**Figure 1.5.** Alignement des segments D1-D2 de 527 espèces de levures et position relative des amorces NL-D2 et NL-4 utilisées pour cibler le marqueur D2. Lorsque le segment D1-D2 de 660 pb est ciblé, l'amorce sens NL-1 est utilisée (O'Donnell and Cigelnik, 1997). Les zones colorées en nuances de violet croissantes correspondent à celles de forte similarité de séquence, les régions très divergentes et non alignées apparaissent en blanc.

Le segment incluant les domaines hypervariables D1 et D2 est également un bon marqueur taxonomique (Kurtzman and Robnett, 1998) mais les études publiées l'utilisent soit avec la méthodologie de PCR-DGGE décrite au chapitre précédent, soit pour identifier l'espèce pour des isolats après amplification par PCR et séquençage de type Sanger. Les amplicons de ces marqueurs n'étaient pas compatibles avec les prestations de la plateforme de séquençage Genewiz. Les marqueurs ITS1 et ITS2 génèrent des tailles d'amplicons assez variables selon les espèces et le marqueur D1-D2 génère des amplicons de taille constante mais d'une longueur de 660 pb, au-delà de la limite de taille préconisée.

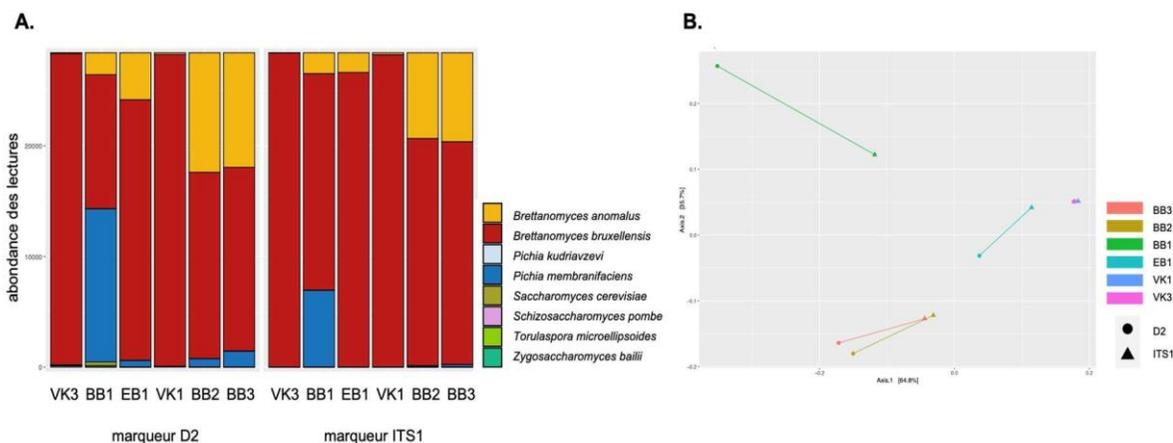
L'examen des alignements de séquences du segment D1-D2 correspondant à 527 espèces de levure (Kurtzman and Robnett, 1998) (Figure 1.5) nous a permis de définir une amorce interne dégénérée permettant d'amplifier uniquement le segment D2 pour ces 527 espèces, pour des amplicons de longueur constante d'environ 330 pb. Il a été décrit que des marqueurs différents peuvent donner des compositions en genres ou en espèces différentes pour un même échantillon (Ercolini, 2004). Cela est dû à une compétition entre les amplicons et les amorces au moment de l'étape d'hybridation lors de la PCR1. L'importance de cet effet est dépendante du marqueur utilisé (Ercolini, 2004). Afin de valider le marqueur D2, nous avons donc comparé les compositions microbiennes de six échantillons d'origine différente, déterminées avec le marqueur D2 ou ITS1, pour la même plateforme de séquençage (celle de Montpellier). Les compositions obtenues sont illustrées dans la (Figure 1.6A) et correspondent à des espèces de levures couramment décrites dans les SCOBY de kombucha (voir l'introduction). Ces compositions montrent des différences selon le marqueur utilisé. Avec le marqueur ITS1, seules trois espèces de la famille des Pichiaceae sont détectées sur l'ensemble des échantillons. Les espèces représentées par les amplicons minoritaires du marqueur D2 sont complètement absentes des compositions obtenues avec le marqueur ITS1 alors que la profondeur de lecture est du même

ordre de grandeur qu'avec le marqueur D2 (nombres moyens de lectures de 81500 et 72000, respectivement). Cela aboutit à des distances assez importantes entre les paires d'échantillons et des distances plus faibles entre les échantillons traités avec le marqueur ITS1 (Figure 1.6B). Nous avons effectué quelques vérifications pour essayer de comprendre pourquoi le marqueur ITS1 semble moins sensible que le marqueur D2. Un examen des séquences universelles des amorces utilisées pour amplifier le segment ITS1 (Bellemain et al., 2010) montre que, pour certaines espèces, les amorces ne sont pas parfaitement complémentaires à leur séquence cible. Il y a un mésappariement qui est assez proche de leur extrémité 3'. De plus, une analyse des séquences ITS1 des espèces détectées dans nos échantillons montre des différences de taille assez importantes allant de 80 à 100 pb pour les Pichiaceae de nos SCOBY, détectées avec les deux marqueurs, à 250 et jusqu'à 400 pb pour les espèces uniquement détectées avec le marqueur D2. Ces différences peuvent expliquer une sensibilité moindre du marqueur ITS1 par rapport au marqueur D2 dans le contexte de SCOBY de kombucha. Il est à noter que les variations de séquence du marqueur D2 des levures sont du même ordre que celles du marqueur V3-V4 des bactéries avec des divergences de 1 à 5% entre les segments D2 d'espèces du même genre et des divergences de 5 à 28% entre les segments D2 des espèces de genre différent. Pour le marqueur ITS1, les séquences des levures habituellement présentes dans les SCOBY sont très différentes entre les genres et même dans certains cas entre les espèces au point qu'elles ne peuvent pas être alignées entre elles. Ces résultats ont donc conforté notre choix d'utiliser le marqueur D2, inédit pour les études métagénomiques en séquençage haut débit, pour analyser et comparer les compositions en levures de tous nos échantillons. Pour l'affiliation des rangs taxonomiques des OTU, l'utilisation de ce marqueur a nécessité de formater et faire installer par les responsables de la plateforme FROGS, une banque représentée par les segments D2 des 527 espèces de levure décrits précédemment (Kurtzman & Robnet, 1998).

Cette banque n'est pas très exhaustive, il faut donc vérifier et ajuster manuellement, si nécessaire, l'espèce attribuée à certains taxons après l'étape d'affiliation par FROGS.

Pour déterminer les compositions en bactéries, nous avons choisi le marqueur V3-V4 qui est l'un des marqueurs les plus utilisés en métagénomique alimentaire (Takahashi et al., 2014). Les amplicons sont de taille constante d'environ 460 pb, compatibles avec les prestations habituelles des plateformes de séquençage.

La banque pour l'affiliation des rangs taxonomiques des OTU est la banque Silva 138 (<https://www.arb-silva.de/documentation/release-138/>), mise à jour le 16 décembre 2019.



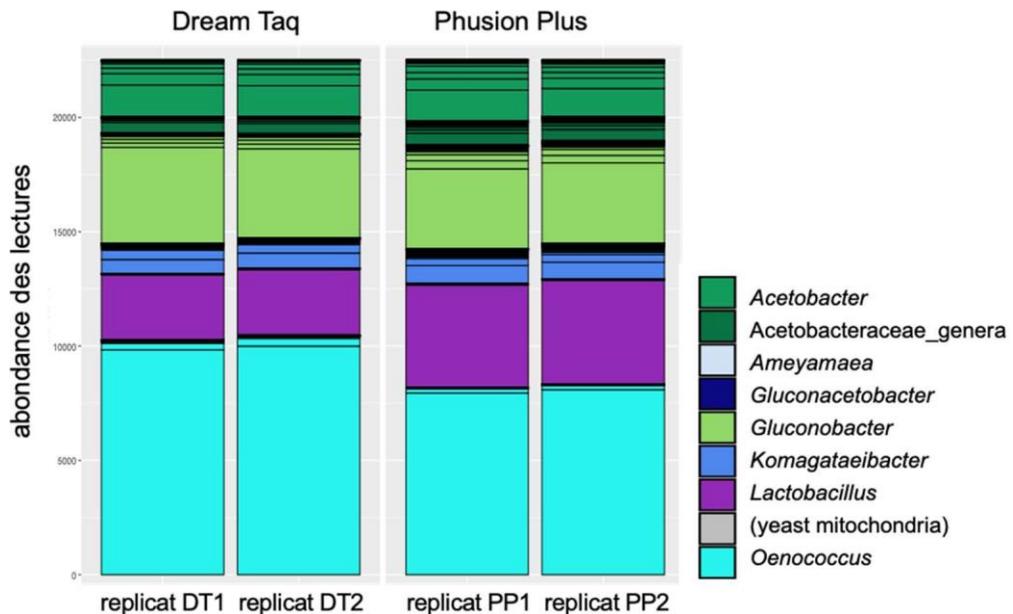
**Figure 1.6.** Comparaison des compositions en espèces de levures avec les marqueurs taxonomiques D2 et ITS1.(A) Comparaison des compositions en espèces de six échantillons d'origine différente : les échantillons sont nommés selon leur origine et disposés selon le marqueur taxonomique utilisé. La hauteur des boîtes représente l'abondance des lectures pour chaque espèce. Le nombre de lecture par échantillon a été raréfié dans le package Phyloseq. (B) Distances entre les six paires d'échantillons, calculées selon la dissimilarité de Bray-Curtis et représentées en 'multidimensional scaling' (MDS). Les couleurs sont utilisées pour différencier l'origine des échantillons et les formes des points pour distinguer le marqueur taxonomique utilisé. Une ligne relie les échantillons de même origine

### 2.3.3 Taq ADN polymérase pour générer les bibliothèques d'amplificons

De nombreuses études de métabarcoding établissent leurs bibliothèques avec une Taq ADN polymérase à haute-fidélité, pour limiter l'impact d'erreurs de PCR sur les données de séquençage. Comme les études du projet de thèse nécessitent de nombreuses amplifications, nous souhaitons vérifier la possibilité d'utiliser une Taq polymérase classique. Nous avons donc comparé les résultats de séquençage de bibliothèques effectuées avec le marqueur le plus long (V3-V4), pour la même préparation d'ADN issu de cellules du liquide de fermentation, lorsque la PCR1 est effectuée avec la Dream Taq polymérase (ThermoFisher Scientific) ou avec la Phusion Plus DNA Polymérase (ThermoFisher Scientific). Pour une taille de 460 pb, le fournisseur annonce que 30% des amplificons présentent au moins une erreur lorsqu'ils sont obtenus avec la Dream Taq polymérase alors que la fidélité de la Phusion Plus DNA Polymérase est 100 fois supérieure.

Dans l'illustration (Figure 1.7), les compositions en OTU affiliés à l'échelle genre sont comparées pour différentes bibliothèques. Ces compositions comprennent des genres de bactéries couramment décrits dans les SCOBY de kombucha (voir l'introduction). La comparaison comprend deux réplicats techniques où la PCR1 a été effectuée avec la Dream Taq polymérase et deux réplicats techniques où la PCR1 a été effectuée avec la Phusion Plus DNA Polymerase. Le séquençage a été effectué avec la plateforme Genewiz. On peut observer que les mêmes OTU ont été constitués dans chacun de ces quatre essais et que pour un genre donné, ce sont les mêmes OTU qui sont majoritaires. Certains OTU n'ont pu être affiliés qu'à l'échelle de la famille (ici, des OTU de la famille des Acetobacteraceae) mais leurs proportions sont similaires dans tous les essais. Quel que soit la Taq polymérase utilisée, d'éventuelles erreurs lors de l'amplification en PCR1 ont donc eu un effet négligeable sur les profils des OTU. Cependant, les proportions de certains genres diffèrent entre les essais avec la Dream Taq et ceux avec la Phusion Plus. Notamment, pour ces derniers, la proportion du genre *Oenococcus*

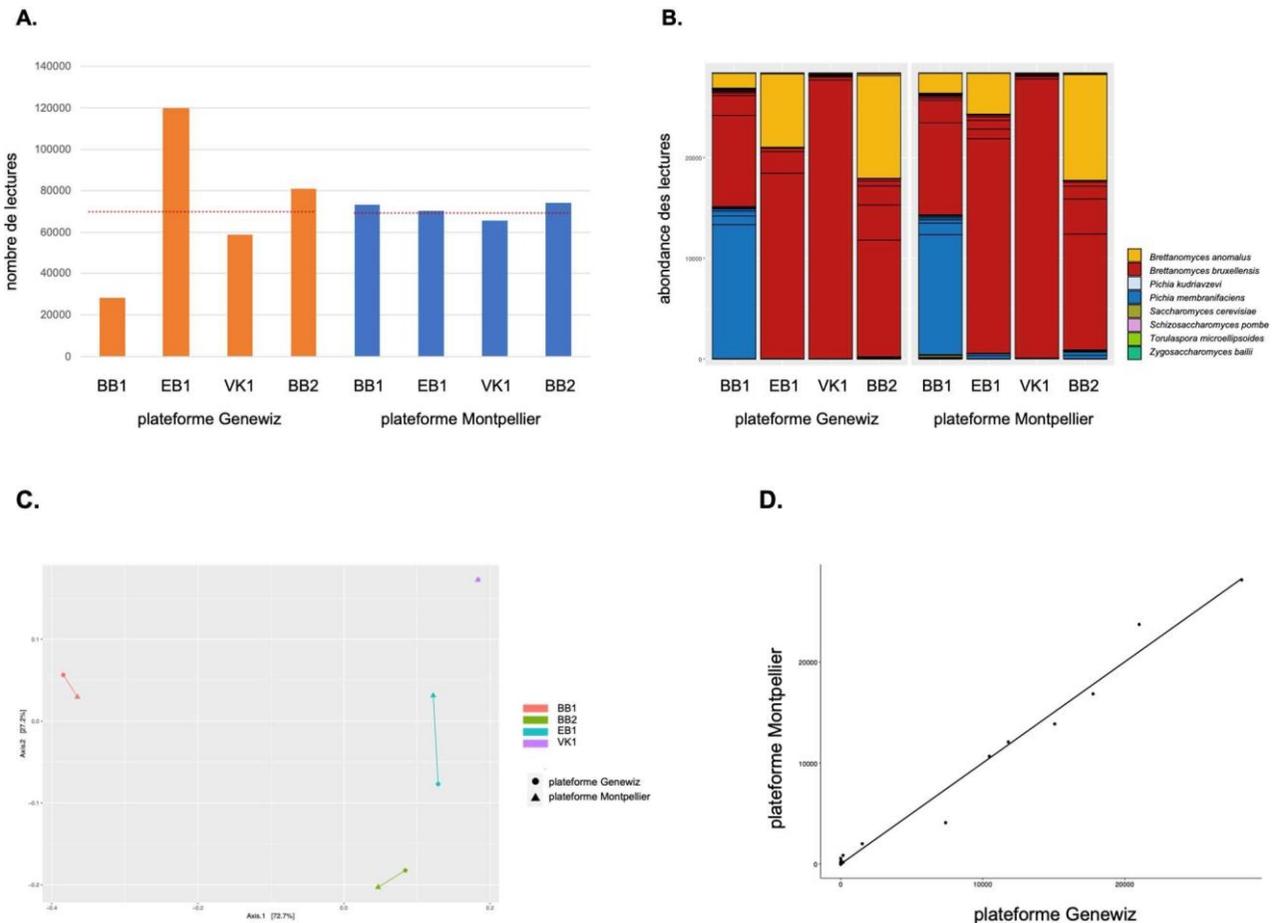
est moins importante, au profit de celle du genre *Lactobacillus*, qui sont les seuls genres représentant l'ordre des Lactobacillales dans nos échantillons. Les séquences consensus des OTU correspondants ne sont similaires qu'à 81%, ce qui permet d'écartier qu'une mauvaise affiliation due à des erreurs de PCR soit à l'origine de ces différences. Ces différences sont certainement dues à des comportements différents des deux polymérases, vis-à-vis de certaines séquences cible. De manière importante, ces différences rappellent que les compositions d'amplicons ne sont qu'un reflet, obtenu dans des conditions données, de la composition effective des cellules d'une communauté. Les bibliothèques nécessaires aux différents projets ont donc toutes été obtenues avec la même Taq polymérase, la Dream Taq polymérase.



**Figure 1.7.** Comparaison des données obtenues avec deux Taq polymérase différentes. Les données sont comparées pour la composition en genres de bactéries, pour le même échantillon et le même marqueur taxonomique V3-V4. La hauteur des boîtes représente l'abondance des lectures de chaque OTU et les couleurs correspondent au genre assigné à chaque OTU. Le nombre de lectures par échantillon a été raréfié dans le package Phyloseq.

## 2.4 Comparaison des données de séquençage générées par les deux plateformes

Les protocoles des deux plateformes utilisées pour le séquençage des bibliothèques d'amplificons diffèrent en plusieurs points, avec notamment des adaptateurs de séquence différente pour étendre les amorces d'amplification de la PCR1. Aussi, nous avons effectué une comparaison dans laquelle les bibliothèques du même marqueur, pour les mêmes échantillons, ont été séquencées avec les deux plateformes. Il s'agit de quatre échantillons d'origine différente : BB1, BB1, VK1 et EB1 pour lesquels une bibliothèque d'amplificons D2 a été générée pour chaque plateforme. Les résultats sont illustrés dans la (Figure 1.8A). Le nombre de lectures générées est plus variable d'un échantillon à un autre, pour les échantillons traités par la plateforme Genewiz. Cela peut s'expliquer par le fait que même si les bibliothèques sont envoyées en même temps, elles ne sont pas forcément traitées dans le même run de séquençage contrairement à la plateforme de Montpellier. Malgré ces variations en nombre de lectures, les profils taxonomiques de deux échantillons de même origine traités par les deux plateformes sont similaires à l'échelle OTU et à l'échelle espèce (Figure 1.8B). Les proportions des espèces majoritaires sont très similaires mais les espèces minoritaires sont mieux détectées pour les échantillons traités par la plateforme de Montpellier. Les distances entre les échantillons, calculées avec l'indice de Bray-Curtis, sont faibles, voire nulles pour les échantillons de même origine (Figure 1.8C). Elles sont inférieures à celle entre les échantillons d'origine différente, même pour les deux échantillons BB2 et EB1 dont les compositions sont les plus similaires. De plus, les compositions en espèces définies après séquençage par les deux plateformes sont corrélées à 99% ( $p < 2.2e-16$ , méthode de Pearson, Figure 8D), nous avons donc considéré qu'il est possible d'inclure dans une même analyse, les données d'échantillons traités par l'une ou l'autre de ces deux plateformes, pour le même marqueur taxonomique.



**Figure 1.8.** Comparaison des données obtenues avec les plateformes de séquençage : (A) Nombre de lectures par échantillon, selon la plateforme de séquençage, la moyenne de chaque lot d'échantillons est indiquée par la ligne rouge pointillée. (B) Comparaison des compositions en espèces des échantillons : les échantillons sont nommés selon leur origine et disposés selon la plateforme de séquençage. La hauteur des boîtes représente l'abondance des lectures de chaque OTU et les couleurs correspondent à l'espèce assignée à chaque OTU. Le nombre de lecture par échantillon a été raréfié dans le package Phyloseq. (C) Distances entre les quatre paires d'échantillons, calculées selon la dissimilarité de Bray-Curtis et représentées en 'multidimensional scaling' (MDS). Les couleurs sont utilisées pour différencier l'origine des échantillons et les formes des points pour distinguer la plateforme de séquençage. Une ligne relie les échantillons de même origine. (D) Corrélation positive entre les abondances des lectures pour chaque espèce, obtenues avec les deux plateformes.

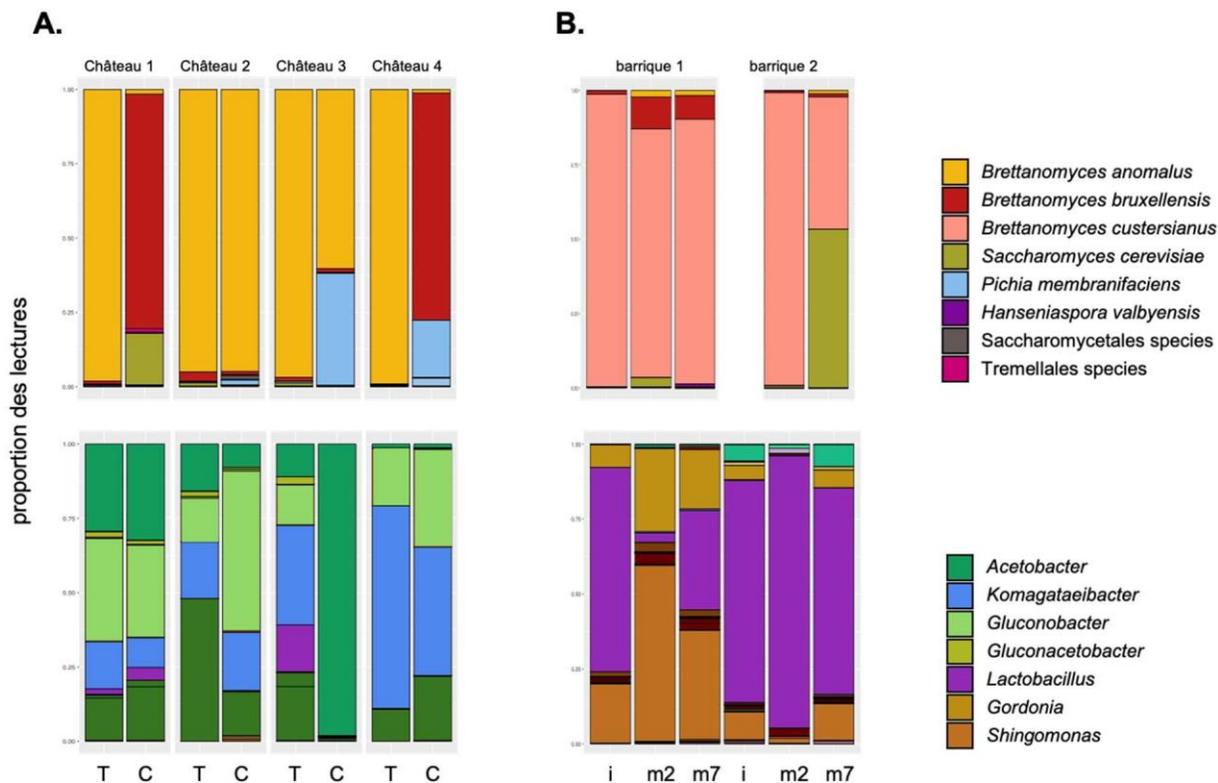
Au total, pendant la thèse, 420 échantillons ont été séquencés : 140 par la plateforme Genewiz et 280 par la plateforme de Montpellier.

### **3. Extension à d'autres communautés impliquant des levures du genre *Brettanomyces***

Certains projets du laboratoire s'intéressent aux levures du genre *Brettanomyces* et plus particulièrement à l'espèce *B. bruxellensis*. Cette espèce est importante pour d'autres boissons fermentées dont certaines bières, le cidre et le vin (Avramova et al., 2018). Pour la fermentation de bières et de cidre, l'espèce est intéressante, dans la production d'arômes ou pour libérer des molécules aromatiques des substrats végétaux fermentés. À l'inverse, pour la fermentation et la maturation des vins rouges, ces arômes sont souvent considérés comme des déviations organoleptiques non désirables et des mesures sont prises pour éviter la contamination aux différents stades de la fermentation, de l'élevage et de la mise en bouteille. Malgré ces mesures, les contaminations par des levures du genre *Brettanomyces* sont problématiques et peu prévisibles, ce qui fait que leurs conséquences ne sont souvent détectées que tardivement lors de l'élevage ou même bien après l'embouteillage. Il est également difficile de savoir pourquoi ces contaminations ne conduisent pas systématiquement à la formation d'arômes indésirables, et pourquoi, dans certains cas, des caractéristiques intéressantes peuvent être attribuées à ces contaminations. Devant la complexité de cette situation, un des axes de recherche à développer consiste à mieux connaître la composition des communautés microbiennes de vin altérés ou non, afin de mieux comprendre les facteurs favorisant ces contaminations et ceux qui conduisent à des effets considérés comme intéressants ou problématiques sur le goût des vins. Nous avons donc testé si notre protocole de métabarcoding pouvait s'appliquer à des échantillons de vin ou de bière. L'Institut de la Vigne et du Vin à Bordeaux,

avec laquelle notre équipe collabore, a pu échantillonner des cellules de différents vins contaminés ou non. Ces échantillons ont été complétés par des échantillons d'un brasseur de bière élevée en fût. Nous avons traité ces échantillons avec notre protocole de métabarcoding. Contrairement aux fermentations de kombucha, les prélèvements sont peu fournis en cellules car lors de l'étape d'élevage, une partie de la flore microbienne a disparu. La plupart des échantillons ont tout de même permis d'obtenir des profils taxonomiques de leurs communautés en levures et en bactéries (Figure 1.9). Pour les échantillons de vin, ces profils présentent des similarités importantes avec ceux obtenus pour les fermentations de kombucha, avec la présence majoritaire des levures *B. bruxellensis* et *B. anomalus* et des bactéries de la famille des Acetobacteraceae (Figure 1.9A). Pour les échantillons de bière, les profils sont différents, avec une espèce de levure majoritaire *Brettanomyces custersianus* qu'on ne retrouve pas dans les communautés de kombucha et des bactéries majoritaires des genres *Gordonia* et *Sphingomonas* (Figure 1.9B). Ce que ces tests ont permis d'apporter à notre projet, est de valider la possibilité d'appliquer le protocole d'extraction à des quantités limitées de cellules, la possibilité de l'appliquer à d'autres boissons que le kombucha et de vérifier le pouvoir de résolution du marqueur D2 pour les espèces de levures. L'investissement consacré à la mise au point du protocole de lyse a permis d'établir un protocole de base qui pourra être appliqué à d'autres projets de l'équipe. Par ailleurs, pour les futurs projets sur ces communautés, ces tests très préliminaires ont aussi permis de mettre en évidence qu'il semble difficile d'interpréter les différences de proportions des taxa, entre vins témoins et vins considérés comme altérés. Les cellules de prélèvements systématiques effectués tout au long des étapes de vinification, du pressage aux stades de fermentation (lorsque la densité de cellule est élevée) et à l'élevage, pourraient être stockées et mises en attente avant d'analyser la composition microbienne par métabarcoding, jusqu'à la détection éventuelle d'altération dues à une contamination par

*Brettanomyces*. Ce design, incluant de multiples comparaisons sur de longues lignées temporelles devrait ainsi permettre de mettre en évidence s'il y a des liens entre la composition microbienne et les effets d'une contamination.



**Figure 1.9.** Compositions microbiennes de vin rouges et de bières selon notre protocole de métabarcoding. (A) Proportions des espèces de levure (marqueur D2) et des genres de bactéries (marqueur V3-V4) dans trois paires d'échantillons de vin rouge, composées d'un vin témoin (T) et d'un vin dont la contamination par *Brettanomyces* est établie (C). Les vins de chaque paire proviennent du même château et les vins des différentes paires proviennent de châteaux différents. (B) Proportions des espèces de levure (marqueur D2) et des genres de bactéries (marqueur V3-V4) dans deux barriques de bière provenant du même brasseur. Les prélèvements ont été effectués lors de la mise en tonneau (i) puis après 2 mois (m2) et sept mois (m7) d'élevage. Dans la légende, seules les taxons majoritaires sont indiqués.

# **Chapitre 2 : Exploration de la diversité taxonomique de SCOBY d'origines différentes**

## **Introduction**

Pour comprendre le fonctionnement d'une communauté microbienne dans un processus biologique donné, dans une niche écologique ou dans un système de fermentation, il est important de connaître tous les taxons impliqués : ceux qui sont strictement nécessaires au processus et ceux qui sont présents sans participer directement à ce processus. De même, il est important de déterminer quel degré de diversité taxonomique caractérise ces communautés : est-ce qu'il s'agit d'une diversité impliquant des taxons très différents pouvant assurer des fonctions similaires nécessaires à ce processus, ou ce processus est-il caractérisé par des taxons bien précis ?

Après avoir mis au point les méthodes de culture de kombucha et de caractérisation taxonomique par métabarcoding, un axe du projet consistait donc rassembler des SCOBY variés et d'analyser leur composition. Lorsque les travaux de thèse ont démarré, quelques études s'étaient déjà intéressées à la caractérisation des microorganismes impliqués dans la fermentation du kombucha, en utilisant des méthodes dépendantes de la culture et la méthode indépendante de métabarcoding. Ces études effectuées par des laboratoires ayant des sources de SCOBY différentes, montraient des différences dans les compositions microbiennes mais aucune comparaison systématique n'avait été réalisée. De plus, les SCOBY analysés étaient de nature très disparate : il s'agissait soit d'un nombre très restreint de SCOBY (Marsh et al., 2014 ; Reva et al., 2015 ; Chakravorty et al., 2016), soit d'échantillons prélevés chez des brasseurs aux pratiques diverses et directement analysés (Coton et al., 2017), soit de SCOBY cultivés au laboratoire mais avec des conditions différentes de thé et de sucres ou de stress (Gaggia et al., 2018 ; Podolich et al., 2017).

Afin d'avoir une connaissance approfondie de la diversité taxonomique des SCOBY, nous avons donc effectué une comparaison systématique de SCOBY

de différentes origines en analysant leur composition microbienne par métabarcoding. Les résultats sont rapportés dans un manuscrit qui va être soumis à une revue à comité de lecture : ces analyses ont permis de définir les taxons communs à tous les SCOBY générés, qui constituent donc le microbiome core. Il s'agit des espèces de levures *B. anomalus* et *B. bruxellensis*, ainsi que les genres de bactéries *Komagataeibacter* et *Acetobacter*. Les SCOBY élevés dans des conditions identiques, présentent une diversité apportée par les proportions relatives des taxons du core, ainsi que par la présence de taxons accessoires, spécifiques à certains SCOBY.

Ce manuscrit est inclus à la fin de ce chapitre du mémoire de thèse. Il est précédé de deux parties qui (i) détaillent comment la collection de SCOBY a été rassemblée et (ii) présentent des résultats préliminaires portant sur la caractérisation de souches isolées à partir de quelques un de ces SCOBY afin de constituer une collection de souches représentatives de la fermentation de kombucha.

## **Résultats**

### **Les starters utilisés pour générer les SCOBY**

Le kombucha est devenue une boisson très populaire, elle est fabriquée dans le monde entier et de très nombreuses marques vendent des boissons sous le nom de kombucha. Comme il n'y n'existe pas de réglementation ou d'appellation kombucha, ces boissons peuvent rassembler des produits assez hétérogènes dans leur composition chimique et biochimique, le seul point commun étant une concentration en éthanol suffisamment faible pour appartenir à la catégorie des boissons sans alcool (au maximum, 1.2 % v/v, en France). La présence de

microorganismes vivants, leur composition taxonomique et le processus de fermentation sont des caractéristiques très variables selon les marques.

Si on le souhaite, on peut produire son propre kombucha, comme cela a été pratiqué pendant des siècles. A notre époque, de très nombreux sites en ligne, d'amateurs ou de professionnels, contribuent à diffuser des connaissances sur le kombucha et dispensent des conseils et des recettes de fermentation (par exemple : <https://revolutionfermentation.com/blogs/kombucha/comment-faire-son-kombucha-maison/>). Il est possible de se procurer un SCOBY auprès d'un brasseur amateur de kombucha et aussi d'en acheter. Différentes marques fabriquent et commercialisent des SCOBY et il est très facile de les acheter en ligne. Généralement, les SCOBY commercialisés sont constitués d'un disque de biofilm plongé dans du liquide fermenté et permettent d'initier directement des back-sloppings selon la méthode traditionnelle. Il est aussi possible de générer un SCOBY en utilisant une boisson kombucha commerciale non pasteurisée pour inoculer une infusion de thé sucrée. Si la boisson a des qualités adéquates, avec un peu de patience (2-3 semaines), un fin biofilm se forme à la surface et un SCOBY complet devient utilisable après un nouveau cycle de fermentation.

Dans ce contexte, nous avons constitué une collection de SCOBY à comparer, en essayant d'avoir une large diversité d'origines (Table 2.1). Bio Brasseurs est bien sûr le fournisseur principal de starter pour les études de la thèse. Un des SCOBY inclus dans cette comparaison (BB1) a été généré directement à partir d'un prélèvement effectué sur le rétentat de filtration d'une cuve de fermentation. Au moment où la thèse a démarré, début 2019, l'offre de SCOBY de kombucha commercialisés en France n'était pas très étendue. Nous avons pu rassembler six SCOBY élaborés par des producteurs professionnels. Le nombre de producteurs a augmenté depuis, et nous pourrions facilement tripler notre collection aujourd'hui. Pour augmenter les comparaisons, un des SCOBY (KF2) a été

recommandé chez le même producteur, une année plus tard et un autre (FE2) a été remis en fermentation après conservation d'une aliquote pendant une année en chambre froide. Pour augmenter notre collection, nous avons également généré des SCOBY à partir de boissons commercialisées : certaines marques, fabriquées en France ou à l'étranger, sont assez largement distribuées en ligne, en grande distribution ou par les chaînes de magasins 'bio'. D'autres marques ont une distribution plus confidentielle et ont pu être découvertes dans des boutiques locales, au gré de déplacements professionnels ou de loisir. Dix boissons ont donc pu être testées comme starter pour générer des SCOBY, en incluant une boisson fabriquée par Bio Brasseurs (BB2) pour comparer le SCOBY généré avec celui issu des lignes de fermentation (BB1). Le partage de SCOBY est très pratiqué par les amateurs de kombucha, via les sites de petites annonces ou le bouche à oreille : quatre starters ont ainsi pu être ajoutés à la collection, dont deux ont été procurés par des membres du laboratoire (AK et PK). Au total, nous avons ainsi rassemblé 23 starters qui ont été mis en fermentation pour générer les SCOBY à comparer.

### **Génération d'une collection de SCOBY comparables**

Les starters rassemblés étaient de formes très différentes, alors que nous souhaitons effectuer des comparaisons sur des communautés cultivées dans les mêmes conditions de fermentation. Nous souhaitons également avoir des communautés qui puissent se propager au cours de plusieurs cycles de fermentation successifs puisque la méthode de back-slopping est une des caractéristiques des SCOBY de kombucha. A partir de ces starters, une première étape a consisté à générer des SCOBY de départ comparables (voir ci-dessous). Dans un deuxième temps, ces SCOBY de départ ont été propagés par back-slopping traditionnel (inoculum constitué du biofilm fille et du liquide de

fermentation) et les prélèvements à effectuer ont été effectués sur trois cycles successifs de fermentation.

Nom	Format du starter	Origine du starter	Origine de fabrication	Biofilm formation	3 backsloppings	Contamination	SCOBY
<b>AK</b>	liquide + biofilm	producteur amateur	France (Ardèche)	√	√	x	√
<b>BB1</b>	suspension cellulaire	brasserie	France (Alsace)	√	√	x	√
<b>BB2</b>	boisson non pasteurisée	commercialisé en France	France (Alsace)	√	√	x	√
<b>BK</b>	boisson non pasteurisée	commercialisé en France	France	√	√	x	√
<b>CP</b>	boisson non pasteurisée	commercialisé en France	Portugal	√	√	x	√
<b>DK</b>	cellules lyophilisées	commercialisé en France	France	x	√	x	x
<b>EB</b>	boisson non pasteurisée	commercialisé en France	France (Aquitaine)	√	√	x	√
<b>FE1</b>	liquide + biofilm	commercialisé en France	Etats-Unis	√	√	x	√
<b>FE2</b>	liquide conservé	commercialisé en France	Etats-Unis	√	√	x	√
<b>FZ</b>	boisson non pasteurisée	commercialisé en France	France (Aquitaine)	√	√	x	√
<b>HK</b>	liquide + biofilm	producteur amateur	France (Alsace)	√	√	x	√
<b>KF1</b>	liquide + biofilm	commercialisé en France	Espagne	√	√	x	√
<b>KF2</b>	liquide + biofilm	commercialisé en France	Espagne	√	√	x	√
<b>KK</b>	boisson non pasteurisée	commercialisé en France	France (Centre)	x	√	x	x
<b>KOK</b>	suspension cellulaire	commercialisé en France	Slovénie	√	√	x	√
<b>LK</b>	liquide + biofilm	commercialisé en France	France	√	√	x	√
<b>LO</b>	liquide + biofilm	commercialisé en France	France	√	√	√	x
<b>MK</b>	liquide + biofilm	brasserie	France (Bordeaux)	√	√	x	√
<b>OK</b>	boisson non pasteurisée	commercialisé en France	Italie	x	√	x	x
<b>PK</b>	liquide + biofilm	producteur amateur	France (Alsace)	√	x	x	x
<b>RI</b>	boisson non pasteurisée	commercialisé en France	Canada	√	√	x	√
<b>UR</b>	boisson non pasteurisée	commercialisé en France	Suisse	√	√	x	√
<b>VK</b>	boisson non pasteurisée	commercialisé en France	France	√	√	x	√

**Table 2.1** Informations sur les 23 starters utilisés pour générer des SCOBY comparables. Le nom des starters ayant générés les 18 SCOBY de cette étude sont en gras.

Selon le type de starter utilisé, le protocole de génération des SCOBY de départ a été adapté comme suit (voir également la Figure n°1 du manuscrit en cours de rédaction). Pour les starters qui sont composés de biofilm et de liquide de fermentation, ils ont été soumis à un seul cycle de fermentation de 14 jours, avant

de démarrer les cycles de fermentations sur lesquelles les échantillons ont été prélevés. Ainsi, ces SCOBY sont dans le même état de fonctionnement malgré des temps de stockages des starters qui peuvent différer selon les fournisseurs.

Si le starter consiste en une suspension concentrée de cellules, il a été ensemencé dans une infusion de thé sucrée pour générer un biofilm, au bout de 21 jours. Les liquides de ces fermentations et les biofilms ont constitué les SCOBY de départ.

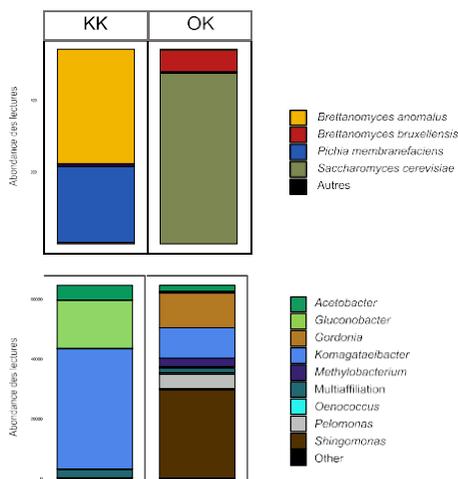
Si ce sont les microorganismes d'une boisson commercialisé qui ont servi de starter, ils ont été ensemencés dans une petite quantité d'infusion de thé sucrée pour générer un fin biofilm au bout de 20 à 30 jours. Une deuxième fermentation a ensuite été effectuée pour obtenir un biofilm plus conséquent et disposer ainsi des SCOBY de départ. Pour deux boissons, il n'a pas été possible de générer de SCOBY de départ, notamment parce qu'aucun biofilm ne s'est formé au cours de différentes tentatives. Un des starters était constitué de microorganismes lyophilisés, il a été ensemencé dans une infusion de thé sucrée, selon les préconisations du fournisseur, mais aucun biofilm ne s'est constitué et son étude n'a pas été poursuivie. Pour deux starters, il n'a pas été possible de les maintenir au long de trois cycles de fermentation (PK et LO).

Nous avons ainsi pu générer 18 SCOBY de départ, propagés sur trois cycles consécutifs, pour analyser les compositions taxonomiques de 90 prélèvements représentés par les cellules des fractions liquides des trois cycles et celles des biofilms mère et fille du troisième cycle (voir la Figure n°1 du manuscrit en court de rédaction). On peut noter que toutes les fermentations présentaient des aspects visuels très similaires, à l'exception de celles des SCOBY KF1 et KF2, dont les biofilms étaient particulièrement épais et rigides. Ces biofilms étaient aussi particulièrement résistants à l'étape de digestion par la cellulase, qui permet de libérer les cellules, au point que nous n'avons pas pu récupérer les cellules des biofilms pour KF1.

Les compositions taxonomiques des échantillons prélevés ont été analysées par métabarcoding, en générant et séquençant les bibliothèques d'amplicons D2 et V3-V4. Ces résultats sont détaillés dans un manuscrit en préparation, ajouté à la fin de ce chapitre.

## Composition de fermentations issues de starters non inclus dans les comparaisons

Plusieurs starters représentés par les boissons étiquetées comme étant du kombucha, n'ont pas permis de générer des SCOBY pouvant être inclus dans nos comparaisons. Aucun biofilm ne s'est constitué, pourtant des microorganismes se sont multipliés, comme il a été possible de le constater, en effectuant l'observation microscopique d'états frais des prélèvements (non montrés).



**Figure 2.1.** Composition microbienne des fermentations de deux starters n'ayant pas abouti à la formation de biofilm. La composition en espèces de levures (en haut) est représentée par la bibliothèque d'amplicons D2. Le nombre de lecture a été normalisé selon l'échantillon présentant la plus faible profondeur de séquençage. La composition en genres de bactéries (en bas) est représentée par la bibliothèque d'amplicons V3-V4. Le nombre de lecture a été normalisé selon l'échantillon présentant la plus faible profondeur de séquençage.

Pour deux tentatives, nous avons néanmoins souhaité analyser leur composition microbienne par métabarcoding, afin de savoir si elle pouvait expliquer l'absence de biofilm. Les résultats obtenus (Figure 2.1) montrent que la fermentation conduite avec le starter OK présente une composition en levures et en bactéries qui n'est pas caractéristique de celle des SCOBY de kombucha, avec une abondance des genres bactériens *Sphingomonas*, *Pelomonas*, et *Methylobacterium* rarement décrits dans les compositions de SCOBY.

Les compositions de la fermentation avec le starter KK sont, quant à elles, très similaires à celles des SCOBY de kombucha, aussi bien pour les espèces de levures que pour les genres de bactéries. Cependant, les deux genres bactériens du microbiome core que nous avons défini, ainsi que l'une des espèces de levures du core, *B. anomalus* ou *B. bruxellensis*, sont présents. Il se peut que les conditions dans lesquelles ces starters ont été cultivés n'aient pas permis d'assembler un SCOBY typique de kombucha et/ou que la présence de certains taxons ait interféré avec cet assemblage. Pour ces deux communautés, une affiliation des taxons de bactéries à l'échelle espèce permettrait peut-être de mieux comprendre pourquoi il n'y a pas eu de formation de biofilm. Ainsi, il serait possible, notamment, de savoir si les espèces du genre *Komagataeibacter* sont des espèces habituellement présentes dans les SCOBY.

### **Isolement et caractérisation d'une collection de souches de levures et de bactéries à partir de différents SCOBY : un aperçu de la diversité taxonomique non révélée par la méthode de métabarcoding**

En parallèle des analyses par métabarcoding, des isollements ont été effectués, à partir de différents prélèvements effectués chez Bio Brasseurs (BB) et pour certains starters de la collection de SCOBY (EB, LO et MK). L'objectif de ces isollements n'était pas de caractériser les SCOBY par des méthodes dépendantes

de la culture, qui dans ce cas doivent être réalisées de manière très exhaustive, mais de constituer une collection de souches qui pourraient être utilisées dans des projets futurs, pour assembler des SCOBY synthétiques. Deux types de milieux gélosés ont permis d'isoler ces souches : le milieu YPD (glucose 2 % m/v, yeast extract 0.5 % m/v, bactopeptone 1 % m/v, agar 2 % m/v) et le milieu MRS (glucose 0.2 % m/v, bactopeptone 0.1 % m/v, beef extract 0.8 % m/v, sodium acetate 0.5 % m/v, yeast extract 0.2 % m/v, dipotassium phosphate 0.2 % m/v, ammonium citrate 0.2 % m/v, Tween 80 0.2 % v/v, magnesium sulphate 0.2 % m/v, manganese sulphate 0.005 % m/v) additionné de 10% (v/v) de jus de raisin filtré. L'aspect des colonies a permis de choisir des isolats correspondant à des aspects variés de taille, de couleur et de texture. Des observations microscopiques d'états frais ont permis de différencier les isolats de levure de ceux de bactéries, car ces milieux ont permis la multiplication des deux types de taxons. 135 souches ont ainsi pu être isolées et stockées dans un sachet conservé à -80°C : 70 isolats de bactérie et 65 isolats de levure.

Pour 42 isolats, l'identification au niveau espèce a pu être effectuée, par séquençage d'amplicons correspondant au gène complet de l'ARNr 16S pour les bactéries, ou au segment D1-D2 du gène de l'ARNr 25S pour les levures. Ces amplicons ont été amplifiés par PCR avec les amorces indiquées dans la (Table 2.2), à partir de lysats de cellules en suspensions denses, chauffées à 95°C pendant dix minutes. Le séquençage a été effectué par la méthode de Sanger automatisée, par la plateforme de séquençage Eurofins. Les résultats des identifications, consignés dans (la Table 2.3), montrent que les isolats caractérisés correspondent principalement à des taxons que nous avons définis comme constituant le microbiome core du kombucha. Cette analyse, à très petite échelle, permet également de voir que les genres de bactéries *Komagataeibacter* et *Lactobacillus* sont représentés par des espèces différentes au sein d'un même SCOBY.

Pour 18 souches de *B. bruxellensis* que nous avons isolées, le séquençage complet du génome a été effectué, en lectures courtes avec la technologie Illumina. Cette analyse a été réalisée dans le cadre d'un autre projet de l'équipe, portant sur une grande collection de souches de *B. bruxellensis*. Ces données de séquençage ont permis d'identifier à quelle sous-population de l'espèce appartiennent les souches que nous avons isolées (Table 2.3). Il est intéressant de remarquer que seule la moitié des isolats de *B. bruxellensis* séquencés ont des caractéristiques génomiques qui permettent de les affilier à la sous-population des souches de kombucha. Il est également remarquable qu'un certain nombre de souches ne sont pas affiliées à une seule sous-population mais constituent des souches mosaïques entre plusieurs sous-populations. Enfin, pour les isolats d'un même SCOBY, le fait qu'ils ne soient pas affiliés à la même sous-population indique que plusieurs souches distinctes de la même espèce sont présentes dans un même SCOBY, pour les quatre SCOBY concernés.

La caractérisation de cette collection est encore très préliminaire mais ces analyses ont leur importance pour la poursuite future du projet, montant l'importance de considérer la diversité taxonomique à différents niveaux.

Marqueur	Taxons ciblés	Orientation	Amorce	Référence	Séquence adaptée
ARN 16S	bactéries et archées	sens	27F	Miller et al, 2013	AGAGTTTGATCCTGGCTCAG
ARN 16S	bactéries et archées	anti-sens	1492R	Miller et al, 2013	GGTTACCTTGTTCAGACTT
D1-D2	levures et champignons	sens	NL-1	O'Donnell & Cigelnik, 1997	GCATATCAATAAGCGGAGGAAAAG
D1-D2	levures et champignons	anti-sens	NL-4	O'Donnell & Cigelnik, 1997	GGTCCGTGTTCAAGACGG

**Table 2.2.** Séquences des amorces pour le séquençage Sanger des amplicons de l'ARNr 16S et du segment D1-D2. Les amorces dégénérées ont été prolongées par des adaptateurs permettant le séquençage par les amorces universelles non dégénérées T7 ou SP6.

isolat	SCOBY	espèce	core/accessoire	sous-population
i001	BB	<i>Brettanomyces bruxellensis</i>	core	tequila/ EtOH-beer-Kombucha
i002	BB	<i>Brettanomyces bruxellensis</i>	core	Wine3
i003	BB	<i>Brettanomyces bruxellensis</i>	core	Wine3
i004	LO	<i>Brettanomyces bruxellensis</i>	core	Kombucha
i005	LO	<i>Brettanomyces bruxellensis</i>	core	Kombucha-Wine3
i006	LO	<i>Brettanomyces bruxellensis</i>	core	tequila/EtOH-beer
i007	LO	<i>Brettanomyces bruxellensis</i>	core	Wine3
i008	LO	<i>Brettanomyces bruxellensis</i>	core	Kombucha
i009	EB	<i>Brettanomyces bruxellensis</i>	core	Kombucha
i010	EB	<i>Brettanomyces bruxellensis</i>	core	Kombucha-Wine3
i011	EB	<i>Brettanomyces bruxellensis</i>	core	tequila/EtOH-beer
i012	EB	<i>Brettanomyces bruxellensis</i>	core	Wine3
i013	EB	<i>Brettanomyces bruxellensis</i>	core	Wine3
i014	EB	<i>Brettanomyces bruxellensis</i>	core	Kombucha
i015	EB	<i>Brettanomyces bruxellensis</i>	core	Kombucha-Wine3
i016	MK	<i>Brettanomyces bruxellensis</i>	core	Wine3
i017	MK	<i>Brettanomyces bruxellensis</i>	core	Kombucha
i018	MK	<i>Brettanomyces bruxellensis</i>	core	Kombucha-Wine3
i019	BB	<i>Brettanomyces anomalus</i>	core	–
i020	BB	<i>Brettanomyces anomalus</i>	core	–
i021	EB	<i>Brettanomyces anomalus</i>	core	–
i022	EB	<i>Brettanomyces anomalus</i>	core	–
i023	EB	<i>Brettanomyces anomalus</i>	core	–
i024	EB	<i>Brettanomyces anomalus</i>	core	–
i025	BB	<i>Saccharomyces cerevisiae</i>	accessoire	–
i026	BB	<i>Saccharomyces cerevisiae</i>	accessoire	–
i027	BB	<i>Schizosaccharomyces pombe</i>	accessoire	–
i028	BB	<i>Zygosaccharomyces lentus</i>	accessoire	–
i029	BB	<i>Hanseniaspora valbiensis</i>	accessoire	–
i030	BB	<i>Kluyveromyces marxianus</i>	accessoire	–
i031	BB	<i>Komagataeibacter hanseni</i>	core	–
i032	BB	<i>Komagataeibacter hanseni</i>	core	–
i033	BB	<i>Komagataeibacter hanseni</i>	core	–
i034	EB	<i>Komagataeibacter hanseni</i>	core	–
i035	BB	<i>Komagataeibacter hanseni</i>	core	–
i036	BB	<i>Komagataeibacter rhaeticus</i>	core	–
i037	BB	<i>Komagataeibacter saccharivorans</i>	core	–
i038	BB	<i>Komagataeibacter saccharivorans</i>	core	–
i039	BB	<i>Komagataeibacter saccharivorans</i>	core	–
i040	BB	<i>Liquorilactobacillus mali</i>	accessoire	–
i041	BB	<i>Liquorilactobacillus nagelii</i>	accessoire	–
i042	BB	<i>Liquorilactobacillus nagelii</i>	accessoire	–

**Table 2.3.** Taxonomie de 42 isolats de levures (noir) et de bactéries (bleu), à l'échelle espèce et à l'échelle sous-population.

## **Diversity in the microbial communities involved in kombucha fermentations performed with SCOBY from different origins.**

(Manuscrit en cours de preparation)

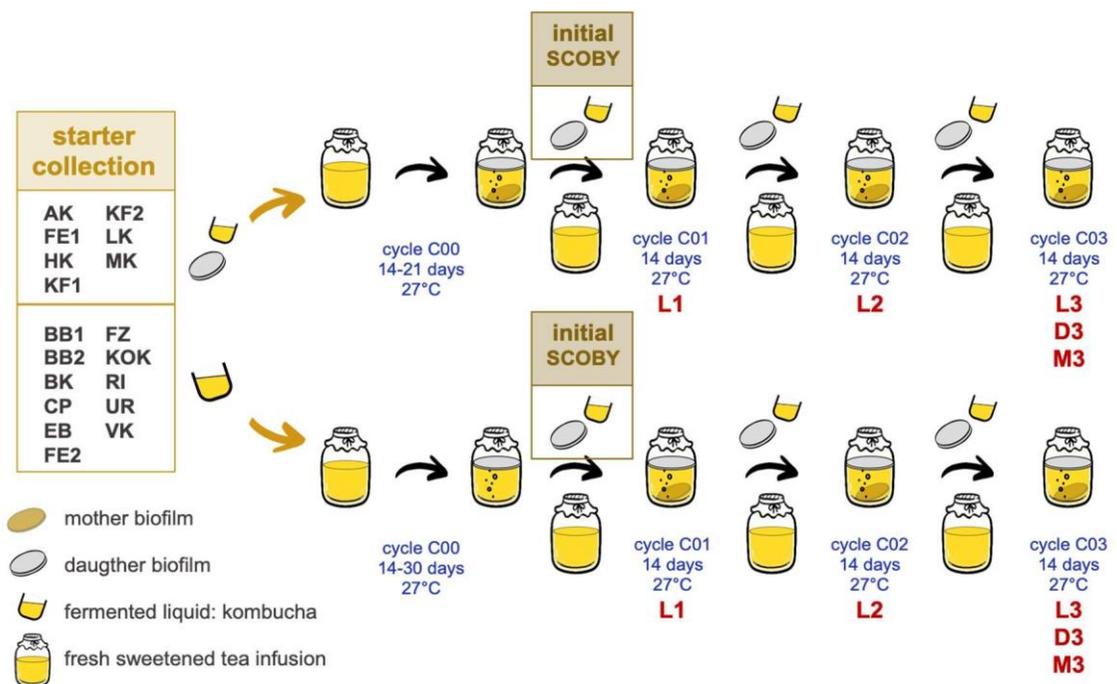
### **Introduction**

Multi-species microbial communities involved in human driven fermentation processes provide valuable models in understanding how the involved microbial communities are shaped through the interactions between microorganisms and between microorganisms and their surrounding environment (Wolfe and Dutton, 2015). A sample taken from a microbial community will be defined by its microbial composition, which consists in the list of the present taxa as well as their relative proportions. The number and proportions of taxa define the sample diversity expressed by alpha diversity indexes (Guiasu and Guiasu, 2010). The comparison of the compositions of samples involved in the same fermentation process but originating from diverse producers and locations, provides insights about the diversity of the microbial communities involved in a given process. This is an important task in order to distinguish the microorganisms that play a central role in the fermentation process in order to further study the interactions between the core and the accessoire taxa.

Microbial communities involved in kombucha fermentation, which are called SCOBY as for Symbiotic Culture Of Bacteria and Yeast, show several characteristics that may shape SCOBY with a high range of diversity and makes them interesting models to study microbe interactions (Huang et al., 2022; May et al., 2019). Kombucha is a popular effervescent beverage generated by the fermentation of a sweetened tea infusion, during an about fifteen days long process (Coton et al., 2017; Jayabalan et al., 2014; Júnior et al., 2022). It has very low alcohol content and tastes slightly sweet and acidic. Kombucha SCOBY

diversity can result from the fact that they are distributed worldwide, they can perform fermentation in diverse conditions of tea type and sucrose concentrations (Laavanya et al, 2021), and that they can be used at different production scales: they were used since centuries in households to produce small quantities of beverage (one to two liters), and they were increasingly used, since the last fifty years, by brewers that handle fermentation tanks reaching 1000 to 2000 liters (Coton et al., 2017; Harrison and Curtin, 2021). The kombucha SCOBYs display a rather complex shape in the fact that the microorganisms take place into two fractions representing two distinct environments: the liquid fraction, that will be consumed as the beverage, and the solid fraction that consists in a floating compact biofilm. The biofilm is made of reticulated cellulose synthesized during the fermentation process and containing embedded microbial cells (Tran et al., 2021). Finally, SCOBYs are propagated from a former fermentation cycle to the next one through a method called back-slopping. Back-slopping consists in using the raw products of a previous fermentation to inoculate the fresh substrate (Leroy and De Vuyst, 2004). The traditional and ancestral back-slopping method for SCOBY makes use of a bipartite starter that includes a portion of both the liquid and biofilm fractions (Jayabalan et al, 2014). Therefore, kombucha is as well the consumed product and a part of the starter used to initiate the next fermentation. The biofilm part of the starter is called the ‘mother’ biofilm. In the end of the fermentation process, a *de novo* formed cellulosic biofilm, the “daughter”, is generated at the air/liquid interface and it is taken as a new “mother” in the starter used to inoculate a fresh tea infusion. This process can be propagated in a continuous and virtually endless manner. On the other hand, the initial “mother” cellulosic biofilm can be reused as part of a starter. Therefore, an increasing number of available starters can be gathered along the successive fermentation cycles and kombucha amateurs take advantage of this characteristic to propagate and to share the SCOBYs among their communities.

With the aim of highlighting diversity among the essential microbial taxa involved in the kombucha fermentation process, we intended to draw comparisons between SCOBYs of different origins, in limiting the potential impact of the conditions in which they were stored or generated, furthermore this information is hardly available. Indeed, for kombucha fermentation, differences in the microbial



**Figure 1.** Generation of the initial SCOBY from starters of different origins and different shapes, followed by three consecutive traditional back-slopping to establish the fermentation series. The red labels indicate the fractions from which the cells were harvested to perform metabarcoding, corresponding to the fermented liquid (L), the daughter (D) and the mother (M) biofilms, at the end of the fermentation that follows the first, second and third back-slopping.

composition of SCOBY were described according to parameters like tea type, sugar amount and fermentation duration (Coton et al., 2017 ; De Filippis et al., 2018b; Gaggia et al., 2018 ; Reva et al., 2015), suggesting that associations between microbial composition and management practices can occur, like it was described for some fermented food (Ojeda-Linares et al., 2022). In addition, a recent study has shown that the viability of some species in kombucha decreases along long-term storage (Grassi et al., 2022).

Accordingly, in our study, we have analyzed the diversity in the microbial compositions of SCOBY that we generated from starters gathered from a wide range of different origins although they were elevated in identical fermentation conditions. We also investigated the microbial compositions along three successive cycles of propagation that were conducted in using the same traditional back-slopping method. Samples were analyzed from the liquid fraction of each cycle and from the mother and the daughter biofilms of the third cycle. Our results illustrate that while maintaining a core composition, the SCOBY differ in the respective proportions between the core taxa. This trait is propagated over cycles, suggesting it defines the SCOBY structure. In addition, their SCOBY differ by the presence of a few accessoire taxa that are not consistently maintained over successive back-slopping.

## **Results**

### **Sample collection**

In our study, we have compared the microbial composition between samples of 18 parallel series of kombucha fermentations. These series consist in the propagation of an initial SCOBY along three back-slopping steps performed in the same conditions (Figure 1 and Table S1). To generate the initial SCOBY of each series, a collection of 23 starters gathered from different origins was established. Each starter was cultivated in a sweet tea infusion and finally 18 of

the generated SCOBYs were appropriate to initiate a fermentation series, according to the following criteria. (i) The initial SCOBY is distributed in the liquid and solid fractions required for the traditional back-slopping method (ii) a daughter biofilm was generated at the end of each fermentation cycle, (iii) no visible contamination occurred along back-slopping and (iv) the biochemical composition of the liquid satisfied the standard of a kombucha beverage. In particular, the pH dropped to 3.5, the saccharose was consumed, ethanol was less than 1% (v/v), and acetate and lactate were produced (Suppl. Figure 1). Cells were collected at the end of each fermentation, from the liquid phase of three consecutive cycles, and from the mother and daughter biofilms of the third cycle. Total DNA was extracted and two amplicon libraries were generated for each sample: an amplicon library targeting the D2 hypervariable region of the 26S ribosomal RNA gene of fungi, and an amplicon library targeting the V3-V4 region of the 16S ribosomal RNA gene of bacteria. After sequencing and processing with FROGS, the D2 library showed 84, 612 mean reads/sample for 76 of the samples. The reads were clustered into 99 OTUs belonging to Ascomycota (98) and Basidiomycota (1). These OTU were further affiliated to 11 yeast genera representing 21 species. After normalization, 12,051 sequences per sample were kept for further analyses. The V3-V4 library showed 40, 550 mean reads/sample for 69 of the samples). After normalization, 7436 sequences per sample were kept for further analyses. The normalized reads are clustered into OTUs corresponding to 25 bacteria genera.

## Overview of the diversity in the microbial compositions

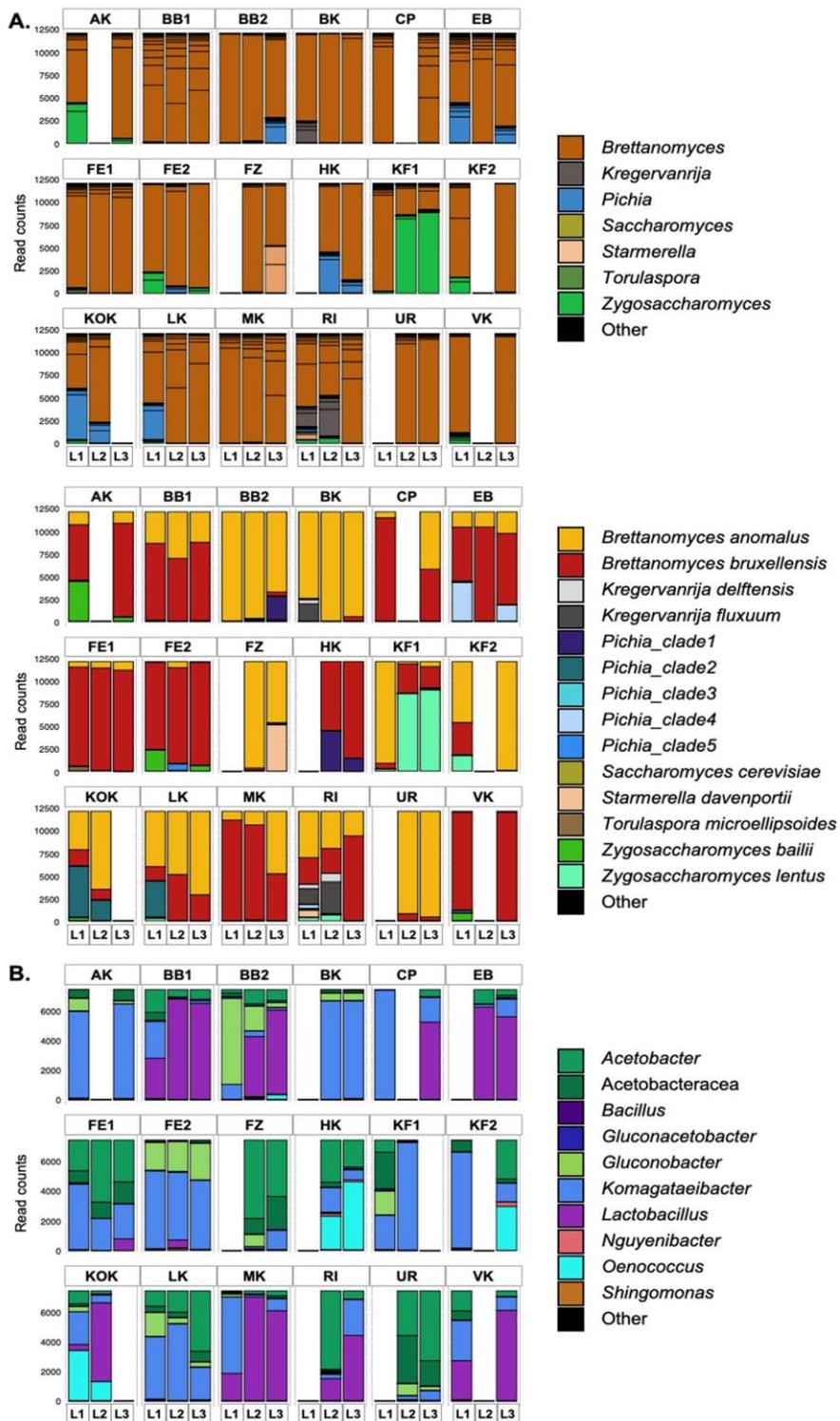
To compare the microbial compositions according to the starter origin, we focus on the libraries performed on the cells of the liquid fractions harvested over the three successive back-slopping. The yeast composition of the liquid fraction is available for 46 of the samples and it was resolved at the genus and at the species levels. The results are illustrated by the profiles of relative abundances of OTU affiliated to the genus level (Figure 2.A, upper panel) and to the species level (Figure 2.A, lower panel).

At the genus range, the taxonomic profiles show that the *Brettanomyces* genus is the only one that is shared by all the samples. It is represented by 24% to more than 99 % of the reads (mean 86.3%). Beside *Brettanomyces*, only four additional genera are present in the samples which overall represent between 1 to 5.5% of the reads. These genera are mostly conserved in many samples of the same starter origins, and a sample generally contains only one of these genera, with the exception of two samples of the RI origin that contains two to three additional genera. Two genera are shared by many samples of different origins: *Pichia* is shared by 13 samples of 9 origins and it is represented by 0.2% to 48.8% of the reads, and *Zygosaccharomyces* is shared by 11 samples of 7 origins and it is represented by 3.1% to 74.4% of the reads. Very few samples display 1 to 37% of the reads corresponding to the genera *Kregervanrija* (three samples of two origins), *Starmerella* (two samples of two origins), *Saccharomyces* (one sample) and *Torulaspota* (one sample). Finally the reads affiliated to the genera *Schizosaccharomyces*, *Kluyveromyces*, *Hanseniaspora* and *Malassezia* are detected in very low proportions, in very few of the samples. The genera number per sample ranges from one to ten with a mean of four genera and the alpha diversity Gini-Simpson index ranges from 0 to 0.53 with a mean of 0.16 and a median of 0.03 (Table S2). These values indicate an overall low diversity of the yeast compositions of the samples and this is mainly due to the fact that they are

dominated by one of the present genera. For all but three samples of KOK and KF origins, *Brettanomyces* is the major genera.

When going to the species range (Figure 2A, lower panel), only the most abundant genera are represented by several species, while the distributions of the minor genera are the same as at the species level. In all the samples, the *Brettanomyces* genus is represented by two species: *B. anomalus* and *B. bruxellensis*. Both species are detected in variable proportions: either the reads affiliated to *B. anomalus* are predominant, in 17 of the samples, or the reads affiliated to *B. bruxellensis* are predominant, in 23 of the samples, or they display rather similar proportions, in six of the samples. The genus *Zygosaccharomyces* is represented by two dominant species and these species are specific to the sample origin and are not shared by the same sample. *Z. lentus* is associated with the RI and KF origins (KF1 and KF2 series). *Z. bailii* is present in a few samples from the AK, VK, KOK, and FE2 origins. Not less than 33 OTU represent the *Pichia* genus and for more than half of them, the species identification by blastn is below 97%. Therefore, to improve the resolution of the *Pichia* distribution, we performed multiple alignments of the corresponding sequences, which highlighted that they are distributed in five clades that we have manually annotated as *Pichia\_clade1* to *Pichia\_clade5*.

Interestingly, this annotation revealed that each clade is specific to one or a few origins, and that a given sample contains only one major *Pichia* clade. For example, in the EB and RI origins, the *Pichia* genus is represented by reads of the sole *Pichia\_clade4*, and in the KOK and LK origins, it is represented by reads of the sole *Pichia\_clade2*.



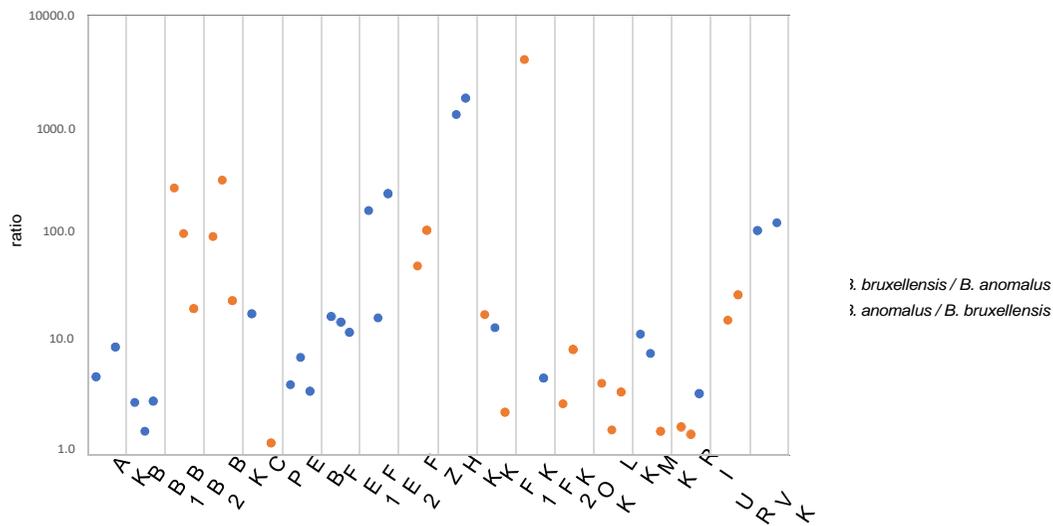
The bacteria composition of the liquid fraction over three successive back-slopping is available for 42 of the samples and it was resolved at the genus or at the family ranges. The taxonomic profiles are represented in (Figure 2B). Six main taxa (five genera and one family) are represented by 99% of the reads, among which three genera are highly shared by the samples. The *Komagataeibacter* genus is predominant, overall represented by 35% of the reads. It is detected in all the samples with read abundances that vary from 0.6 to 97.7%. The second genus shared by all samples is *Acetobacter*, which is represented by 18% of the reads with abundances that vary from 0.7 to 71%. The *Gluconobacter* genus is shared by six of the 42 available samples. It is represented by 7.4% of the reads with abundances that vary from 0 to 72%. In the VK origin (samples VK1 and VK3), no reads affiliated to the *Gluconobacter* genus were detected. All samples share taxa that belong to the Acetobacteraceae family, represented by 7% of the reads that could not be resolved at the genus level. The sample abundances vary from 0.6 to 43%. Two genera display highly contrasted distributions: *Lactobacillus* (represented by 0 to 91% of the reads) and *Oenococcus* (represented by 0 to 62% of the reads) are abundant in some of the samples and almost absent in the others. All the other genera are overall represented by less than 0.3% of the reads, with the read abundances in the samples that vary from 0 to 4%.

**Figure 2.** Microbial diversity and dynamics across the liquid fraction of the fermentation series of SCOBY generated by starters of 18 different origins. Read counts are normalized based on the sample that had the lowest number of sequences. Taxa are distinguished by a color scheme described at right. (A) Yeast diversity revealed by D2 metabarcoding analysis and resolved at the genus range (upper panel, abundance of OTU colored according to the genus) or the species range (lower panel). (B) Bacteria diversity revealed by V3-V4 metabarcoding analysis and resolved at the genus range.

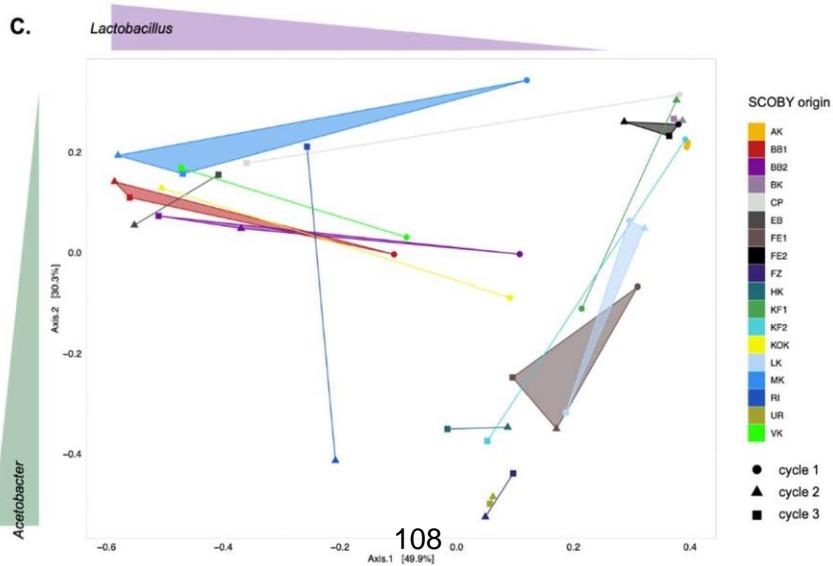
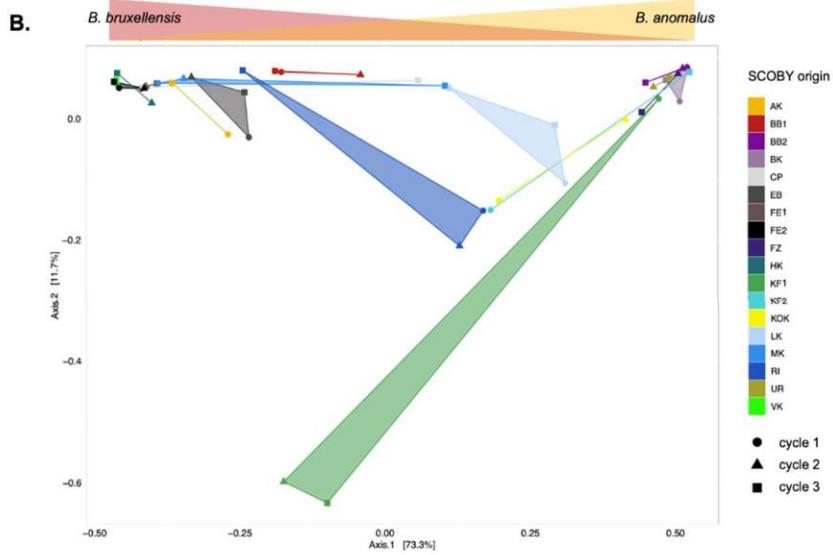
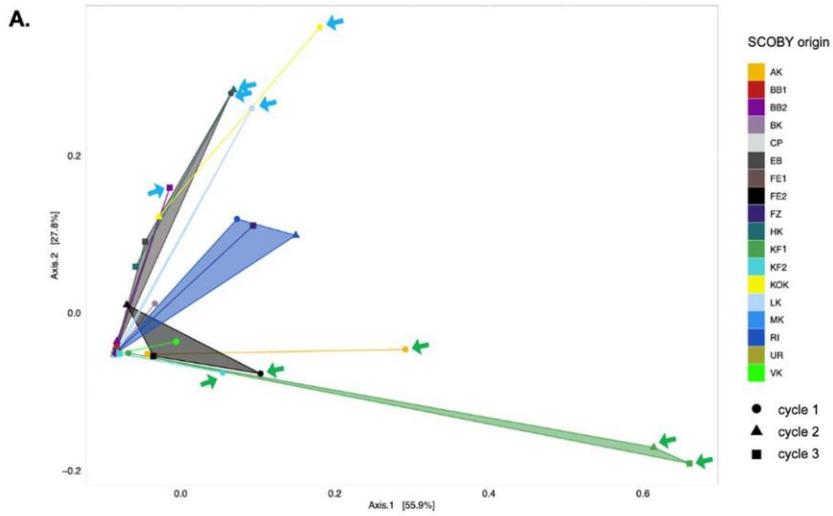
### **Microbial composition dynamics over successive back-slopping**

The microbial compositions were analyzed for three consecutive cycles. As mentioned before, the yeast compositions at the species range indicates that the ratio between the core species *B. anomalus* and *B. bruxellensis* are variable. Interestingly, for 13 of the 18 origins, the available samples of the same origin display similar values of *B. bruxellensis*/*B. anomalus* ratio whatever the back-slopping of the series (Figure 3). This means that this ratio is rather conserved across back-slopping and it thus suggests it defines a signature of the SCOBY structure. The diversity between samples, computed as the Bray-Curtis indexes and represented in multidimensional scaling (Figure 4), clearly shows that all the SCOBYs obtained from the starters of different origins are tightly grouped when diversity is computed at the genus range, with few distant samples due to the abundant representation of one additional genus with variable proportions across back-slopping (Figure 4A). SCOBYs are split in two main groups when diversity is computed at the species range (Figure 4B), with the samples of the two groups being distributed according to the respective proportions in *B. anomalus* and *B. bruxellensis* species. Six SCOBYs are clearly dominated by *B. anomalus* (BB2, BK, FZ, KOK, LK, UR), seven SCOBYs are clearly dominated by *B. bruxellensis* (AK, BB1, EB, FE1, FE2, HK, VK). Among the remaining SCOBYs, three exhibit proportions that are more fluctuating across back-slopping (CP, KF2, MK), and two have some of their samples with very high proportions of additional species (KF1, RI).

For bacteria compositions, the Bray-Curtis dissimilarity separates the SCOBYs in three main groups (Figure 4C) according to the evolution of the *Lactobacillus* and *Acetobacter* contents. For the SCOBYs characterized by the presence of the *Lactobacillus* genus, the contents are increasing from the first to the second cycle and are rather stable across the third cycle, for the available samples.



**Figure 3.** Diversity in the *B. bruxellensis* / *B. anomalus* ratio across SCOBY origins and its conservation across successive back-slopping. Each point corresponds to one sample with available *B. bruxellensis* / *B. anomalus* ratio (blue, computed when *B. bruxellensis* is dominant over *B. anomalus*) or *B. anomalus* / *B. bruxellensis* (orange, computed when *B. anomalus* is dominant over *B. bruxellensis*). The samples of the same origin are in the same interval and they are in the order of the fermentation cycles C01 to C03.



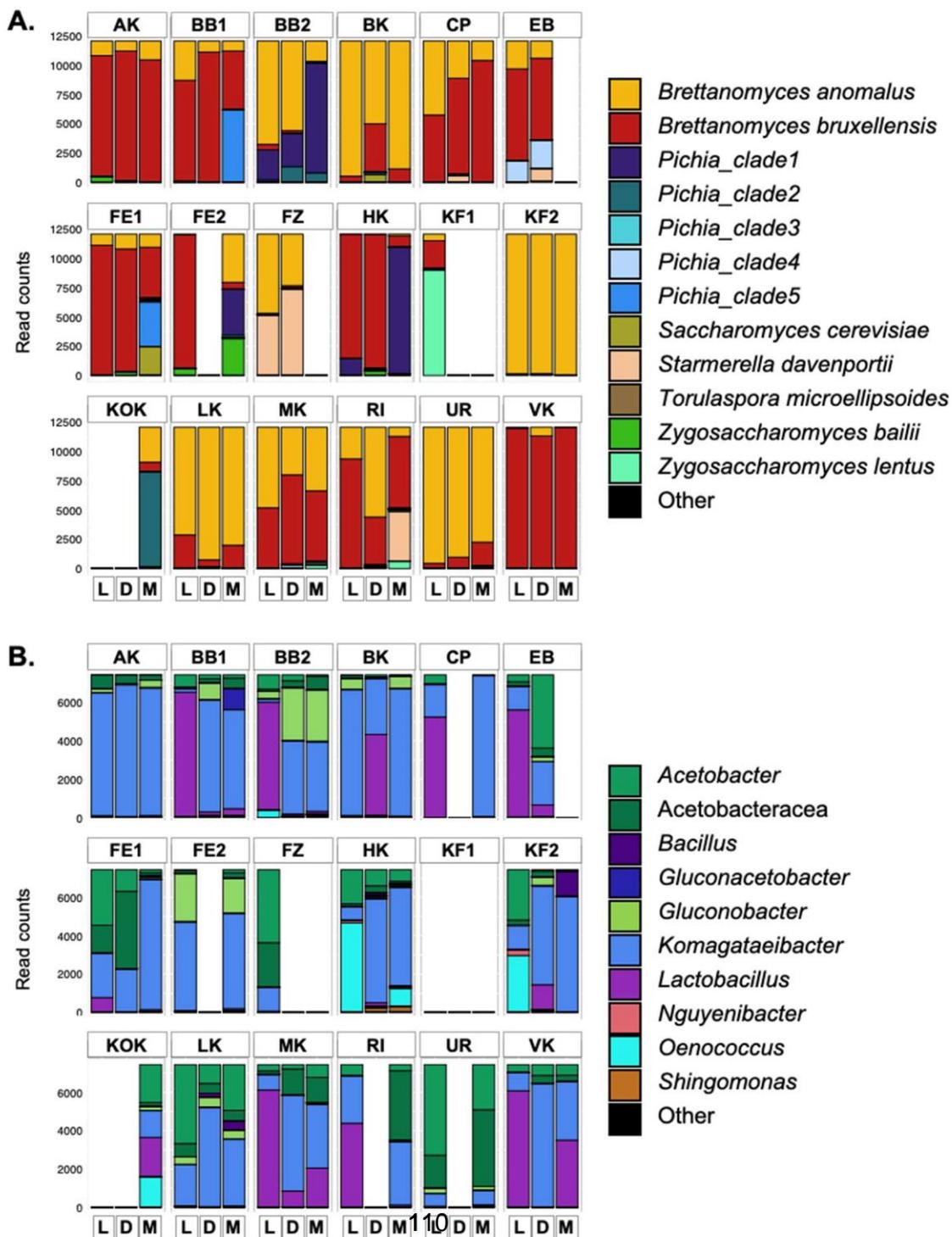
**Figure 4.** Spatial ordinations of the liquid fractions across successive back-slopping, based on Bray-Curtis dissimilarity matrices. Samples are colored according to the starter origin and the marker shape distinguishes the samples according to the fermentation cycle. The colored lines and polygons connect the samples from the same origin. (A) Dissimilarity computed at the genus range for the yeast compositions. The colored arrows point to the samples that display high proportions of the *Pichia* (blue) or *Zygosaccharomyces* (green) genera. (B) Dissimilarity computed at the species range for the yeast compositions. The distribution along axis\_1 is mainly driven by the decreasing abundance of *B. bruxellensis*, represented by the light-red triangle, and by the increasing abundance of *B. anomalus*, represented by the light-yellow triangle. (C) Dissimilarity computed at the genus range for the bacteria compositions. The distribution along axis\_1 is mainly driven by the decreasing abundance of *Lactobacillus*, represented by the light-purple triangle, and the distribution along axis\_2 is mainly driven by the increasing abundance of *Acetobacter*, represented by the light-green triangle.

### **Microbial compositions in the solid and liquid fractions**

The solid fraction represented by the cellulosic biofilm is a hallmark of the kombucha of fermentation. In the traditional process of fermentation, it is represented by two parts: (i) the mother biofilm that is used in the starter to inoculate a fresh tea infusion, and (ii) the daughter biofilm that is generated during the fermentation process and that is transmitted throughout the next back-slopping, and therefore it becomes the mother.

At the time of harvesting, the daughter biofilm was floating at the liquid-air interface of the fermentation while the mother biofilm had sunk to the bottom of the fermentation vessel. In order to compare their microbial composition, cells of the mother and daughter biofilms of the third cycle of each fermentation series were harvested after enzymatic digestion of the cellulose network and submitted to DNA extraction. The sequencing results of the D2 and V3-V4 libraries are available for 30 and 26 samples, respectively (Figure 5). For the available D2 library compositions, it is possible to compare the composition of 15 daughter

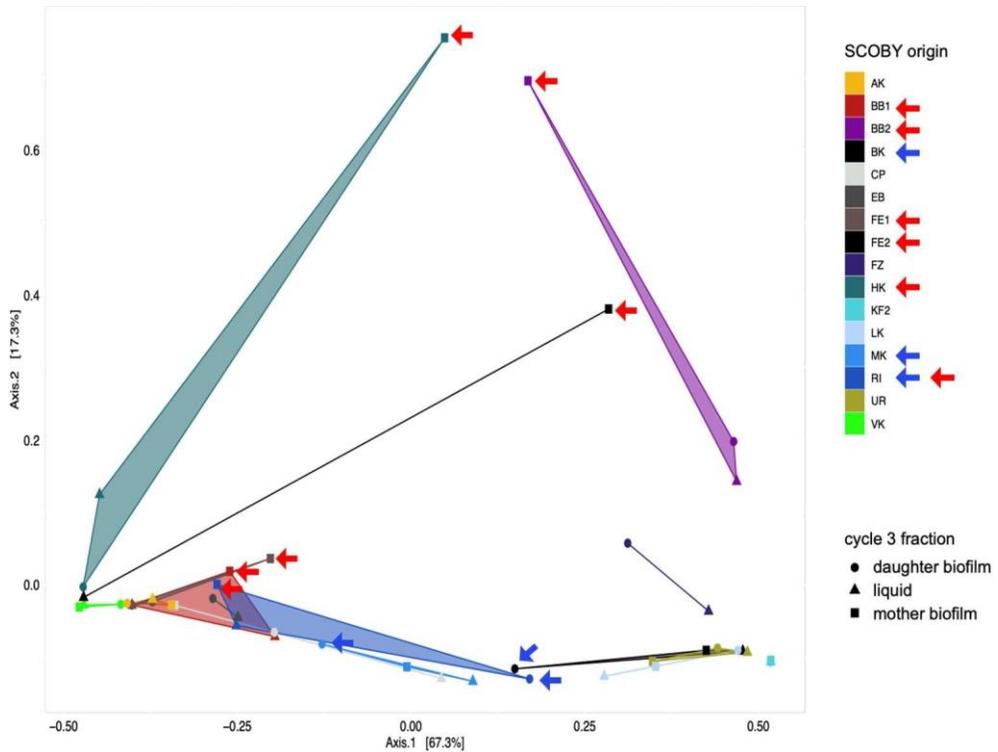
biofilms with the composition of the corresponding liquid fraction of the same cycle (Figure 5A)



For most of the comparisons, the reads affiliated to the predominant species display rather similar proportions, which is illustrated by the short distances on the multidimensional scaling plot established according to the Bray-Curtis beta-diversity indexes (Figure 6). The most divergent compositions are mainly characterized by an inversion in the *B. bruxellensis* / *B.anomalus* ratio, for BK, CP, MK and RI origins. Some accessoire minor taxa are also detected in slightly higher proportions in the libraries of the daughter biofilms.

According to the comparison with the corresponding liquid fraction of the same cycle, the available mother biofilm compositions can be split into two groups: for nine of them, the composition is similar to the corresponding liquid fraction of the same cycle. For six of them, the composition is significantly different, as illustrated by the long distances in Figure 6: they are characterized by high proportions (31 to 89.5% of the reads) of the *Pichia*\_clade1 and clade5 or the *Starmerella davenportii* species.

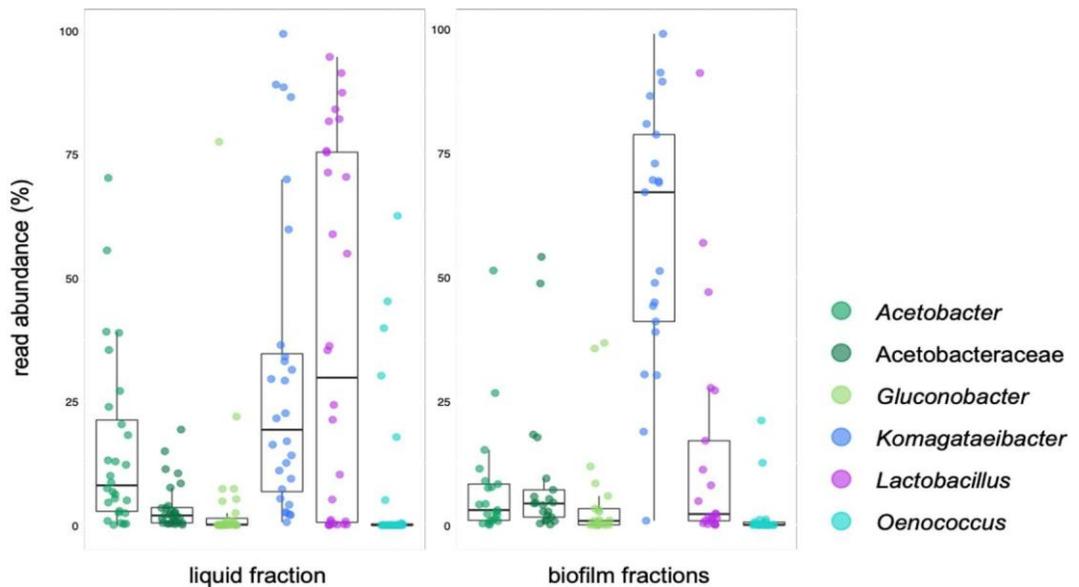
**Figure 5.** Microbial distributions between the liquid (L) and biofilm daughter (D) and mother (M) fractions during the third cycle of the fermentation series of SCOBYs generated by starters of 18 different origins. Read counts are normalized based on the sample that had the lowest number of sequences. Taxa are distinguished by a color scheme described at right. (A) Yeast diversity revealed by D2 metabarcoding analysis and resolved at the species range. (B) Bacteria diversity revealed by V3-V4 metabarcoding analysis and resolved at the genus range.



**Figure 6.** Spatial ordination plot between the yeast compositions in the different fractions of the fermentation cycle following the third back-slopping. Multidimensional scaling is based on a Bray-Curtis dissimilarity matrix built at the species level. The color distinguishes the samples according to their origin and the marker shape distinguishes the samples according to the fraction. The colored lines and polygons connect the samples from the same origin. The colored arrows point to the daughter (blue) or mother (red) biofilm compositions that are the most divergent from the corresponding liquid fraction. KF1 and KOK samples are not included in the plot because only one sample is available for these origins.

Among the available V3-V4 library compositions (Figure 5B), the visual comparison between the composition of 11 daughter biofilms and the corresponding liquid fraction of the same cycle, shows that similar compositions are only displayed for the samples of four origins: AK, FE2, LK and UR. These samples are the only one that do not contain reads affiliated to the *Lactobacillus* or to the *Oenococcus* genera. For the samples that contain *Lactobacillus* or *Oenococcus*, these genera are mainly detected in the liquid fraction. In these

samples, the relative abundance of the *Komagataeibacter* genus is the most affected. This results in an apparent preferential distribution of *Komagataeibacter* in the solid fraction, and an apparent preferential distribution of *Lactobacillus* and *Oenococcus* in the liquid fraction (Figure 7), however the confirmation of these preferences would require to estimate the total abundances of bacteria in the respective fractions.



**Figure 7.** Relative abundances of the major bacteria taxa plotted according to the liquid or biofilm (daughter and mother) fractions. Significantly higher abundance is observed for the *Komagataeibacter* genus in the biofilm fractions (p-val < 0.005) and for the *Lactobacillus* genus in the liquid fraction (p-val < 0.01). The samples of AK, FE2, LK and UR origins were not included in the plot because the *Lactobacillus* genus was not detected.

## Discussion

### **A small set of taxa defines the kombucha SCOBY core microbiome**

In this study, we have elucidated the microbial compositions of a collection of SCOBYs that have spontaneously assembled from kombucha starters of different origins. The conditions applied to the starters were strictly identical, in order to be able to attribute the observed diversity to the sole initial microbial compositions of the starters, including the taxa present and their respective physiological state at the time of inoculation. To our knowledge, among the many available studies on the microbial compositions of kombucha (Coton et al., 2017; Harrison and Curtin, 2021; Huang et al., 2022; Laavanya et al., 2021; Savary et al., 2021; Tran et al., 2020b), our study is the first one to have applied this experimental design to such a number of samples. We then submitted these microbial communities to two additional back-slopping in order to assess the stability of the deciphered microbial compositions, and we have also investigated the spatial distribution of the taxa between the liquid and the biofilm fractions of the SCOBY. We have developed an unprecedented taxonomic marker based on the D2 hypervariable region of the yeast 26S rRNA coding gene for metabarcoding that is resolved by high throughput sequencing. It differs from the ITS (Internal Transcribed Spacer) based markers usually used for fungi, in exhibiting a constant amplicon size (approx. 330 pb long) and an interspecies sequence variation pattern similar to the V3-V4 bacteria marker, while ITS sequences are so divergent that they cannot be aligned. It has a good resolution at the species level and we performed some comparisons indicating it shows a higher sensitivity than ITS1 in detecting yeast species that carry long ITS1 segments like *S. cerevisiae* or *S. pombe* (not shown). For the SCOBYs that were included in the study, we detected ten among the eleven most abundant yeast genera reported in the literature and all the bacteria most common genera in

kombucha. Only four taxa are shared between all the samples whatever the starter origin: the yeast *Brettanomyces* genus further represented by both the *B. anomalus* and *B. bruxellensis* species, and the *Komagataeibacter* and *Acetobacter* bacteria genera. These patterns of compositions suggest that a core microbiome can be defined for kombucha, and that both the yeast and the bacteria core microbiome are suitable with the model of “minimal core” microbiome as described by Hamady and Knight (Hamady and Knight, 2009) and defined as few components being shared by all, and many components being shared by a few others.

### **High fluctuations in the microbial compositions across back-slopping**

Some yeast accessory genera are shared by many SCOBYs (*Pichia* and *Zygosaccharomyces*) while some are shared by few SCOBYs. For bacteria, the highly abundant *Lactobacillus* accessory genera defines two main groups of SCOBYs. Among the samples of our study, the different groups that can be delimited in combining the diversity in bacteria and yeast, correspond to the four different SCOBY archetypes described by Harrison and Curtin (2021). These archetypes correspond to variations on the core microbiome that represents either the main component of the SCOBYs, or in which the *Zygosaccharomyces* or *Starmerella* genera are dominant over *Brettanomyces*, or in which *Lactobacillus* is dominant over *Komagataeibacter*. In addition, our analyses performed at the species range highlighted an additional level of diversity for the yeast core that is dominated either by *B. anomalus* or by *B. bruxellensis* core species. Importantly, unlike the yeast accessory genera which tend to decrease in the liquid fraction, the dominant core species are stable across back-slopping. This raises the issue of the capacity of the accessory yeast taxa to permanently colonize the kombucha niche. In order to understand the highly overlooked interactions between the core and the accessory taxa, it will be of high interest to investigate if the sustainable presence of accessory taxa is dependent on the fermentation conditions, if the

biofilm fraction in which they remain abundant may act as reservoir, or if they need being recurrently reintroduced from the environment to recolonize the SCOBYs. In our study, the back-slopping steps were performed under sterility which would have prevented this kind of events that are undoubtedly common when handling SCOBYs in brewery or in households.

## **Material and methods**

### **The SCOBY collection**

We gathered a collection of 23 kombucha starters. The various starters were either purchased on French commercial platforms, or obtained from private households, or by kombucha brewers, all located in France. Because the gathered initial starters display disparate stages and shapes (Suppl. Table 1), they were handled with an appropriate protocol as follows, in order to generate a collection of comparable SCOBYs (Figure 1). The traditional back-slopping method for kombucha brewing requires an inoculum represented by the final products of a previous fermentation: a SCOBY made from both the biofilm and the liquid fractions. When the initial starters had both the biofilm and liquid fraction available, they were subjected to a single fermentation step to be ready to start the successive fermentation series. In order to generate a SCOBY from unpasteurized marketed kombucha beverages sold by various brands, we get the purchased bottle to settle during one hour after opening. We then inoculate one volume of the bottom phase to five volumes of sweet tea infusion and incubate the cultures for 20 to 30 days at 27°C. When this step resulted in the formation of a floating biofilm, it was used to re-inoculate a fresh sweet tea infusion, along with a fraction of the liquid phase. After 14 days of incubation at 27°C, these cultures resulted in a *de novo* biofilm suitable to start the back-slopping series, along with a fraction of the liquid phase. When the initial starter was a concentrated microbial suspension, it was subjected to a single fermentation step following the provider

specifications. After 14 to 21 days of incubation at 27°C, the *de novo* biofilm was ready to start the back-slopping series, along with a fraction of the fermented liquid.

### **Routine back-slopping, fermentation series and sampling**

A volume of 20 L of tea infusion was routinely prepared in a 30 L heating container. A mix of clean tap water and demineralized water (1:9 v/v) was brought to the boil and maintained for 10 minutes. 0.6 % w/v green tea leaves (Fannings vert, Jardins de GAÏA) were added and let steep for 15 minutes. After removal of the tea leaves, 4% w/v sugarcane (ActiBio), was added and stirred until dissolution. 1.1 L aliquots of the sweetened tea infusion were distributed in 1.7 L glass containers that were previously rinsed twice with simmering water and covered by a cotton towel fixed with a rubber band. Once it had cooled to room temperature, inoculation was performed according to the traditional back-slopping method for kombucha brewing, with a SCOBY represented by 100 mL of the liquid fraction and the “daughter” biofilm from a former kombucha fermentation. Fermentations were conducted at 27°C during 14 days under static conditions, and re-inoculated according to the traditional back-slopping method into a fresh sweetened tea infusion every 14 days. Our samples were harvested from parallel fermentation series, each series was conducted during three successive back-slopping (Figure 1).

Samples were taken at day 14: 50 ml of the liquid phase was collected after gentle homogenization within the container, and 5 g of the solid phases (mother or daughter) were collected by cutting an approx. 2 cm wide band of the biofilm on the edge of the disc with a sterile blade. To release the cells from the biofilm, the pieces of biofilm were treated with 240 mg/ml cellulase (Sigma-Aldrich, Denmark) for 4 h at 37°C. The cells of the liquid or the solid phase were passed through a 100 µm cell strainer (Corning, USA) to remove remaining small

cellulose aggregates. The filtrates were centrifuged at 4500g for 10 min at 4°C, and the drained cell pellets were kept frozen at -20°C until DNA extraction. 10 ml of the supernatant of the liquid phase was filtered via 0,45 µm sterile membrane and preserved at -20°C until biochemical dosages.

### **Biochemical composition**

The liquid fraction supernatants were defrosted at ambient temperature for two hours. pH was measured with a BioRAD pHmeter. Ethanol contents were determined using the K-ETOH enzymatic kit (Megazyme, LIBIOS, France). Sugar contents (sucrose, fructose, glucose) were determined using the K-SUFRG enzymatic kit (Megazyme, LIBIOS, France). Lactic and acetic acid contents were determined using the K-DLATE and the K-ACETRM enzymatic kits, respectively (Megazyme, LIBIOS, France). Samples dilutions and concentration calculations were done following the manufacturer's instructions. For each dosage, blank and standard solutions were prepared with samples and all the analyses were carried out in triplicate.

### **DNA extraction**

The Qiagen DNeasy PowerSoil Kit (Qiagen, France) was used to extract the total genomic DNA according to the manufacturer's instructions with two major modifications. (i) The Power Beads were replaced by 25 mg of 425-600 µm glass beads (Sigma, USA), and (ii) the lysing procedure was splitted in to three steps in order to perform sequential lysis as follows. A fraction of only 40 % of the volume of the lysis buffer was added to the lysing tube, supplied with 1/100 (w/v) of Sodium Dodecyl Sulfate. A first homogenization treatment was applied during 30 sec. at a 6 m/s speed with a FastPrep instrument (MP Biomedicals). The lysing tube was centrifuged for 2 min at 15000 g, the supernatant was collected and kept on ice. In a second step, a fraction of 30 % of the volume of the original lysis

buffer was added to the lysing tube and a new homogenization treatment, centrifugation and supernatant collection were performed in the same conditions. The second step was repeated: in total, three serial homogenization steps were performed, to ensure breakdown of the most solid cells and preservation of the DNA extracted from the most delicate cells. The pooled supernatants of each step were subjected to the supplier protocol for DNA purification.

### **Amplicon libraries and MiSeq sequencing**

To prepare the multiplexed Illumina libraries, we employed a strategy based on a two-step PCR approach: a first PCR using the locus-specific primers including the Illumina adapter overhangs (with 30 cycles), and a second PCR for the incorporation of Illumina dual-indexed adapters (with 12 cycles). Bead purifications were carried out after each PCR. Quantification, normalization and pooling were performed before sequencing on Illumina MiSeq.

For PCR1, the 16S V3-V4 region was amplified for bacteria and the D2 hypervariable region of the 26S ribosomal RNA gene for fungi. The D2 region was targeted with the PCR primers NL-D2 and NL-4, while the V3-V4 region was targeted with the PCR primers Pro341F and Pro785R (Suppl. Table 2). All PCR primers were extended with overhangs suitable for the sequencing platform (Suppl. Table 2).

The PCR1 reaction was performed using 50 ng of DNA template, and 1,25 units of DreamTaq polymerase (Thermo Scientific) in a final volume of 50  $\mu$ l. For each forward or reverse primer, an equimolar mix of was added to the PCR mix to a final concentration of 0.5  $\mu$ M. The amplification was performed with first a denaturation step at 95 °C for 2 min followed by 7 cycles of [denaturation (94 °C, 30 s) annealing (55 °C, 15s) extension (72 °C, 40 s)] and 23 cycles of [denaturation (94 °C, 30 s) annealing (62 °C, 20s) extension (72 °C, 40 s)]. A final extension step (72 °C, 1 s) then held at 12°C.

## **Bioinformatic and diversity analyzes**

Sequence processing was performed using the FROGS pipelines as developed by (Escudié et al., 2018) under Galaxy (Afgan et al., 2018). Paired reads were assembled by VSEARCH with a rate of mismatches in overlapping regions set at 10% (Magoč and Salzberg, 2011). Based on barcode, all the forward and reverse reads were assigned to different samples, then truncated by cutting off the primer and barcode. Primers were removed using Cutadapt 1.18 (Martin, 2011). Sequences were clustered using the Swarm algorithm (Mahé et al., 2015) with an aggregation distance of 3 into Operational Taxonomic Units (OTU). Chimera removal was conducted using VSEARCH with *de novo* UCHIME method (Edgar et al., 2011). The amplicon reference alignment and taxonomic assignment were performed using blastn as alignment tool. For the 16S OTU, the taxonomic assignment was performed using the non-redundant nucleotide database SILVA database version 138 (Quast et al., 2013). For the 26S OTU, the taxonomic assignment was performed using a custom database of 527 Ascomycete and Basidiomycete species sequences established according to (Kurtzman and Robnett, 1998) (Suppl. File 1). The taxonomic assignments were then manually refined as follows: if the value of blastn was lower than 97%, the OTU consensus sequence was subjected to BLASTn search against NCBI and if a better blast result was obtained, the corresponding taxa was used for assignment. As an example, two OTU were re-assigned to *Starmerrella davenportii* instead of *S. stellata*. Multi-affiliations of OTU were dealt with by assigning the lowest common taxonomy level to multi-affiliated OTU. Data were normalized based on the sample that had the lowest number of sequences using the `rarefy_even_depth` function of the R (v. 4.1.0) phyloseq package (v. 1.24.2) (McMurdie and Holmes, 2013).

## References

Les références de l'article sont répertoriées dans la section commune des références du manuscrit de thèse.

Name	Starter shape	Starter supply	Starter manufacture	Biofilm formation	3 backsloppings	Contamination
AK	liquid + biofilm	household	France (Ardèche)	√	√	x
BB1	cell suspension	brewery	France (Alsace)	√	√	x
BB2	unpasteurized beverage	commercial	France (Alsace)	√	√	x
BK	unpasteurized beverage	commercial	France	√	√	x
CP	unpasteurized beverage	commercial	Portugal	√	√	x
DK	lyophilized cells	commercial	France	x	√	x
EB	unpasteurized beverage	commercial	France (Aquitaine)	√	√	x
FE1	liquid + biofilm	commercial	US	√	√	x
FE2	liquid conserved	commercial	US	√	√	x
FZ	unpasteurized beverage	commercial	France (Aquitaine)	√	√	x
HK	liquid + biofilm	household	France (Alsace)	√	√	x
KF1	liquid + biofilm	commercial	Spain	√	√	x
KF2	liquid + biofilm	commercial	Spain	√	√	x
KK	unpasteurized beverage	commercial	France (Centre)	x	√	x
KOK	cell suspension	commercial	Slovenia	√	√	x
LK	liquid + biofilm	commercial	France	√	√	x
LO	liquid + biofilm	commercial	France	√	√	√
MK	liquid + biofilm	brewery	France (Bordeaux)	√	√	x
OK	unpasteurized beverage	commercial	Italy	x	√	x
PK	liquid + biofilm	household	France (Alsace)	√	x	x
RI	unpasteurized beverage	commercial	Canada	√	√	x
UR	unpasteurized beverage	commercial	Switzerland	√	√	x
VK	unpasteurized beverage	commercial	France	√	√	x

Suppl. Table 1. Starter origins, shapes

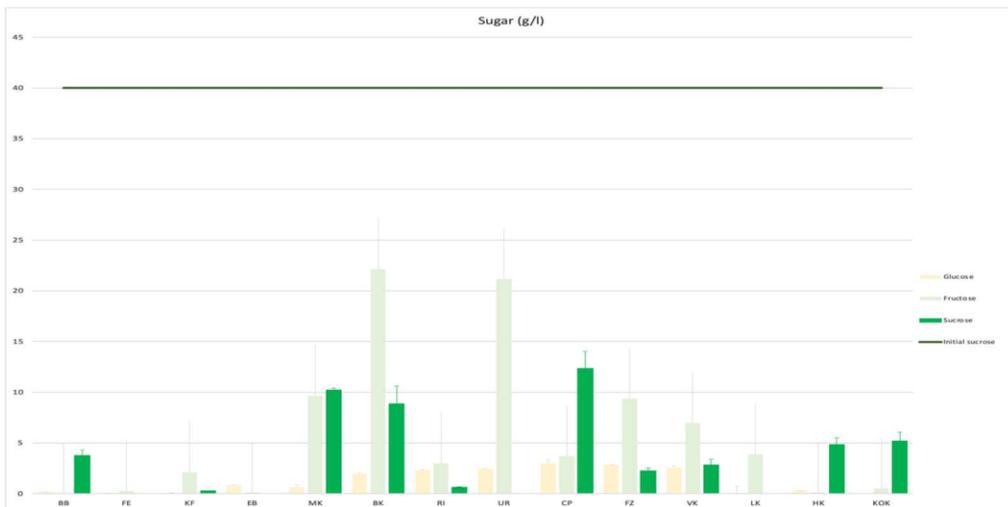
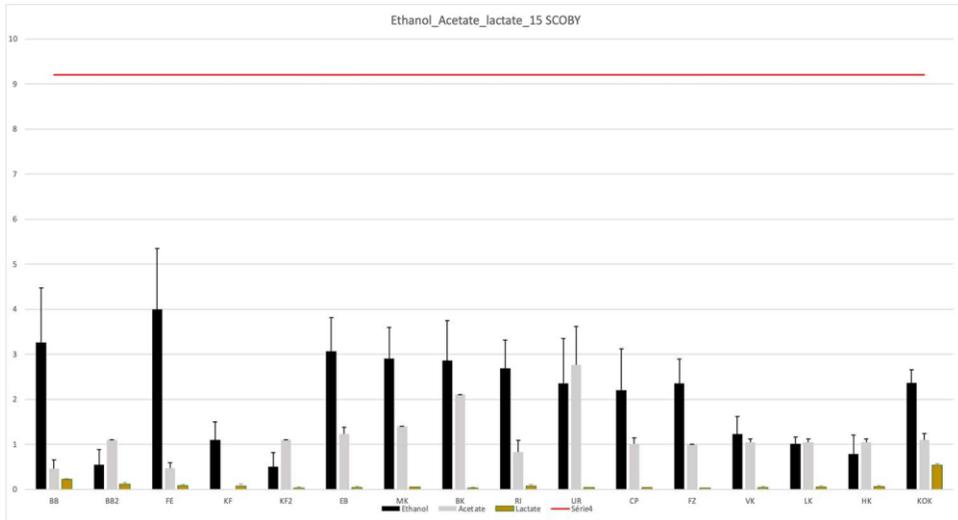
Sample	Observed (Genus)	Gini-Simpson (Genus)	Observed (Species)	Gini-Simpson (Species)
AK1	4	0.47	5	0.59
AK3	3	0.08	4	0.26
BB11	3	0.02	4	0.43
BB12	4	0.01	5	0.49
BB13	2	0.01	4	0.41
BB21	3	0.00	4	0.01
BB22	2	0.03	4	0.05
BB23	4	0.35	7	0.42
BK1	9	0.32	14	0.35
BK2	2	0.00	3	0.01
BK3	3	0.00	5	0.08
CP1	3	0.00	4	0.10
CP3	3	0.00	5	0.50
EB1	6	0.46	9	0.61
EB2	2	0.01	4	0.24
EB3	5	0.26	8	0.52
FE11	4	0.09	6	0.18
FE12	2	0.00	4	0.11
FE13	2	0.00	3	0.14
FE21	4	0.31	5	0.32
FE22	3	0.13	6	0.22
FE23	7	0.09	8	0.10
FZ2	5	0.01	7	0.05
FZ3	10	0.50	12	0.51
HK2	5	0.47	8	0.47
HK3	3	0.21	4	0.21
KF11	4	0.04	5	0.14
KF12	8	0.42	12	0.43
KF13	9	0.39	14	0.41
KF21	4	0.25	6	0.58
KF23	2	0.01	3	0.01
KOK2	3	0.53	4	0.63
KOK3	3	0.31	4	0.45
LK1	4	0.48	8	0.62
LK2	5	0.01	8	0.49
LK3	3	0.00	5	0.36
MK1	4	0.00	6	0.15
MK2	2	0.01	3	0.23
MK3	4	0.00	5	0.49
RI1	9	0.52	14	0.73
RI2	8	0.54	11	0.74
RI3	3	0.00	4	0.35
UR2	2	0.00	3	0.12
UR3	1	0.00	2	0.07
VK1	6	0.18	9	0.20
VK3	1	0.00	2	0.02

**Suppl. Table 2.** Diversity indexes for the yeast compositions of samples of the liquid fractions.

Library	Targeted taxa	Orientation	Original primer	Reference	Adapted sequence
V3-V4	bacteria	sense	Pro341F	Takahashi et al, 2014	(adapater1*)-CCTACGGGNBGCASCAG
V3-V4	bacteria	reverse	Pro805R	Takahashi et al, 2014	(adapater2*)-GACTACNVGGGTATCTAATCC
D2	fungi	sense	NL-D2	This study	(adapater1*)-AGAGACCGATAGCRAACAAGTA
D2	fungi	reverse	NL-4	O'Donnell & Cigelnik, 1997	(adapater2*)-GGTCCGTGTTCAAGACGG

Adaptation	Sequence	Orientation	Sequencing platform
adapater 1-g	ACACTCTTTCCCTACACGACGCTCTTCCGATCT	sense	Genewiz
adapater 2-g	GACTGGAGTTCAGACGTGTGCTCTTCCGATCT	reverse	Genewiz
adapater 1-a	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNN	sense	INRAE
adapater 1-b	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNDV	sense	INRAE
adapater 1-c	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNRM	sense	INRAE
adapater 2-a	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNN	reverse	INRAE
adapater 2-b	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNNH	reverse	INRAE
adapater 2-c	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNNNW	reverse	INRAE

### Suppl. Table 3. Primers



**Suppl. Figure 1.** Changes in the chemical composition ( $\text{g L}^{-1}$ ) in the liquid fraction of 16 SCOBYs from the collection. Upper panel: ethanol and organic acids. Colors reflects the compounds content: Ethanol (black), Acetate (grey) and Lactate (brown). The red line is the limit of 1,2% ethanol (v/v). Lower panel: Sucrose, glucose and fructose. Colors reflects the compounds content: Sucrose (green), Glucose (light orange) and Fructose (light green). The green line is the amount of sucrose added to the tea infusion inoculating the SCOBYs

# **Chapitre 3 : Dynamique des communautés microbiennes**

## 1. Introduction

Dans la littérature disponible sur les produits fermentés, lorsque l'évolution des compositions microbiennes est étudiée dans le temps, il s'agit généralement de son évolution entre le début et la fin du même cycle de fermentation. Peu d'études ont été menées pour les SCOBY de kombucha et ces études ont montré que, au cours d'un cycle, les proportions relatives des taxons majoritaires montrent de légères variations et aucun des taxons minoritaires ne montre d'expansion remarquable. Ce sont des études qui ont été menées sur un cycle d'un brassin industriel (Coton et al., 2017), ou sur un cycle de brassins de laboratoire correspondant à des SCOBY naturels (Marsh et al., 2014) ou synthétiques (Savary et al., 2021). Les taxons majoritaires détectés dans la fraction liquide de ces études étaient les genres *Brettanomyces* et *Zygosaccharomyces*, pour les levures, et *Gluconacetobacter* (qui correspond très probablement à l'actuel genre *Komagataeibacter*) pour les bactéries. Le genre *Lactobacillus* ne représentait un taxon majoritaire que dans l'étude de (Coton et al., 2017). Ces études, limitées aux variations au sein d'un seul cycle et effectuées avec des SCOBY peu variés suggèrent néanmoins une certaine stabilité de leurs compositions microbiennes. A titre de comparaison avec d'autres produits fermentés, de grandes variations dans les proportions des espèces bactériennes sont observées au cours des 48 heures du cycle de fermentation du kéfir de lait (Blasche et al., 2021). Cette différence de comportement est peut être simplement observée à cause des temps très rapprochés entre les prélèvements, ce qui n'a pas été reproduit pour les études sur les SCOBY de kombucha. Pour d'autres produits, fermentés spontanément par des communautés qui se constituent à partir de la microflore indigène, comme le kimchi (Zabat et al., 2018), il n'est pas surprenant d'observer de grandes variations de composition microbienne au cours d'un cycle, représentant la sélection des taxons impliqués dans la fermentation. Comme décrit dans

l'introduction, ce type de fermentation n'est pas comparable, dans son type de propagation des communautés microbiennes, à celle du kombucha.

A notre connaissance, aucune étude ne s'est intéressée, de manière approfondie et systématique, à l'évolution des compositions microbiennes lorsqu'elles sont propagées au cours de cycles successifs de fermentation, sur des échelles longues représentant une vingtaine de cycles (une année de production) ou plus. Cependant, dans le cas particulier où les communautés microbiennes sont propagées d'une fermentation à l'autre par back-slopping, la question de la stabilité des performances des communautés microbiennes est cruciale pour le maintien d'une production au cours du temps et la stabilité des compositions microbiennes pourrait être une cause importante du maintien de ces performances. En effet, comme détaillé dans l'introduction, pour les fermentations de type kombucha, kéfir et Gwell, les communautés ne s'assemblent pas de manière spontanée à partir de la microflore présente sur le substrat initial, et il n'existe pas de starters commerciaux équivalents à ceux des produits pour la fermentation des bières, des produits laitiers fermentés ou du pain. Si un cycle de fermentation échoue, le producteur ne pourra pas l'utiliser comme starter pour initier un nouveau cycle et cela peut conduire à une diminution conséquente de ses volumes de production (von Gastrow et al., 2022). Une des causes de la perte d'un starter de kombucha peut être la contamination accidentelle et facilement visible, par des insectes ou des moisissures. Une autre cause peut être la diminution des performances de la fermentation : lorsque cette situation se présente, il semblerait que le stockage temporaire au froid permette de réutiliser à nouveau le starter 'reposé' (source : BioBrasseurs). Les causes de ces baisses de performance ne sont pas connues. Si les compositions microbiennes sont en cause, s'agit-il d'une dérive progressive des compositions, qui aboutit à un déséquilibre avec une diminution des taxons importants dans la fermentation de kombucha (les genres de levures et de bactéries du core), au détriment de genres accessoires ? Ou s'agit-

il de contaminations microbiennes issues de l'environnement, qui interfèrent avec le fonctionnement des SCOBY ? Dans ce contexte, nous avons voulu explorer et comparer les compositions microbiennes au cours de fermentations successives. Nous avons mené cette exploration sur deux échelles : à grande échelle chez BioBrasseurs (cuves de 1000 à 2000 litres) et à petite échelle (1.2 litres) au laboratoire.

Comme nous l'avons décrit dans le premier chapitre, mais nous avons développé plusieurs applications pour utiliser les rétentats de filtration des cuves de fermentation de BioBrasseurs, qui sont des suspensions denses de l'ensemble des micro-organismes de la fraction liquide de la cuve, pour établir le design des expériences décrites dans ce chapitre :

- un rétentat de filtration est facile à homogénéiser, contrairement au contenu liquide de la cuve, il constitue donc un prélèvement de choix pour comparer les compositions de fermentations différentes. Les cellules des rétentats sont donc systématiquement récupérées pour déterminer leur composition microbienne par métabarcoding et suivre ainsi l'évolution de la production chez BioBrasseurs. Les résultats obtenus sont décrits dans la première partie de ce chapitre.

- un rétentat peut être utilisé comme starter en le répartissant dans des fermentations parallèles représentant des répliquats biologiques. Ces répliquats sont ensuite utilisés soit pour suivre la composition microbienne au cours des repiquages successifs, soit pour voir les effets de conditions de conservation appliquées sur le starter. Les résultats obtenus sont décrits dans les deuxième et troisième parties de ce chapitre.

## 2. Résultats

### 2.1 Composition microbienne des SCOBY des sites de production de BioBrasseurs

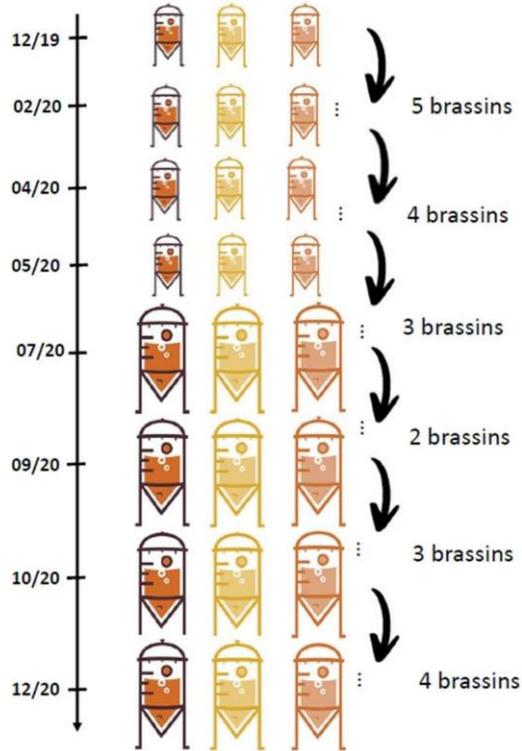
#### 2.1.1 Schéma expérimental et collection des prélèvements

BioBrasseurs pratique une production du kombucha qui met en œuvre jusqu'à huit brassins simultanément en fermentation. Les cuves de fermentation ont des volumes variant de 1000 à 2000 litres. Pour les fermentations successives, la pratique du back-slopping traditionnel est appliquée, en utilisant comme starter le SCOBY de la fermentation précédente, constitué du biofilm d'une épaisseur d'environ 5 cm et pesant entre 10 et 15 kg, additionné de 50 à 100 litres de kombucha non filtré. Ainsi des lignées parallèles de brassins ont été établies, définies par le même biofilm qui est propagé d'un cycle à l'autre, avec du liquide du brassin précédent. De manière pratique, à la fin de chaque cycle de fermentation, par analogie au procédé artisanal, le biofilm est retiré, le kombucha fermenté et transvasé dans une nouvelle cuve pour une filtration grossière, et/ ou une aromatisation afin d'être embouteillé et commercialisé. Environ 5% du volume est gardé dans la cuve pour enchaîner les cycles de back-slopping en ajoutant une nouvelle infusion de thé sucrée. Le biofilm est replacé dans la cuve et un nouveau cycle est lancé. La rotation des SCOBY se fait de cette façon dans tous les brassins. Au site de production, les biofilms mère et fille restent entremêlés, aussi, la masse du biofilm augmente au cours des fermentations successives. Lorsque la masse a doublé, le biofilm est divisé en deux biofilms considérés comme apparentés, qui peuvent être mis en fermentations parallèles. Nous avons choisi de suivre trois lignées parallèles de brassins. Les prélèvements ont débuté en décembre 2019 et se sont poursuivis jusqu'à l'été 2021. Nous avons mis en place une routine de prélèvement des rétentats de filtration, sur le site de

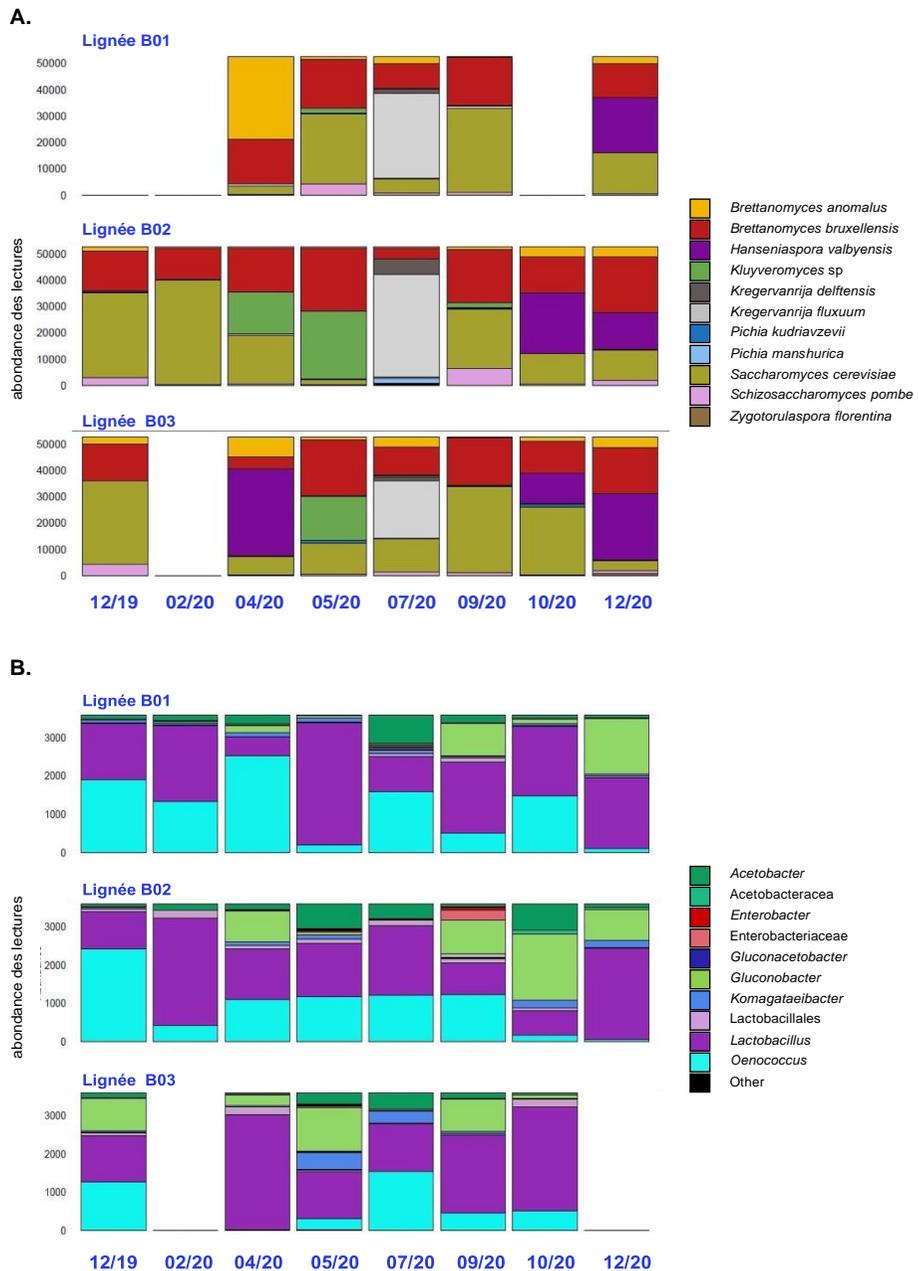
production, de manière à limiter au maximum son impact sur la charge de travail des opérateurs. Ainsi, des bocaux en verre d'une contenance de 400 ml ont été mis à disposition dans l'atelier de production près du dispositif de filtration. Ces bocaux sont étiquetés à l'avance en mentionnant la date, le nom du brassin et le SCOBY qui a servi pour l'inoculer. Le stockage des prélèvements s'est fait dans la pièce à température dirigée. Lors de ma mission hebdomadaire chez BioBrasseurs, les bocaux sont récupérés, la conformité de l'étiquetage des bocaux aux fiches de suivi de production est vérifiée puis un transfert dans des tubes de 50 ml est effectué afin de faciliter le transport au laboratoire. Une fois au laboratoire, les culots cellulaires sont récupérés par une centrifugation et conservés à -20°C. Les données relatives aux lignées, les dates, les brassins et les SCOBY impliqués sont archivées dans le fichier de suivi de la rotation des SCOBY.

L'année où les prélèvements ont été effectués est particulièrement intéressante car elle a été marquée par deux événements majeurs : un arrêt temporaire dû au confinement lors de la crise sanitaire (mars 2020) et un changement du local de production qui a fait évoluer l'outil de travail (juillet 2020). Les cuves ont doublé de volume de 1000 à 2000 litres. Les conteneurs de stockage en polyéthylène haute densité compatibles avec l'industrie alimentaire ont été remplacés par des cuves en acier inoxydable où la température est régulée. Sur une durée d'une année, nous avons ainsi pu récolter un total de 87 échantillons pour trois lignées de production.

Pour l'analyse des compositions microbiennes, 24 échantillons ont été choisis, représentant huit temps de prélèvement tout au long d'une année sur les trois lignées (Figure 3.1).



**Figure 3.1.** Suivi de la composition des lignées de production chez BioBrasseurs. Les prélèvements ont été réalisés sur le rétentat de filtration de 87 brassins appartenant à trois lignées de fermentation parallèles, représentées par des couleurs différentes. Seuls les 24 brassins pour lesquels la composition microbienne a été déterminée sont représentés. Les dates de prélèvement sont indiquées.



**Figure 3.2.** Composition microbienne des rétentats de trois lignées de brassins, suivies le long d’une année, désignées B01 à B03. Les prélèvements d’une même lignée sont disposés sur la même ligne et les prélèvements effectués à la même date sont disposés dans la même colonne. (A) Composition des espèces de levures, représentées par la librairie d’amplicons D2. Le nombre de lecture a été normalisé selon l’échantillon présentant la plus faible profondeur de séquençage (52485 lectures). (B) Composition des genres de bactéries, représentés par la librairie d’amplicons V3-V4. Le nombre de lecture a été normalisé selon l’échantillon présentant la plus faible profondeur de séquençage (3580 lectures).

### 2.1.2 Résultats des compositions microbiennes

L'ADN total a été extrait des cellules récoltées et les bibliothèques d'amplicons D2 et V3-V4 ont été synthétisées et séquencées. Après séquençage, nous avons obtenu des profils taxonomiques pour les bibliothèques D2 de 20 échantillons et pour les bibliothèques V3-V4 de 22 échantillons, représentés dans la (Figure 3.2).

Dans les compositions microbiennes de levures (Figure 3.2A), nous observons que l'espèce de levure *Brettanomyces bruxellensis*, du core défini au chapitre 2, est présente dans tous les échantillons mais avec des abondances différentes variant de 30 à 50 %. Dans des conditions industrielles, nous observons également la présence, en proportions parfois importantes, d'autres genres ou espèces qui ont été considérés comme accessoires dans notre étude du chapitre 2 : *Saccharomyces cerevisiae* (10 à 80 %), *Hanseniaspora valbyensis* (20 à 70 %), *Kregervanrija pseudodelftensis* (40 à 75 %) et *K. fluxuum* (5 à 10 %), ainsi que le genre *Kluyveromyces* (7 à 45 %) pour lequel l'identification des OTU à l'échelle espèce n'a pas pu être effectuée car les espèces *K. lactis* et *K. marxianus* ont leur séquence du segment D2 identique.

En examinant l'évolution des profils au cours des prélèvements, lors du premier point de prélèvement, on remarque que les deux lignées disponibles montrent des profils très semblables composés essentiellement de *S. cerevisiae* (60 %), *B. bruxellensis* (30 %) ainsi que de *B. anomalus* et *S. pombe* en moindre abondance. Dans le deuxième prélèvement, le seul profil disponible ressemble aux précédents avec une diminution de l'abondance de *B. anomalus* et *S. pombe*. Le premier prélèvement du printemps, qui correspond à la reprise de l'activité, suite au confinement, présente des profils très différents des précédents et différents sur les trois lignées. Les trois profils sont caractérisés par une diminution importante de *S. cerevisiae* et cette diminution se fait en faveur d'un taxon différent pour les 3 lignées : l'espèce *B. anomalus* pour la lignée B01, une espèce du genre

*Kluyveromyces* dans la lignée B02, et l'espèce *Hanseniaspora valbiensis* détectée au niveau de la lignée B03. Ces trois taxons représentent 55, 25 et 70 %, respectivement. La présence prépondérante du genre *Kluyveromyces* est confirmée dans les prélèvements ultérieurs, dans les trois lignées, bien qu'en proportion moindre pour la première lignée. Ce genre n'avait été détecté dans nos études précédentes que dans un seul échantillon et avec une proportion très faible (0.14 % dans le profil de l'échantillon FZ3, chapitre 2). Pour la suite des prélèvements, ce genre redevient très minoritaire voire non détecté. Un autre changement drastique dans les profils est observé lors du prélèvement de juillet 2020, avec la prédominance dans les trois lignées, de l'espèce *K. pseudodelftensis* (40 à 75 %) et de l'espèce *K. fluxuum* (3 à 10 %). Les prélèvements qui présentent ces compositions inhabituelles, correspondent aux premiers brassins effectués dans les nouvelles cuves de fermentation, dans les nouveaux locaux de production. Ces deux espèces de Pichiaceae ont couramment été détectées dans les SCOBY de kombucha que nous avons analysés mais avec des proportions plutôt faibles, inférieures à 1%, sauf pour quelques échantillons présentant des proportions de *K. fluxuum* allant jusqu'à 28 % (chapitre 2, échantillons BK1 et RI2). Après deux mois de back-slopping, ces deux espèces ne sont plus détectées dans les profils de septembre 2020. Le dernier changement notable est observé pour les deux derniers prélèvements analysés : toutes les lignées montrent des compositions très similaires avec une abondance importante de l'espèce *H. valbyensis* (25 à 50 %). Nous n'avons pas analysé systématiquement les compositions des prélèvements ultérieurs, sauf pour un prélèvement effectué en juillet 2021, utilisé pour initier les lignées de laboratoire et les tests de conservation (voir paragraphes suivants). Ce prélèvement présente exactement les mêmes compositions que ceux des trois prélèvements de décembre 2020.

D'une façon générale, dans ces échantillons qui peuvent être considérés comme des suites temporelles, il est possible d'établir des liens évidents entre les

variations de compositions observées et le moment des prélèvements. En effet, toutes les lignées présentent le même type de changement majeur au même moment, ce qui est en faveur d'une réponse à un facteur externe comme des conditions saisonnières ou de l'environnement de travail.

Dans les compositions microbiennes de bactéries (Figure 3.2.B), nous observons que les profils taxonomiques des prélèvements sont très variables aussi bien entre les lignées qu'en fonction des cycles. Contrairement aux profils des levures, aucun point commun ne peut être mis en évidence pour expliquer ces fluctuations qui semblent être aléatoires.

Dans les rétentats de production, les genres majoritaires sont les bactéries acétiques *Gluconobacter* (10 à 55 %) et les bactéries lactiques *Lactobacillus* (25 à 80 %) et *Oenococcus* (5 à 70 %). Ces genres font partie de ceux décrits comme accessoires au chapitre 2 et constituent donc une caractéristique des SCOBY de BioBrasseurs. Les genres du core, *Komagataeibacter* et *Acetobacter*, n'ont été observés ici, qu'en faible proportion (1 à 20 % et 2 à 25 %).

Des entérobactéries ont été détectées à 10% d'abondance pour le prélèvement de septembre 2020 de la lignée B02. Seul un non-respect des bonnes pratiques d'hygiène peut justifier une telle contamination. Des analyses par un laboratoire accrédité sont systématiquement effectuées sur les produits finis de BioBrasseurs, visant la recherche des pathogènes alimentaires. Comme aucun signalement de pathogène n'a été rapporté, nous supposons que soit le genre d'entérobactéries détecté n'est pas considéré comme un pathogène, soit il n'a pas survécu aux conditions acides, légèrement alcoolisées du kombucha. Cette contamination était temporaire et unique à la deuxième lignée.

### 2.1.3. Discussion

Les prélèvements étudiés ont montré de grandes différences dans leur composition microbienne. Il faut souligner que la composition microbienne d'un

rétenant de filtration est représentative de la fraction liquide de la fermentation. Or nous avons décrit précédemment que la distribution des genres bactériens *Komagataeibacter*, *Lactobacillus* et *Oenococcus* est différente dans la fraction biofilm et dans la fraction liquide (chapitre 2). De même, certains genres de levure, comme le genre *Pichia*, sont particulièrement abondants dans le biofilm-mère des fermentations de laboratoire. Nous ne savons pas dans quelle mesure la composition du biofilm fluctue elle aussi, au cours des cycles et s'il peut servir de réservoir à certains taxons qui peuvent devenir majoritaires dans la phase liquide au cours de certains cycles, lorsque les conditions les favorisent. Au vu de la taille et de la structure très hétérogène du biofilm en cuve de production, avec les biofilms mère et fille imbriqués, il nous semble néanmoins difficile d'envisager un suivi de sa composition.

Il est intéressant de remarquer que les compositions des SCOBY des différentes lignées peuvent présenter des différences importantes alors que chez BioBrasseurs, les SCOBY de toutes les lignées de production sont issues du même SCOBY initial, mis en fermentation en 2016. Entre ce SCOBY initial et les SCOBY actuels, il y a eu de nombreux back-slopping ainsi que des divisions du biofilm lorsque celui-ci a pris trop de poids. Ce n'est donc pas la composition initiale du SCOBY initial qui conditionne les compositions des lignées mais les nombreuses étapes effectuées et des événements de production qui selon l'événement, peuvent être soit spécifiques d'un brassin ou d'une lignée, soit influencent tous les brassins parallèles. Par exemple, pour la présence des levures du genre *Kluyveromyces*, au printemps 2020, sur l'ancien site de production, il est possible de suspecter une colonisation, temporaire et saisonnière, par ce taxon qui pourrait venir de l'environnement du site de production. La configuration des installations de l'outil de production, qui impliquait l'usage d'opérations manuelles complexes pour la récupération du biofilm et le pompage de la fraction liquide, seraient en faveur d'une colonisation de l'extérieur. La géométrie des

nouvelles cuves et du système de pompage a simplifié ces opérations manuelles. A l'inverse, en juillet 2020, lors du déménagement et au doublement du volume des cuves, des opérations de rassemblement de biofilms apparentés ont été effectuées afin de satisfaire les conditions de quantité de starter. Cette opération aurait pu causer l'expansion temporaire des espèces du genre *Kregervanrija* qui est habituel dans les niches écologiques fermentaire du vin, du cidre et du kombucha (Coton et al., 2017 ; Kurtzman, 2011). Une analyse des autres prélèvements disponibles, effectués en 2021 et 2022 permettra de confirmer si, dans les installations du nouveau site de production, les compositions des lignées de production sont plus stables et homogènes, ou si des fluctuations persistent. Dans ce cas, l'observation de ces fluctuations sur de plus longues périodes permettra peut-être de dégager un profil de ces fluctuations, en particulier pour les compositions en bactéries, afin de pouvoir les associer à des événements saisonniers ou des changements dans le protocole de production. Par ailleurs, ces fluctuations de compositions microbiennes ne semblent pas avoir affecté les propriétés organoleptiques de la boisson commercialisée. Il serait néanmoins intéressant de connaître les conséquences de ces variations sur la composition biochimique du kombucha, en particulier celle des composants aromatiques.

## 2.2 Effet des back-slopping successifs à l'échelle laboratoire

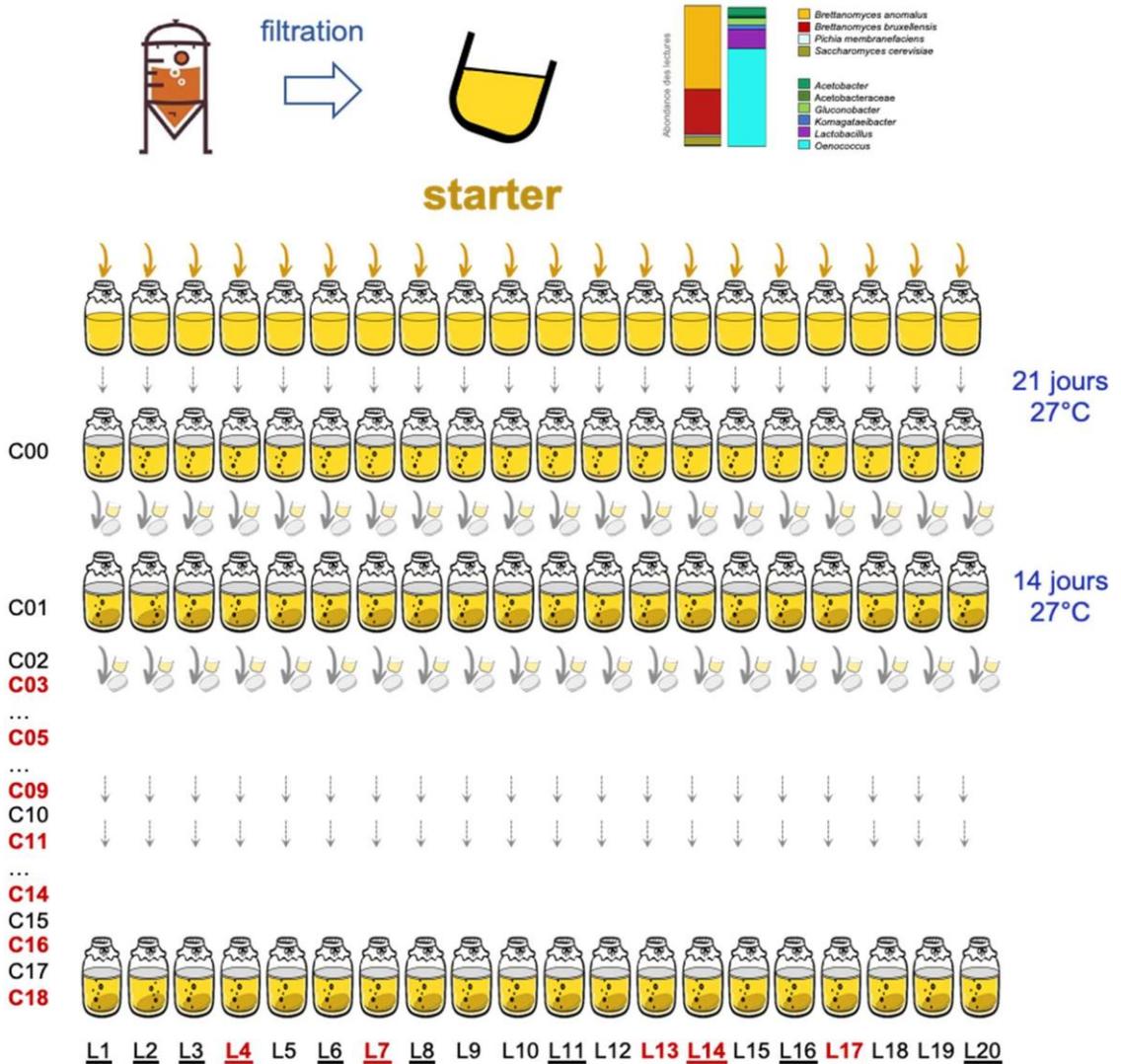
A l'échelle du laboratoire, pour comparer les compositions microbiennes des cycles successifs de fermentation, nous avons reproduit des lignées de fermentation parallèles avec un schéma de propagation des SCOBY équivalent à celui des propagations des lignées sur le site de production de BioBrasseurs (Figures 3.3 et 3.6). En plus de la différence d'échelle, les différences avec les échantillons des lignées de production sont les suivantes :

- nous avons choisi d'effectuer les back-slopping non seulement en conditions d'hygiène mais aussi en conditions de stérilité, en travaillant près de la flamme d'un bec Bunsen. L'utilisation d'un poste de sécurité microbiologique aurait été idéale, mais les bocaux de 1.7 litres sont trop encombrants pour garantir la stérilité des manipulations dans ces conditions. L'objectif est de limiter au maximum la contamination microbienne afin d'évaluer les seuls effets de dérives éventuelles des compositions microbiennes.
- une autre différence est le nombre plus important de lignées suivies : 20 et 11 lignées de fermentations parallèles ont été suivies, pour les deux essais effectués,
- une dernière différence se situe dans le fait que, au laboratoire, les lignées d'un même essai sont en fait des réplicats biologiques. En effet, comme illustré par le design expérimental, un même starter est utilisé pour initier toutes les lignées d'un même essai, ce starter est constitué par l'un des rétentats de filtration, prélevé sur les lignées de production chez BioBrasseurs.

## 2.2.1 Premier essai : suivi de 20 lignées de fermentation parallèles

### 2.2.1.1 Description du dispositif expérimental et schéma de prélèvement

Le starter à l'origine des vingt lignées est un rétentat de filtration de BioBrasseurs qui a été prélevé en avril 2020, juste après la reprise des activités suite au confinement général lors de la crise sanitaire de la Covid19. Sa composition microbienne a été établie à partir des librairies D2 et V3-V4 (Figure 3.3). A partir de cette suspension microbienne, nous avons élaboré des lignées identiques, comme illustré par la Figure 3.3. Un litre d'une préparation d'infusion sucrée et refroidie a été versé dans 20 bocaux traités conformément aux pratiques détaillées (chapitre 1). Le SCOBY starter est homogénéisé et réparti dans les 20 bocaux, à raison de 50 ml par bocal (5 % du volume total), dans les conditions de stérilité représentées par le travail près de la flamme d'un bec Bunsen.



**Figure 3.3.** Design expérimental pour élaborer 20 lignées de fermentation parallèles (désignées L1 à L20) à partir du même starter, à l'échelle laboratoire. Le starter est un rétentat de filtration issu de BioBrasseurs, la composition de ses librairies de levures et de bactéries est représentée. 18 cycles (C01 à C18) ont été effectués. Les cycles pour lesquels des prélèvements systématiques des cellules de la fraction liquide ont été effectués sont signalés en rouge, les lignées pour lesquelles la composition microbienne a été analysée par metabarcoding, sont signalées en rouge. Les lignées qui ont conservé un biofilm d'épaisseur normale après le dixième back-slopping sont soulignées.

Les bocaux sont incubés pendant 21 jours à 27°C (cycle C00), à l'issue desquels un biofilm s'est formé. A la fin de ce cycle, un inoculum constitué du liquide et du biofilm formé est effectué pour le premier back-slopping initiant le cycle de fermentation C01. La durée d'incubation à partir du cycle C01 est de 14 jours à 27°C. Afin de ne pas induire un facteur de variation externe pouvant affecter la communauté, nous avons choisi de ne prélever que de la fraction liquide de toutes ces lignées. En effet, nous avons considéré qu'il est préférable de ne pas découper les biofilms filles pour en analyser la composition microbienne. Les cellules de la fraction liquide ont été prélevées pour toutes les lignées, au long des sept cycles : C03, C05, C09, C11, C14, C16 et C18, et nous avons collecté 140 culots cellulaires stockés à -20°C.

Au fur et à mesure des back-slopping, nous avons également observé l'aspect visuel des biofilms.



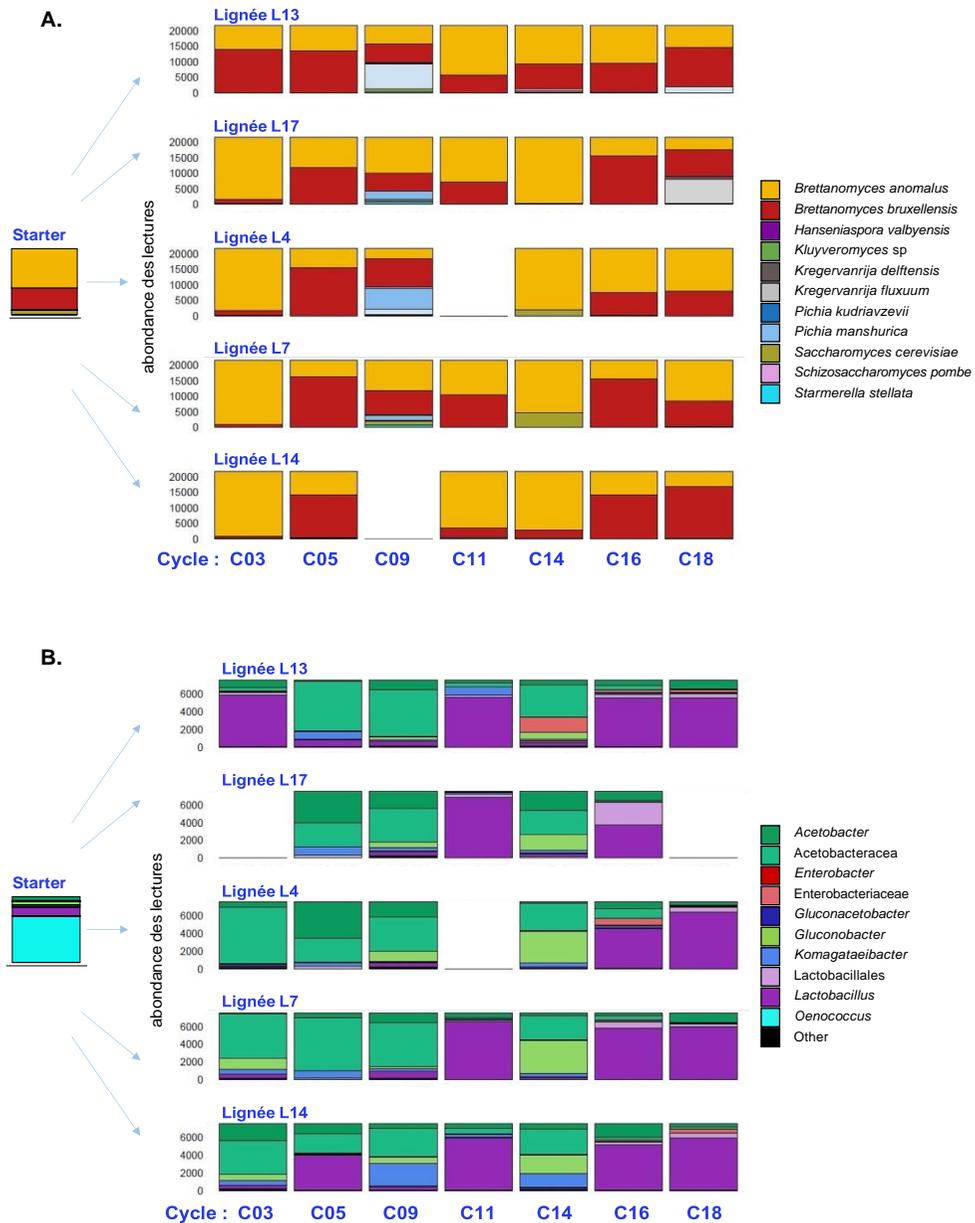
**Figure 3.4.** Observation des épaisseurs des biofilms-filles formés en surface, lors du cycle suivant le dixième back-slopping : à droite, la formation d'un biofilm d'épaisseur habituelle (0.5 cm, flèche large) ; à gauche, la formation d'un biofilm plus fin (flèche fine) dans les lignées (L5, L9, L10, L12, L13, L15, L17, L18, L19).

A la fin du dixième cycle, nous avons constaté des différences les 20 lignées : cette différence concerne l'épaisseur du biofilm fille et a permis de distinguer deux groupes de lignées : un groupe avec un biofilm d'épaisseur habituelle (0.5 cm) et un groupe ayant un biofilm plus fin (Figure 3.4). Cette différence d'épaisseur s'est maintenue sur les six cycles suivants, puis nous avons stoppé la propagation des 20 lignées à la fin du 18e cycle car tous les biofilms-filles étaient de plus en plus fins et irréguliers malgré l'augmentation de la durée d'incubation de 14 à 21 jours. Pour l'analyse des compositions microbiennes, nous avons choisi de comparer tous les prélèvements de deux lignées ayant un biofilm fille fin à partir du 10e cycle (lignées L13 et L17) avec ceux des trois lignées ayant biofilm fille habituel au 10e cycle (lignées L4, L7, L14). Pour ces prélèvements, l'ADN total a été extrait et les bibliothèques d'amplicons D2 et V3-V4 ont été synthétisées et séquencées.

#### 2.2.1.2 Résultats des compositions microbiennes

Le starter utilisé pour générer les 20 lignées présente la composition microbienne de levures suivante : 58 % des amplicons de la région D2 ont été affiliés à *B. anomalus*, 30 % à *B. bruxellensis*, 10 % à *S. cerevisiae* et 2 % à *P. membranifaciens*. Pour les bactéries, 60 % des amplicons V3-V4 sont affiliés à *Oenococcus*, 20 % à *Lactobacillus*, 10 % à *Acetobacter*, 7 % à *Gluconobacter* et 3 % à *Komagataeibacter* (Figures 3.3 et 3.5). Ce sont tous des taxons couramment trouvés dans les SCOBY de kombucha mais l'abondance du genre *Oenococcus* est particulièrement importante.

A partir de ce starter, les compositions en levures sont très semblables au cycle C03, pour les cinq lignées analysées (Figure 3.5.A). Ces compositions du cycle C03 sont marquées par l'absence de *S. cerevisiae* et *P. membranifaciens* et par l'abondance de *B. anomalus*, largement majoritaire (96 %) par rapport à *B. bruxellensis*, sauf pour la lignée L13 qui présente 60 % de *B. bruxellensis*.



**Figure 3.5.** Composition microbienne de la fraction liquide de 5 lignées de laboratoire parallèles, propagées par back-slopping, suivies au cours de 18 cycles de fermentation consécutifs. Les prélèvements d’une même lignée sont disposés sur la même ligne et les prélèvements effectués au même cycle sont disposés dans la même colonne. (A) Composition des espèces de levures, représentées par la librairie d’amplicons D2. Le nombre de lecture a été normalisé selon l’échantillon présentant la plus faible profondeur de séquençage (21673 lectures). (B) Composition des genres de bactéries, représentés par la librairie d’amplicons V3-V4. Le nombre de lecture a été normalisé selon l’échantillon présentant la plus faible profondeur de séquençage (7540 lectures).

Il semble donc que dès le cycle C03, des légères variations de composition en levures se sont établies entre les 20 lignées. Au cours des cycles successifs, la faible abondance de *S. cerevisiae* se maintient et une partie des variations observées concerne des fluctuations du rapport entre les abondances de *B. bruxellensis* et *B. anomalus*. Ces fluctuations passent par le cycle C05 où *B. bruxellensis* est devenue majoritaire, suivie plus tard par deux cycles où *B. anomalus* est devenue progressivement très majoritaire (C11 et C14). Cette description est globale avec des variations entre les lignées. Enfin, aux cycles C16 et C18, *B. bruxellensis* est redevenue majoritaire. Le passage du cycle C05, avec *B. bruxellensis* majoritaire, au cycle C11 avec *B. anomalus* majoritaire, est marqué par un changement observé sur tous les prélèvements disponibles du cycle C09 : le genre *Pichia*, qui n'était pas détecté précédemment, est soudain devenu abondant, avec des proportions de 10 à 35 %. Cette colonisation est passagère puisque ce genre n'est plus détecté dans les cycles suivants.

Toutes les lignées sont affectées au même cycle, il donc est possible d'émettre l'hypothèse que la fluctuation transitoire du rapport entre les abondances de *B. bruxellensis* et *B. anomalus*, observée aux prélèvements ultérieurs des cycles C11 et C14, soit une conséquence de cette colonisation transitoire. La connaissance des compositions des cycles C08 et C10 auraient peut-être permis d'apporter du support à cette hypothèse, en précisant à quel cycle la colonisation de *Pichia* a eu lieu et quelle est la synchronie avec la fluctuation du rapport entre les abondances de *B. bruxellensis* et *B. anomalus*. La production, par les espèces du genre *Pichia*, de toxines killer actives sur plusieurs espèces de levure, dont *B. bruxellensis*, serait une piste à explorer (Belda et al., 2017; Santos et al., 2009).

Enfin, il semble qu'il n'y ait aucun lien apparent entre les profils de variations des compositions ni en levures, ni en bactéries (qui sont détaillés ci-dessous) et les changements d'épaisseur du biofilm pour les lignées L13 et L17.

Les variations de compositions microbiennes sont très marquées pour les compositions bactériennes (Figure 3.5.B) : elles sont toutes différentes dès le cycle C03 des quatre lignées disponibles, celle de la lignée L13 étant la plus différente du fait de la prédominance du genre *Lactobacillus*. Dans les profils de variations, une observation remarquable concerne les proportions du genre *Lactobacillus*. Toutes les compositions du cycle C11 sont marquées par une augmentation soudaine de ce genre qui est prédominant et atteint des proportions de 75 % à 90 %. Pour toutes les compositions du cycle C14, au contraire, ce genre est très minoritaire (8 %), pour atteindre et conserver des proportions majoritaires dans toutes les compositions des derniers cycles (75 à 85 %). Contrairement aux variations des compositions en bactéries sur le site de production chez BioBrasseurs, un profil de variation semble plus évident à déterminer dans les lignées de laboratoire. L'observation d'une certaine synchronie entre les fluctuations observées pour les compositions de levures et de bactéries, entre les cycles C09, C11 et C14, soulève la question de l'effet d'un événement externe qui aurait non seulement affecté les compositions de levures mais aussi celles de bactéries, voire d'une interdépendance entre les deux types de composition. Par ailleurs, la présence d'entérobactéries détectables pour les trois lignées, à différents cycles, semble indiquer plusieurs contaminations indépendantes et suggère que les conditions de travail stérile soient difficiles à satisfaire avec l'encombrement des bouches et l'usage du bec Bunsen,

Une observation générale concerne le genre *Oenococcus* : très abondant dans le starter (60 %), il n'est plus du tout détecté dans nos prélèvements, il semble que les conditions des lignées de laboratoire ne soient pas favorables à son maintien dans la communauté microbienne. Une autre observation remarquable est l'abondance importante des OTU de la famille des Acetobacteraceae (10 à 80 %), dont l'affiliation à l'échelle genre n'est pas possible avec les données des amplicons V3-V4.

Évidemment, cette observation souligne les limites de ce marqueur taxonomique pour ces taxons, de par son profil de polymorphisme ou/et de par l'état des connaissances sur les taxons de cette famille, répertoriées dans la banque d'affiliation. Néanmoins une telle abondance de ces OTU n'avait pas encore été observée dans nos analyses. Cette famille contient des genres que nous avons définis comme genres du core microbien : *Acetobacter* et *Komagataeibacter*, ainsi que des genres que nous avons définis comme accessoires, tels *Gluconobacter*, *Gluconacetobacter*, *Ameyamaea*, *Neokomagataea*, *Nguyenibacter*, *Swingsia*. Une analyse plus exhaustive des compositions microbiennes, basées sur le séquençage shotgun du contenu génomique de la communauté permettrait peut-être d'améliorer la résolution des taxons de cette famille, non seulement pour nos projets dans la fermentation de kombucha mais aussi pour l'amélioration du contenu des banques.

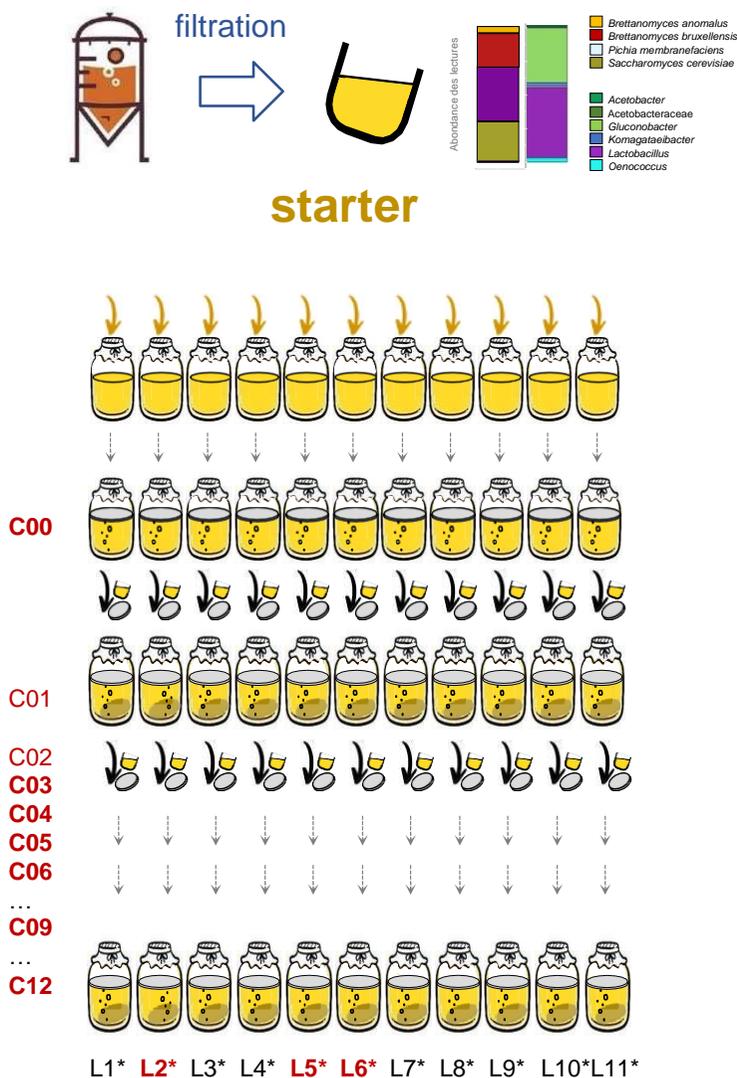
## 2.2.2 Deuxième essai : suivi de 11 lignées de fermentation parallèles

### 2.2.2.1 Description du dispositif expérimental et schéma de prélèvement

Le dispositif expérimental est très similaire à celui du premier essai et il est résumé dans la (Figure 3.6) Les différences sont les suivantes :

- un nombre moins important, de 11 lignées parallèles, pour une meilleure homogénéité des opérations. Notamment, il est possible de préparer un seul batch d'infusion sucrée pour effectuer les 11 back-slopping parallèles, alors que deux batches étaient nécessaires pour les 20 lignées. Les lignées de cette nouvelle série ont été désignées L1\* à L11\*.
- nous avons adopté une autre stratégie de prélèvements : en plus des prélèvements sur toutes les lignées à des intervalles de deux à trois cycles, pour trois des onze lignées, nous avons effectué des prélèvements systématiques à chaque cycle,

y compris après le cycle d'initiation C00 correspondant à l'inoculation des bocaux d'infusion avec la suspension microbienne starter. Le but de ces prélèvements rapprochés était de mieux repérer les effets des événements externes sur les compositions microbiennes.

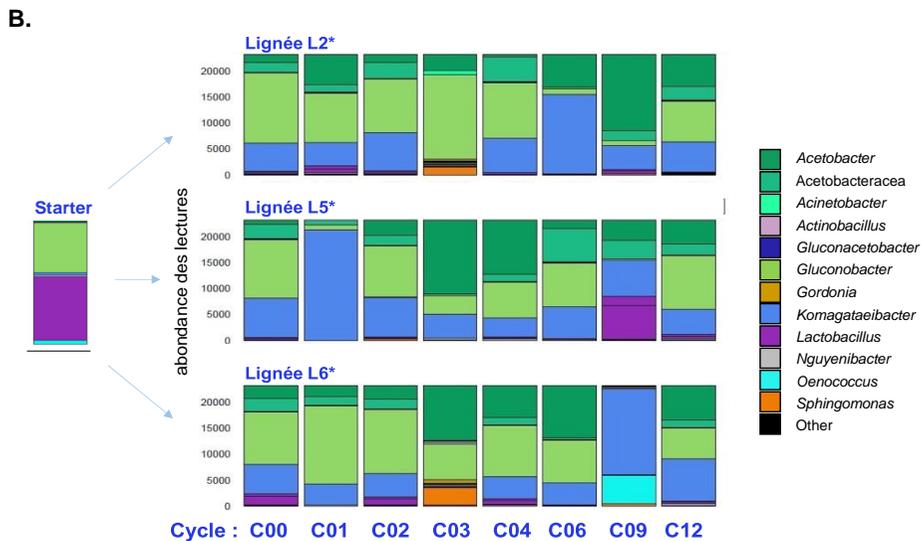
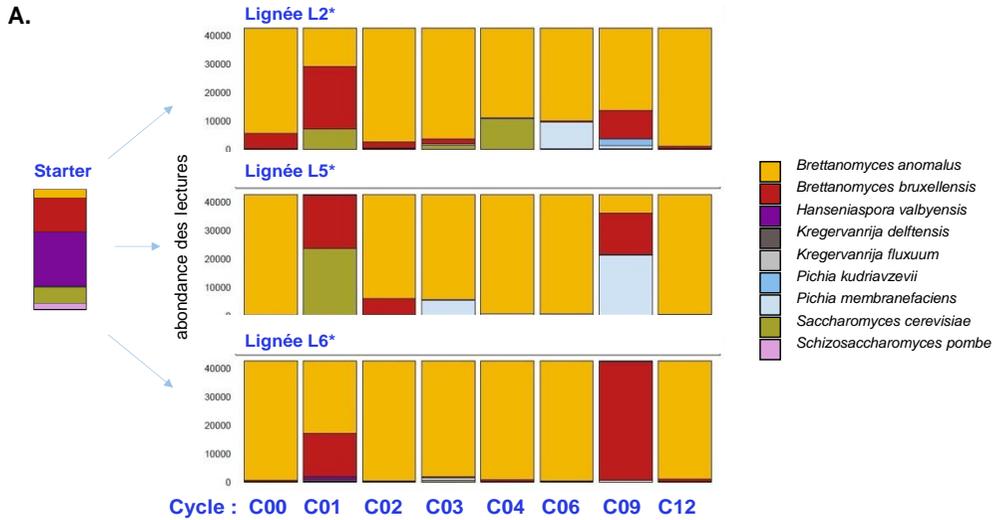


**Figure 3.6.** Design expérimental pour le suivi de la deuxième série des lignées de laboratoires. Les prélèvements ont été systématiques pour les trois lignées L2\*, L5\* et L6\*, sur les cycles consécutifs C00, C01, C02, C03 et C04 et pour les cycles plus espacés C06, C09 et C12.

- le starter de départ à l'origine de ces lignes parallèles identiques a une composition différente, révélé par le séquençage des librairies D2 et V3-V4 effectué sur les cellules du starter. Le profil taxonomique de ce starter, prélevé en production chez BioBrasseurs en juillet 2021, est similaire à celui des lignées en production en décembre 2020. Les abondances des taxons de levures correspondent à 45 % de de séquences de *H. valbyensis*, 25 % de *B. bruxellensis*, 15 % de *S. cerevisiae*, 10 % de *B. anomalus* et 5 % de *S. pombe*. Quant à la composition bactérienne, le profil est composé de 60 % de séquences affiliées au genre *Lactobacillus*, 20 % de *Gluconobacter*, 10 % de *Komagataeibacter* et 10 % d'*Acetobacter* et 0.3 % d'*Oenococcus*. La diversité microbienne de ce starter est donc plus importante que celui utilisé pour les 20 lignées, aussi bien en nombre de taxons que par l'absence d'un taxon prédominant. Les valeurs comparées des indices d'alpha-diversité sont les suivants : le starter des 11 lignées contient 15 taxons de levures dont les proportions donnent un indice de Gini-Simpson de 0.77 et 29 taxons de bactéries dont les proportions donnent un indice de Gini-Simpson de 0.63. Pour le starter des 20 lignées, ces valeurs n'étaient que de 9 taxons de levures dont les proportions donnent un indice de Gini-Simpson de 0.59 et 19 taxons de bactéries dont les proportions donnent un indice de Gini-Simpson de 0.59.

#### 2.2.2.2 Résultats des compositions microbiennes

Que ce soit pour les levures ou les bactéries, il faut tout d'abord remarquer que les compositions des trois lignées au cycle C00 sont très similaires (Figure 3.7). Les légères variations observées peuvent être dues à la précision de la méthode de métabarcoding. Cette observation conforte notre protocole de génération de lignées à partir d'un rétentat de filtration qui est effectivement suffisamment homogène pour permettre de constituer des répliquats biologiques.



**Figure 3.7.** Composition microbienne de la fraction liquide de 3 lignées de laboratoire parallèles, propagées par back-slopping, suivies au cours de 12 cycles consécutifs de fermentation. Les prélèvements d'une même lignée sont disposés sur la même ligne et les prélèvements effectués au même cycle sont disposés dans la même colonne. (A) Composition des espèces de levures, représentées par la librairie d'amplicons D2. Le nombre de lecture a été normalisé selon l'échantillon présentant la plus faible profondeur de séquençage (44980 lectures). (B) Composition des genres de bactéries, représentés par la librairie d'amplicons V3-V4. Le nombre de lecture a été normalisé selon l'échantillon présentant la plus faible profondeur de séquençage (23130 lectures).

Malgré cette homogénéité au cycle C00, au cours des cycles suivants, chaque lignée présente des fluctuations de composition qui lui sont propres, indiquant que la composition initiale n'est pas le seul paramètre à avoir un effet sur les variations d'un cycle à l'autre. Ce schéma de prélèvement montre que, au moins pour les compositions de levures, les variations de compositions peuvent être très rapides, comme illustré par les compositions très particulières du cycle C01 par rapport à celles des cycles précédant et suivant.

Les profils taxonomiques des levures sont caractérisés par une diminution importante de la diversité, pour les 3 lignées suivies et pour tous les cycles. L'espèce *H. valbyensis*, très représentée dans le starter et l'espèce *S. pombe* ne sont pas détectées dans ces lignées de fermentation au laboratoire. A l'exception des cycles C01 et C09, l'espèce *B. anomalus* est largement prédominante dans certaines lignées avec des proportions entre 20 et 70 %, contrairement au starter où elle ne représentait que 10 % des séquences.

Les variations de compositions des bactéries sont surtout marquées par une faible abondance du genre *Lactobacillus*, abondant dans le starter. Il faut également remarquer la présence transitoire au cycle C03 de deux lignées, du genre *Sphingomonas*, peu décrit dans les SCOBY de kombucha. Sa présence a été rapportée dans un kombucha local en Slovaquie, qui, de notre point de vue, semble avoir une composition plutôt inhabituelle (Ivanišová et al., 2019). Ce genre était également majoritaire dans une des fermentations conduites avec des starters de différentes origines, décrite dans l'étude du chapitre 2. Ce starter avait été écarté car il n'avait pas abouti à la formation d'un SCOBY utilisable en back-slopping traditionnel.

Enfin, les compositions en levures et en bactéries, de tous les prélèvements du cycle C09, se distinguent des celles des autres cycles et sont variables entre les lignées. Les échantillons disponibles des cycles C07, C08 ainsi que ceux des cycles suivants devront être analysés par métabarcoding, pour compléter ce jeu

De données et éventuellement avoir des éléments permettant de mieux comprendre ces variations.

### 2.2.3 Discussion

Une conclusion importante des résultats des expériences décrites dans cette partie du projet de thèse, est qu'il est possible de générer des fermentations qui constituent des réplicats biologiques de qualité, malgré la nature solide d'une partie de l'inoculum, essentiellement impossible à homogénéiser : il suffit de pouvoir reconstituer cet inoculum à partir d'une suspension homogène des micro-organismes. Une étude de (May et al., 2017) avait décrit un protocole de broyage du biofilm avec un mixer, pour obtenir une suspension homogène qui peut être répartie pour constituer des réplicats biologiques. Cette solution nous semble un peu agressive, sachant que le biofilm est souvent décrit comme une niche écologique qui protège les communautés microbiennes (May et al., 2019). Dans ce contexte, il est possible que les changements importants et ponctuels, que nous avons observés dans les compositions du cycle C01 du suivi des 11 lignées, soient causés par une forme de stress représenté par le premier transfert de type back-slopping.

La possibilité d'obtenir des réplicats biologiques sous la forme de lignées de fermentation est essentielle pour développer des projets visant à comprendre l'assemblage et la stabilité de ces communautés et leurs profils de variation des compositions microbiennes. Il sera en effet possible de comparer l'évolution des lignées constituées à partir de starters de composition différentes et également de tester les effets, sur les compositions microbiennes, de l'application de traitements ou de changements de conditions.

Dans ce contexte, nous avons initié une étude préliminaire dans laquelle les cellules d'un même rétentat de filtration ont été soumises à des conditions

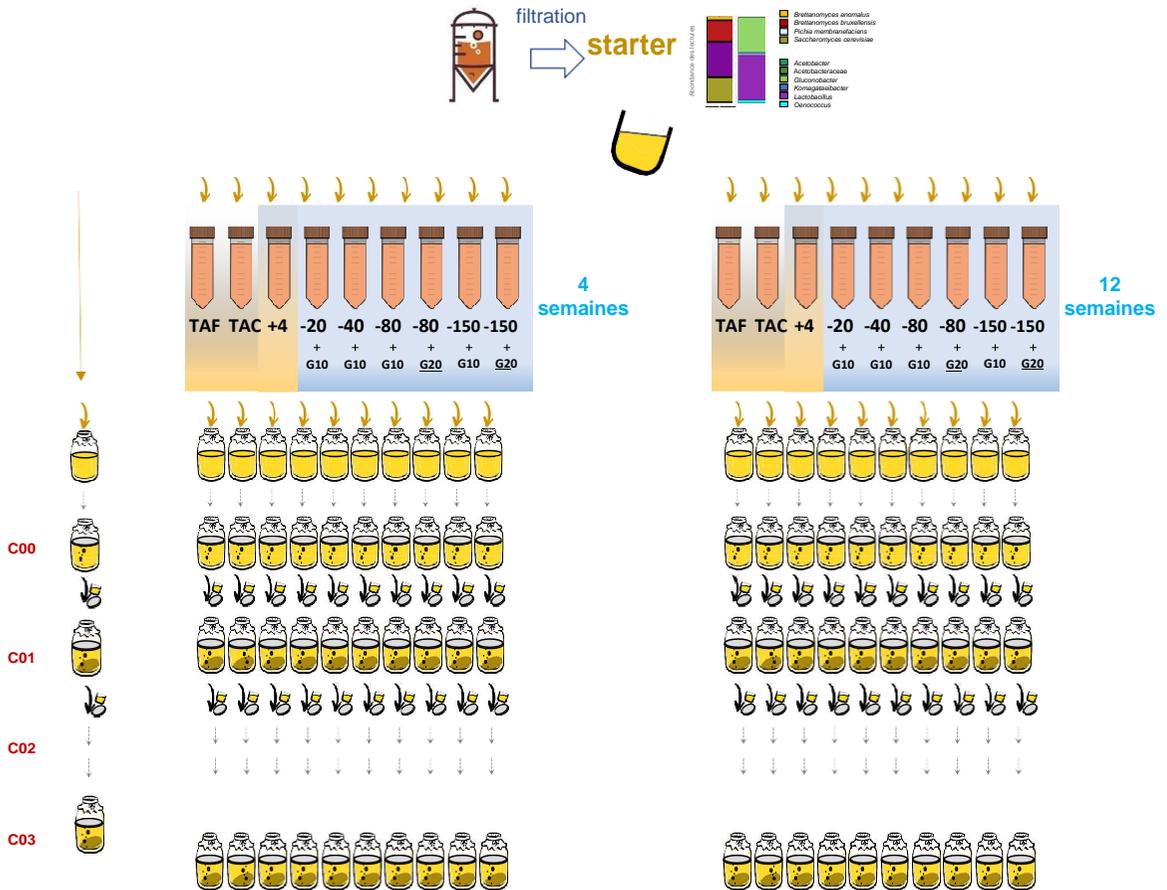
différentes avant d'être utilisées comme starter pour initier les lignées de fermentation.

### 2.3 Effets de différentes conditions de conservation du starter

Les effets de la conservation du starter nous a semblé être la condition la plus pertinente à tester pour initier une étude d'intérêt pour les producteurs de kombucha, amateurs ou brasseurs. En effet, au fur et à mesure des back-slopping, le biofilm mère, qui est retiré en fin de fermentation, peut être conservé et constituer un starter supplémentaire. Les producteurs amateurs peuvent ainsi conserver des SCOBY "de secours" ou les offrir. Chez les brasseurs, en plus de garder un SCOBY de réserve, l'objectif majeur est de générer des SCOBY performants permettant de maintenir ou d'augmenter la capacité de production et d'assurer une fermentation dans les délais habituels. Une pratique adoptée par BioBrasseurs, du moment où la performance du SCOBY diminue, consiste à stocker temporairement le SCOBY au froid avant de le remettre en fermentation. Nous avons donc soumis un même starter à différentes durées et températures de stockage avant d'initier des lignées de fermentation.

#### 2.3.1 Description du dispositif expérimental et prélèvements

Le rétentat de BioBrasseurs qui a servi de starter pour tester l'effet de différentes conditions de conservation est celui utilisé pour la série à 11 lignées de laboratoire, sa composition microbienne a été établie à partir des librairies D2 et V3-V4 (Figure 3.6). La composition après sa mise en culture directement, sans étape de conservation est également connue (Figure 3.7). Pour tester les différentes conditions de fermentation, des aliquotes de 40 ml du starter homogénéisé ont été réparties dans plusieurs tubes falcon de 50 ml et additionnées ou non d'un cryoprotectant : le glycérol (Figure 3.8).



**Figure 3.8.** Design expérimental pour comparer les effets sur les fermentations, de différentes conditions de conservation d'un même starter. Le rétentat de filtration a été soit mis directement en fermentation (colonne de gauche), soit conservé dans différentes conditions de température et de concentration en glycérol, pendant 4 ou 12 semaines, avant d'être mis en fermentation. Le cycle de démarrage C00 a été suivi de trois étapes de back-slopping consécutives.

Conditions de conservation : température ambiante fluctuante (TAF) ; température ambiante contrôlée (TAC) ; 4°C (+4) ; -20°C en présence de glycérol à 10 % v/v (-20 + G10) ; -40°C en présence de glycérol à 10 % v/v (-40 + G10) ; -80°C en présence de glycérol à 10 % v/v (-80 + G10) ; -80°C en présence de glycérol à 20 % v/v (-80 + G20) ; -150°C en présence de glycérol à 10 % v/v (-150 + G10) ; -150°C en présence de glycérol à 20 % v/v (-150 + G20).

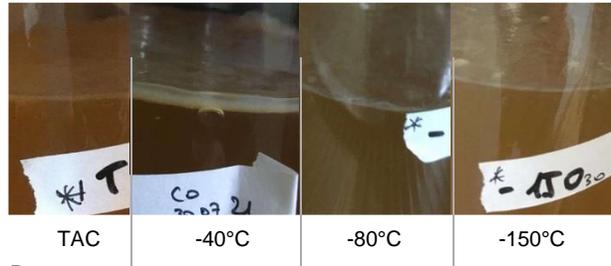
Ces tubes ont ensuite été placés dans différentes conditions de température de stockage : la température ambiante fluctuante dans le laboratoire, la température ambiante stable de 20°C dans une pièce climatisée et la température de 4°C en chambre froide. Différentes températures de congélation et de cryoprotecteur ont

également été testées : -20°C, -40°C, -80°C et -150°C en présence de glycérol 10 % (v/v) et -80°C et -150°C en présence de glycérol 20 % (v/v). Les tubes ont tous été congelés à la même température de -40°C, avant d'être répartis dans les congélateurs aux différentes températures de congélation. Pour chaque condition de conservation, quatre réplicats ont été réalisés et stockés.

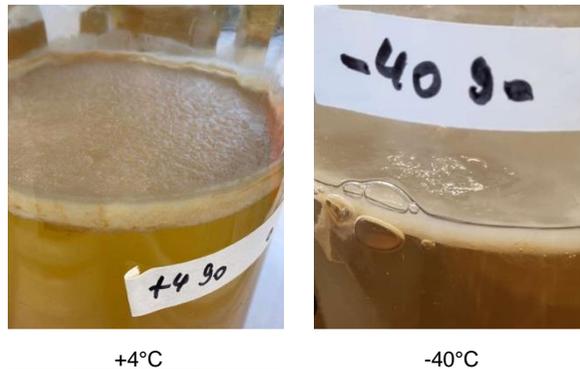
Deux durées de stockage ont pour l'instant été testées : 4 semaines et 12 semaines et il reste encore deux réplicats disponibles de chaque condition, conservés à cette date depuis 16 mois. Au moment d'utiliser ces stocks comme starters, les tubes congelés ont été transférés et placés par paliers de 12 heures dans le congélateur de température inférieure (de -150°C à -80°C, puis de -80°C à -40°C puis de -40°C à -20°C), avant une étape finale de 12 heures à -20°C. Ils ont ensuite été rapidement décongelés en 20 minutes dans un bain-marie à 32°C. Les cellules de ces différentes conditions de stockage ont ensuite été transvasées dans 1.1 litres d'infusion de thé sucrée. Au bout de 21 jours, les back-slopping ont été initiés, pour des durées de 14 jours sur trois cycles consécutifs. Les compositions microbiennes ont été établies sur les prélèvements en fin de cycle, à partir des librairies D2 et V3-V4 (Figures 3.10 et 3.11).

En plus des compositions microbiennes, l'aspect visuel des biofilms fille a été observé quotidiennement pendant leur formation (Figure 3.9). Il a ainsi été possible d'observer que les cinétiques de formation de ces biofilms est différente selon les conditions de conservation appliquées au starter. Par exemple, un biofilm cellulosique d'une épaisseur régulière d'environ 0.2 cm a été produit dès 5 jours de fermentation, dans les bocal correspondants aux starters conservés à -20°C et -40°C : c'est une vitesse de formation plus rapide que ce que l'on observe généralement avec un rétentat de filtration frais. Au même moment, à la surface de tous les autres bocal, une fine pellicule commence à peine à se développer. Ces observations montrent que le redémarrage de la communauté après conservation est différent si les températures de conservation sont différentes.

A la fin du cycle C00, les biofilms avaient également des aspects finaux différents, notamment pour ceux dont les starters ont été conservés à température ambiante : les biofilms filleont plutôt un aspect et une consistance de membrane beige et opaque, flottant à lasurface du liquide et adhérant aux parois du bocal, aspect qui n'est jamais observémême au tout début de formation des biofilms. L'aspect et la consistance habituelsdes biofilms fille correspond plutôt à celui d'un gel translucide et très légèrementbleuté, un peu comme un gel réticulé d'agarose à 1.8 % (w/v). Cet aspect est visible dès le début de sa formation, lorsqu'un biofilm fille se forme par des 'patches' isolés qui s'étendent et se rejoignent au bout de 5 jours. A.



B.



**Figure 3.9.** Aspects caractéristiques observés lors de la formation des biofilms filles. (A) Le retard dans la formation du biofilm à partir du starter conservé à température ambiante contrôlée (TAC) ou ceux congelés à  $-80^{\circ}\text{C}$  et  $-150^{\circ}\text{C}$ , en comparaison de celui conservé à  $-40^{\circ}\text{C}$ , est visible sur série de photos prises cinq jours après l'inoculation des starters conservés pendant 4 semaines. Les conditions de conservation sont indiquées sous les photos. (Le cadrage des photos est assez médiocre car nous avons laissé les bocaux en place dans l'incubateur sans les bouger afin de ne pas perturber la formation des biofilms.) (B) En fin du cycle C00 de démarrage de fermentation, l'aspect inhabituel du biofilm généré à partir d'un starter conservé en chambre froide, en comparaison à l'aspect habituel d'un biofilm fille généré avec un starter conservé à  $-40^{\circ}\text{C}$ . Dans cet exemple, ils'agit des

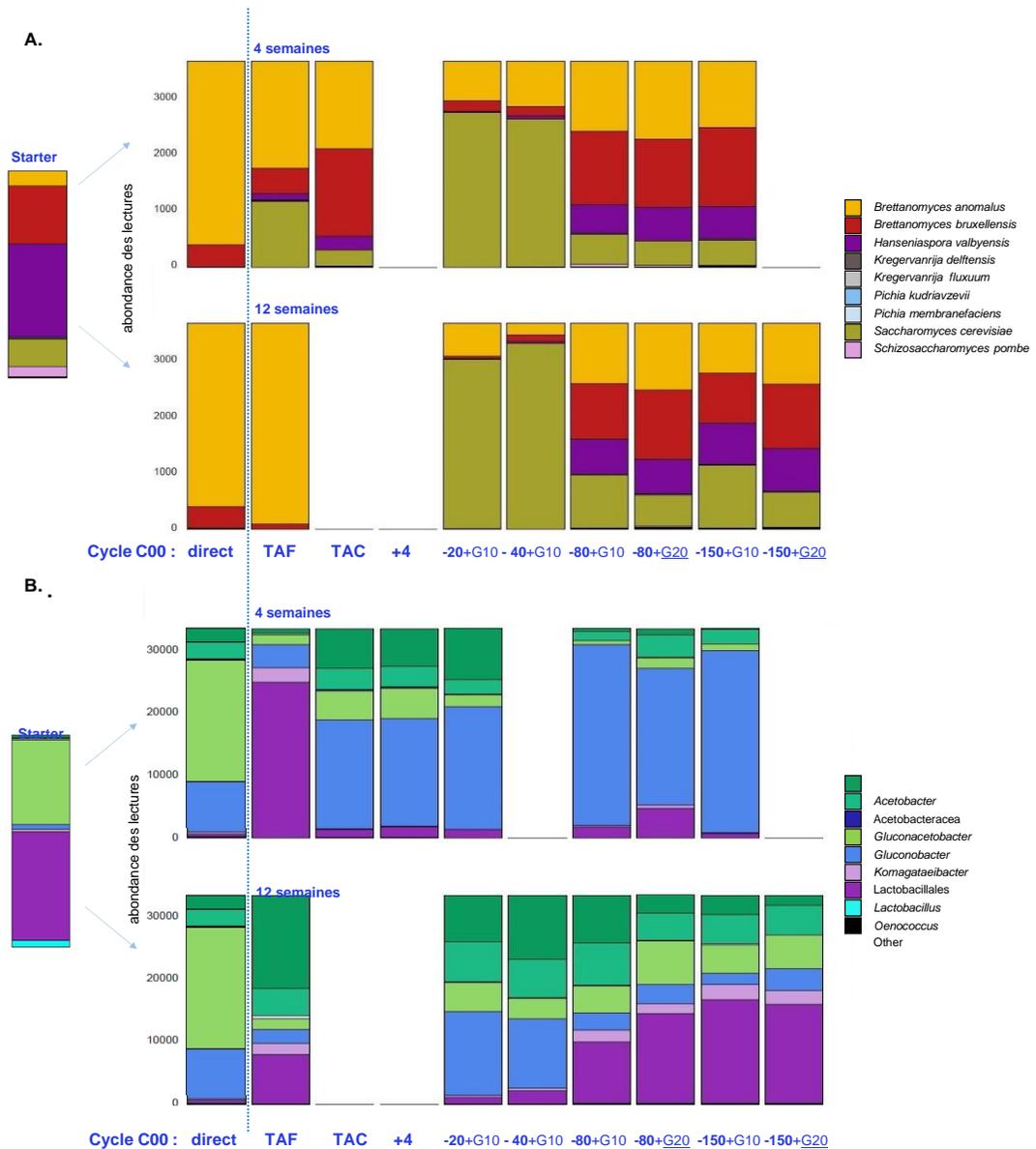
starters conservés pendant 12 semaines.

### 2.3.2 Résultats des compositions microbiennes

Les compositions ont été comparées à celles du starter initial (le rétentat de filtration issu du site de production de BioBrasseurs et récolté en juillet 2021) et celles de ce starter mis en fermentation sans conservation. A la fin du cycle C00 de démarrage, les profils taxonomiques de levures et de bactéries sont tous différents, selon les traitements de conservation appliqués ou non au starter (Figure 3.10). Pour les compositions en levures (Figure 3.10.A), une observation déjà remarquée avec le starter inoculé sans étape de conservation préalable (lors du suivi des 11 lignées de laboratoire) est l'absence des espèces *S. cerevisiae* et *H. valbyensis*, présentes dans le starter initial, ce qui contraste avec les compositions des fermentations issues des starters conservés à  $-80^{\circ}\text{C}$  et  $-150^{\circ}\text{C}$  qui sont très proches de celle du starter initial. Ceci mène à établir un parallèle avec les souchiers des collections de laboratoire qui sont conservés durant des années à la température optimale de  $-80^{\circ}\text{C}$ . Il n'y a pas d'effet visible, ni de la durée de conservation quelle que soit la température de conservation, ni de la concentration en glycérol pour les températures de conservation de  $-80^{\circ}\text{C}$  et  $-150^{\circ}\text{C}$ , sur la composition en levures.

On peut également remarquer la présence particulièrement importante de *S. cerevisiae* pour les fermentations issues de starters conservés à  $-20^{\circ}\text{C}$  et  $-40^{\circ}\text{C}$ , où elle atteint une proportion allant jusqu'à 85 %.

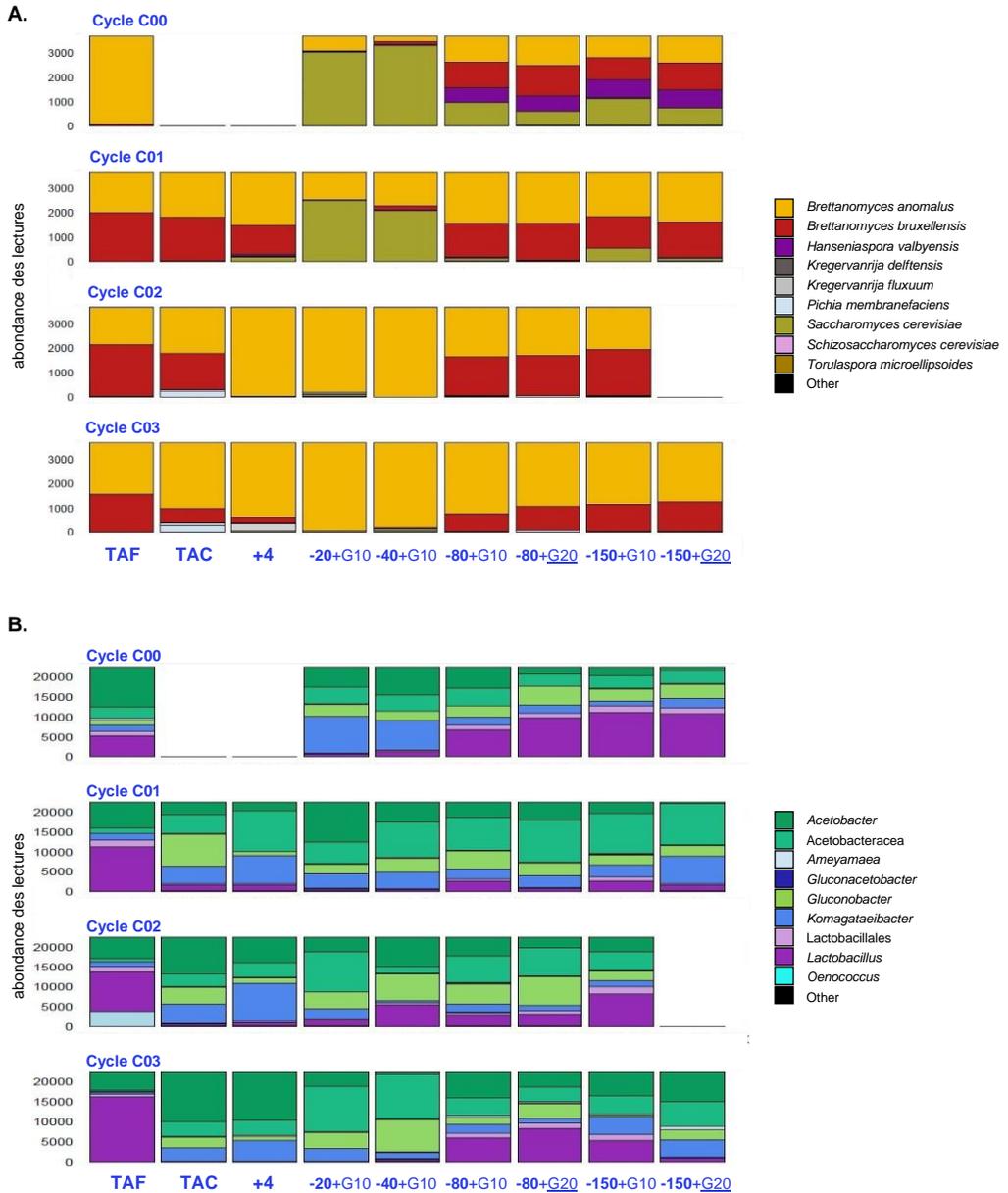
Cette abondance de *S. cerevisiae* et la formation rapide de biofilm (par les bactéries) sont corrélées et il pourra être intéressant d'étudier la relation entre les activités métaboliques des diverses espèces impliquées dans les SCOBY et la vitesse de formation du biofilm



Pour les compositions de bactéries (Figure 3.10.B), un comportement comparable à celui observé pour les compositions en levures peut être décrit : les fermentations issues des starters conservés à -80°C et -150°C sont assez proches de celle du starter initial, avec l'importance du genre *Lactobacillus* notamment ainsi que la proportion du genre *Gluconobacter*. Dans ce test, la durée de conservation des starters à ces températures de congélation semble avoir un effet sur les compositions des fermentations de démarrage.

Afin de connaître les effets éventuels des compositions au cycle C00 sur celle des cycles successifs, nous avons analysé les compositions microbiennes lors de trois back-slopping successifs effectués avec les SCOBY issus des starters conservés durant 12 semaines (Figure 3.11).

Les observations sont à examiner avec prudence puisque nous n'avons pas de réplicats de chaque condition alors que les compositions des lignées de laboratoire générées avec le même starter initial avaient montré des fluctuations dès les premiers cycles de back-slopping (Figure 3.7). Pour les compositions de levures (Figure 3.11.A), les espèces *H. valbyensis* et *S. cerevisiae*, présentes au cycle C00, deviennent indétectables dès les cycles C01 et C02, respectivement. L'espèce *B. anomalus* devient progressivement prédominante dans toutes les compositions du cycle C03. Il semble donc que la composition au cycle C00 de la communauté de levure ait assez peu d'effet à long terme, au fur et à mesure de la succession des cycles de back-slopping. Néanmoins, toutes les compositions des fermentations successives à partir des starters conservés à -20°C et -40°C, sont caractérisées par une très faible abondance de l'espèce *B. bruxellensis*. Il serait intéressant d'étudier les effets des températures de conservation sur différentes souches de cette espèce afin d'établir si cette observation est due à des fluctuations de composition aléatoires ou si elle se conserve particulièrement mal à ces températures de congélation.



**Figure 3.11.** Suivi des compositions microbiennes de la fraction liquide, au cours de trois cycles consécutifs de fermentation avec propagation par back-slopping, pour 9 séries de fermentation qui se distinguent par les conditions de stockage des starters. Les starters ont été conservés pendant 12 semaines. Les prélèvements parallèles du même cycle sont disposés sur la même ligne et ceux d’une série de back-slopping correspondant à la même condition de conservation du starter sont disposés en colonne. (A) Composition des espèces de levures, représentées par la librairie d’amplicons D2. Le nombre de lecture a été normalisé selon l’échantillon présentant la plus faible profondeur de séquençage (3691 lectures). (B) Composition des genres de bactéries, représentés par la librairie d’amplicons V3-V4. Le nombre de lecture a été normalisé selon l’échantillon présentant la plus faible profondeur de séquençage (22492 lectures).

Avec cette série de starters, les différentes compositions en bactéries, au fur et à mesure des back-slopping successifs, définissent deux groupes distincts (Figure 3.11.B), en fonction de l'abondance importante ou non du genre *Lactobacillus*. Ces groupes correspondent aux températures de conservation du starter ainsi qu'à la composition au cycle C00 : les fermentations dont les starters ont été conservés à -80°C et -150°C et à température ambiante fluctuante sont celles où le genre *Lactobacillus* est bien maintenu au fur et à mesure des cycles successifs (abondance des lectures de de 25 à 75 % avec une moyenne de 45 %). Les réplicats constituant les 11 lignées de laboratoire n'avaient pas montré un tel profil de diversité dans les compositions bactériennes, ce qui suggère que ces observations sont effectivement en partie liées aux conditions de conservation des starters et pas seulement à des fluctuations aléatoires de compositions.

### 2.3.3 Discussion

A notre connaissance, une seule étude très récente a testé les effets de la conservation sur les taxons de kombucha, avec un design expérimental très différent (Grassi et al., 2022). Les compositions microbiennes ont été comparées pour des kombuchas mis en bouteille et stockés à 4°C ou à température ambiante, en utilisant des méthodes dépendantes de la culture et des analyses de PCR-DGGE (dont les principes sont rappelés dans l'introduction de la thèse). Les résultats ont mis en évidence une mauvaise conservation de la viabilité de l'espèce de levure *S. cerevisiae* et des bactéries lactiques, alors que l'espèce de levure *B. anomalus* et les bactéries acétiques conservent une très bonne viabilité. Il est difficile de comparer ces résultats aux nôtres, non seulement parce que les méthodes de détection n'ont pas la même résolution, que la remise en fermentation des microorganismes n'a pas été testée mais aussi parce que la composition microbienne de la communauté étudiée était très différente de la nôtre, avec notamment l'absence de l'espèce de levure *B. bruxellensis*.

Notre étude préliminaire a permis de montrer la possibilité d'utiliser notre design expérimental, basé sur des lignées de fermentation parallèles, pour mettre en évidence l'impact de l'application volontaire de différentes conditions sur le déroulement de la fermentation. Cet impact était visible macroscopiquement, avec les différentes cinétiques de formation des biofilms fille. L'analyse des compositions microbiennes a également révélé un impact de ces conditions mais il est encore assez difficile de faire le lien entre ces différences de compositions et ces différences de déroulement de la fermentation. Cela est dû au fait que l'abondance des taxons en fin de fermentation ne reflète que partiellement leur activité au cours du cycle complet. Les résultats disponibles devront être complétés pour une expérience à plus grande échelle, incluant des cinétiques des compositions microbienne et biochimique, notamment au cours du cycle C00 de démarrage. Techniquement, ce type d'expérience ne pourra être conduite qu'avec peu de conditions mais de nombreux réplicats. Ces réplicats seront nécessaires, non seulement pour vérifier la reproductibilité des résultats, mais aussi parce lorsqu'un prélèvement est effectué, la fermentation en cours est perturbée et ne peut donc pas être poursuivie pour des prélèvements ultérieurs. Il faudra donc prévoir autant de fermentations parallèles que de points de prélèvements. Des données de méta-transcriptomique seraient également intéressantes à générer afin de suivre l'activité des taxons par leurs cinétiques d'expression.

Pour mieux répondre au questionnement des producteurs de kombucha, il serait important de mimer au mieux des starters équivalents à ceux du back-slopping traditionnel comprenant la fraction biofilm et la fraction liquide, afin de tester différentes conditions de fermentation. Pour ce type d'expérience, le design que nous avons développé à partir des rétentats de filtration, pourra être utilisé afin de générer des réplicats de biofilm fille généré au cycle C00 de démarrage et de les placer dans différentes conditions de conservation. Il faudra néanmoins vérifier si les compositions homogènes que nous avons observées pour les fractions liquides

des cycles C00 des lignées parallèles, le sont également pour celles des biofilms fille correspondantes.

### 3. Conclusions

Les résultats des trois types d'expériences décrites dans ce chapitre ont permis non seulement de valider des protocoles de suivi de lignées de fermentation, mais ils également permis de mettre en évidence que des facteurs extérieurs peuvent avoir un impact détectable sur la composition microbienne des SCOBY.

Ces protocoles vont servir de base au développement d'essais à plus grande échelle afin de tester les effets de nombreuses conditions, comme la concentration en saccharose, le type et la qualité de thé, la température de fermentation, la présence d'agents toxiques ou l'ajout de taxons microbiens au SCOBY, tels que des moisissures ou des souches de levures dont les propriétés 'killer' ont été caractérisées.

Ces essais ont également montré que la reprise d'activité après congélation du starter apparaît comme une condition de choix pour étudier les cinétiques d'activité des taxons des SCOBY et comprendre les rôles respectifs, ceci dans une perspective plus fondamentale visant à mieux connaître les interactions entre microorganismes dans une communauté.



# **Conclusion générale et perspectives**

Mon travail de thèse a débuté avec la mise en place des protocoles nécessaires pour effectuer des fermentations de kombucha au laboratoire, puis des protocoles pour caractériser les produits finis : quelques caractéristiques chimiques par dosages avec des kits enzymatiques et surtout la composition des communautés microbiennes des levures et des bactéries. Une partie importante de ces mises au point sont décrites dans le chapitre 1.

La maîtrise de ces protocoles a permis d'effectuer des études comparatives des compositions microbiennes de SCOBY, établies selon plusieurs design expérimentaux, décrits dans les chapitres 2 et 3. Avec la comparaison des 18 SCOBY reconstitués à partir de starters de différentes origines (chapitre 2), nous avons défini des micro-organismes importants dans la fermentation dukombucha : des taxons omniprésents constituant le core (*B. bruxellensis*, *B. anomalus*, *Komagataeibacter* et *Acetobacter*) et d'autres considérés comme accessoires, spécifiques à certains SCOBY. Nous avons constaté une distribution différente des genres bactériens *Komagataeibacter*, *Lactobacillus* et *Oenococcus* dans les fractions biofilm et liquide. La dynamique des communautés microbiennes a été décrite pendant trois cycles fermentaires consécutifs. Nous avons pu observer que la composition du consortium microbien évolue différemment d'un SCOBY à un autre, en termes d'abondance relative des taxons reconnus comme accessoires, qui prédominent dans certains cas. Il n'est pas clair, si ces variations observées sur trois cycles correspondent à une stabilisation progressive des compositions microbiennes ou si elles correspondent à des fluctuations récurrentes et donc susceptibles de se produire régulièrement au cours des générations successives. Pour cette dernière hypothèse, il est bien sûr important de décrire les profils de ces variations et éventuellement d'identifier les types de causes de ces fluctuations.

Dans ce contexte, les résultats des comparaisons décrites au chapitre 3, ont permis d'apporter des éléments sur différents aspects de ces variations. La comparaison des lignées de fermentations qui constituent des réplicats biologiques propagés sur 12 à 18 cycles consécutifs, a mis en évidence que, à partir de compositions en taxons similaires lors des premiers cycles, chaque lignée étudiée a présenté des fluctuations progressives de composition qui lui sont propres, indiquant que la composition initiale n'est pas le seul paramètre à avoir un effet sur les variations d'un cycle à l'autre. Ces comparaisons ont également montré que toutes les lignées étudiées peuvent présenter une variation soudaine de la composition microbienne, exactement au même cycle, suggérant la survenue d'un événement externe dont on peut détecter les effets. Les résultats de comparaisons, cette fois menés sur des lignées dont les starters ont été soumis à des traitements différents (ici des conditions de conservation différentes), ont en effet montré des effets visibles macroscopiquement, avec différentes cinétiques de formation des biofilms fille et des effets détectables au niveau des compositions microbiennes, également impactées par ces différences de conditions.

Dans la même optique de suivi de la composition au cours des cycles de fermentation, les résultats observés pour trois lignées de SCOBY en conditions réelles de production chez Bio Brasseurs ont montré de grandes différences dans leur composition microbienne. Là aussi, des variations sont soit spécifiques d'un brassin ou bien elles touchent tous les brassins parallèles, en relation avec des événements temporaires liés à l'interruption occasionnelle de la production ou les variations saisonnières.

Avec les dérives des compositions microbiennes observées dans les lignées suivies chez BioBrasseurs, il serait intéressant de savoir quels effets ont ces variations de composition, sur les métabolites présents dans le produit fini et sur ses propriétés sensorielles. Dans les lignées plusieurs taxons ont présenté des

variations, comme la dominance de *B. anomalus* aux dépens de *B. bruxellensis* ou la dominance de *Lactobacillus* par rapport à *Oenococcus*, l'abondance permanente de *S. cerevisiae* dans toutes les lignées et les abondances ponctuelles marquées par la prédominance de *Kluyveromyces* au printemps, de *Kregervanrija* lors du déménagement ou de *H. valbyensis*, quelques mois après le déménagement. Chez Bio Brasseurs, des tests sont effectués pour mesurer le pH, les sucres solubles, exprimés en degrés Brix avec un réfractomètre et le taux d'éthanol par le dosage avec un kit enzymatique. Ces mesures se font systématiquement sur tous les brassins afin de surveiller rapidement les métabolites clés d'une fermentation de type kombucha. A la fin de chaque fermentation, le brasseur et la cogérante, ayant développé une expertise subjective, valident le goût et si le brassin passe aux étapes de commercialisation. Il s'agit d'une analyse descriptive, basée sur l'évaluation des attributs : apparence, couleur, odeur, goût et l'acceptation générale d'un produit. Elle permet de recueillir les données sur sa qualité et elle est menée par des consommateurs (Zubaidah et al., 2018). Aucun changement des propriétés globales n'a été constaté sur l'appréciation du kombucha des brassins pour lesquels nous avons détecté des variations importantes de composition microbienne. Néanmoins, une étude métabolomique pointue, permettra de voir si ces variations peuvent être corrélées à des changements des profils métabolomiques, aussi bien sur les brassins de production que sur des lignées de laboratoire. Afin d'étudier les principaux métabolites du kombucha, des méthodes chromatographiques en phase liquide ou en phase gazeuse, couplées ou non à des méthodes de spectrométrie de masse ont été décrites dans différentes études portant sur le kombucha (Huang et al., 2022; Savary et al., 2021; Tran et al., 2022b; Villarreal-Soto et al., 2020). Ces méthodes sont plus appropriées pour des études systématiques, en comparaison aux dosages enzymatiques que nous avons réalisés, qui sont très limités en termes de nombre de métabolites analysés.

Comme les relations entre métabolites et saveurs sont complexes à prédire, une dégustation semble être l'outil indispensable pour compléter ces caractérisations. Une évaluation quantitative à travers un panel sensoriel pourrait être développée avec un groupe de dégustateurs composé de 15 à 20 personnes, sélectionnés pour leur capacité à détecter, reconnaître et mémoriser des textures, des odeurs, des saveurs et des arômes caractéristiques. Ces aptitudes sensorielles sont développées après avoir fait une formation de plusieurs semaines à l'aide de protocoles normalisés (NF EN ISO 8586 – NF EN ISO 13299). Cette norme spécifie les critères pour choisir les personnes possédant des aptitudes sensorielles particulières parmi les sujets qualifiés, ou parmi des spécialistes en produits. Il s'agit d'un entraînement intensif dont l'objectif est d'augmenter les connaissances des panélistes pour devenir sujets experts capables de réaliser un profil sensoriel. Un profil est obtenu en attribuant des scores à des descripteurs sensoriels en fonction de l'intensité. Pour simplifier, les membres du panel sensoriel se familiarisent avec le vocabulaire spécifique qui a été retenu pour décrire les produits de la gamme étudiée. Ces descripteurs/ attributs sensoriels sont calibrés en mesurant l'intensité de chacun des critères avec un niveau de performance proche à celui d'un instrument de mesure. L'établissement d'un profil sensoriel, se fait par une attribution d'un coefficient à chaque descripteur sensoriel en fonction de la perception du produit fini optimal désiré. Un panel entraîné permet de vérifier la conformité et la stabilité sensorielle des lots et de détecter et mesurer l'impact d'un changement de formulation ou de procédé sur les caractéristiques sensorielles d'un produit. Néanmoins, l'absence des standards qui peuvent décrire le kombucha limitent la mise en place d'un panel sensoriel de kombucha dans le court terme.

Les résultats des analyses sensorielles et métabolomiques peuvent rassembler les données nécessaires pour mieux comprendre les différentes interactions

microbiennes et métaboliques. L'exploration de ces interactions est une perspective pour la poursuite des travaux initiés pendant ma thèse. Pour ma thèse, les profils taxonomiques microbiens ont été générés par métagénomique ciblée avec une approche par métabarcoding. Cette approche présente un certain nombre de limites que je vais décrire ci-dessous. D'une part, il est assez difficile de faire le lien entre les abondances relatives des amplicons générés par métabarcoding et les abondances absolues des taxons dans un échantillon. Cela est dû au nombre de copies variables des opérons codant les ARN ribosomiaux et porteurs des régions ciblées. Ce nombre de copies étant variable et imprédictible d'une souche à l'autre de la même espèce, les corrections des calculs d'abondances restent peu fiables (Louca et al., 2018). Si les approches par métabarcoding sont pertinentes pour détecter des changements dans les études comparatives, il est plus difficile d'envisager leur utilisation dans une perspective de compréhension des interactions, entre taxons et avec l'environnement et avec les compositions en métabolites. L'utilisation des méthodes métagénomiques non ciblées, en séquençage courte ou longue lecture, permettraient de résoudre ces limites. Elles permettent également de détecter les proportions relatives et les variations entre les taxons, qui ne peuvent pas être évaluées si les taxons sont détectés avec des marqueurs différents (en général pour les taxons appartenant à des domaines ou des règnes différents) et dont les interactions sont essentielles au fonctionnement des communautés.

D'autre part, une limite de l'approche par métabarcoding réside dans la résolution taxonomique des OTU, au rang taxonomique espèce pour les levures mais seulement et au rang genre (et même famille) pour les bactéries. En ayant recours à des approches métagénomiques non ciblées, il est possible d'obtenir des séquences qui peuvent être assemblées et comparées aux génomes des bases de données (De Filippis et al., 2018a ; Oulas et al., 2015), cette résolution pourrait

donc être améliorée et pourrait même atteindre le rang taxonomique de la souche. Nos analyses d'isolats de levures ayant montré qu'un même SCOBY peut contenir plusieurs souches de l'espèce *B. bruxellensis* (chapitre 2), nous pourrions préciser dans quelle mesure une telle diversité est présente pour d'autres taxons des SCOBY. Enfin, dans le contexte des interactions microbiennes, il serait possible de vérifier si les variations de compositions microbiennes concernent également la composition en souches : lorsque l'abondance d'une espèce fluctue au cours des cycles successifs, est-ce la même souche qui prédomine après chaque fluctuation ou des souches différentes, qui seraient plus aptes à résister aux nouvelles conditions ? Une approche non ciblée en séquençage longues lectures serait particulièrement appropriée afin de pouvoir assembler des longs contigs pouvant correspondre aux chromosomes des bactéries et des levures est d'estimer le nombre de souches présentes pour une même espèce.

A défaut de pouvoir assembler des génomes complets, les données générées par métagénomique non ciblée permettront non seulement d'étudier des communautés de taxons, mais également des communautés de gènes et d'éléments génétiques tels que les plasmides et les prophages. La prédiction des gènes et leur annotation fonctionnelle permettent de connaître les fonctions potentielles présentes dans une communauté, telles que des voies de biosynthèse, des transporteurs ou la production d'antimicrobiens. La distribution de ces fonctions dans les taxons présents permettra également de prédire leur implication potentielle dans des interactions. Cette analyse pourra être complétée par un suivi des fluctuations éventuelles des fonctions, si elles suivent celles des abondances en taxons et si elles sont associées à des enrichissements en groupes de gènes caractéristiques ou de voies métaboliques potentielles particulières. Ces données permettront également de fournir de nouvelles informations sur les génomes des taxons identifiés, par exemple d'élargir les répertoires de gènes constituant les

pangénomes. Ce type d'approche a déjà été décrit pour des SCOBY de kombucha et a permis de caractériser les génomes des seuls taxons présents dans ces SCOBY dont la composition nous semble cependant atypique, avec les seules espèces *Zygosaccharomyces bailii* et *Komagataeibacter rhaeticus*, de levure et bactérie, respectivement (Arikan et al., 2020).

Comme il n'est pas clair si une fonction peut être assurée, même quand elle est présente dans les génomes et éventuellement enrichie dans les communautés de gènes, une analyse exhaustive des produits des gènes peut rassembler des données supplémentaires pour mieux comprendre les interactions microbiennes. Ainsi, le métatranscriptome peut être analysé, en utilisant des approches de RNAseq (De Filippis et al., 2018a). Ce métatranscriptome reflète l'activité des taxons, d'où la possibilité de différencier les micro-organismes métaboliquement actifs et passifs, notamment dans des conditions différentes (Wolfe et al. 2014) et son analyse permet également d'améliorer la résolution d'une communauté au niveau souche. En perspective de poursuivre les travaux initiés dans la thèse, cette approche métatranscriptomique serait particulièrement intéressante pour étudier l'aspect fondamental de la cinétique d'assemblage d'une communauté microbienne. En effet, lorsque nous avons reconstitué des SCOBY à partir des starters conservés dans différentes conditions, nous avons remarqué que les prédominances des taxons ne sont pas les mêmes selon les conditions de conservation, avec notamment une prédominance de *S. cerevisiae* pour les conditions -20°C et -40°C, corrélée à une très faible proportion de *B. bruxellensis* dans les cycles ultérieurs. Une comparaison des cinétiques d'expression pourrait permettre d'établir s'il y a une succession différente de l'activité des taxons, selon les conditions de conservation des starters, et quels types de gènes sont concernés.

Lorsque nous avons mis en évidence les variations de composition microbienne, entre SCOBY naturels différents et entre les cycles d'un même SCOBY naturel initial, un questionnement que nous n'avons pu résoudre est celui des taxons qui deviennent soudainement abondants, alors qu'ils n'étaient pas détectés lors des cycles précédents. Étaient-ils présents à bas bruit dans les SCOBY, mais avec une abondance trop faible pour être détectés ? Et ont-ils pu devenir prépondérants à l'occasion de changements de conditions, ou s'il s'agit de contaminations microbiennes ? Afin que ce questionnement ne vienne pas interférer avec les expériences de cinétiques d'assemblage des SCOBY que nous souhaiterions développer, ces expériences pourraient être conduites avec des starters définis, constitués de co-cultures dont la composition est connue. La collection d'isolats décrite au chapitre 2, ainsi que le souchier de l'espèce *B. bruxellensis* disponible au laboratoire, permettraient de développer ces SCOBY de kombucha synthétiques qui ont été déjà décrits dans certaines études (Huang et al., 2022 ; Savary et al., 2021 ; Tran et al., 2022a). Cependant, il ne s'agirait pas ici d'effectuer une stratégie de réduction de la complexité, visées par ces études, mais de disposer de systèmes bien définis de par leurs contenus en multiples taxons voire en multiples souches de chaque taxon. Une première étape constituer ces systèmes, sera le choix des souches à mélanger et leurs proportions initiales, qui pourra se faire en fonction des résultats apportés par les analyses de métagénomiques non ciblées. L'étape suivante consistera à vérifier la stabilité de ces systèmes définis, au cours de back-slopping successifs et s'ils présentent des variations, de les comparer ces profils de variation à ceux des SCOBY naturels. Ces communautés synthétiques seraient ainsi des modèles de choix et pourraient être confectionnés à façon pour comprendre les interactions entre les taxons en les analysant par les approches métagénomique, métatranscriptomique et métabolomiques proposées ci-dessus, sans oublier l'évaluation des propriétés sensorielles de kombuchas issus de leurs activités.



## Références

- Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., Blankenberg, D., 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. <https://doi.org/10.1093/nar/gky379>
- Aldrete-Tapia, J.A., Escalante-Minakata, P., Martínez-Peniche, R.A., Tamplin, M.L., Hernández-Iturriaga, M., 2020. Yeast and bacterial diversity, dynamics and fermentative kinetics during small-scale tequila spontaneous fermentation. *Food Microbiol.* 86, 103339. <https://doi.org/10.1016/j.fm.2019.103339>
- Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., Huse, S.M., 2009. A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLOS ONE* 4, e6372. <https://doi.org/10.1371/journal.pone.0006372>
- Angly, F.E., Felts, B., Breitbart, M., Salamon, P., Edwards, R.A., Carlson, C., Chan, A.M., Haynes, M., Kelley, S., Liu, H., Mahaffy, J.M., Mueller, J.E., Nulton, J., Olson, R., Parsons, R., Rayhawk, S., Suttle, C.A., Rohwer, F., 2006. The Marine Viromes of Four Oceanic Regions. *PLOS Biol.* 4, e368. <https://doi.org/10.1371/journal.pbio.0040368>
- Arıkan, M., Mitchell, A.L., Finn, R.D., Gürel, F., 2020. Microbial composition of Kombucha determined using amplicon sequencing and shotgun metagenomics. *J. Food Sci.* 85, 455–464. <https://doi.org/10.1111/1750-3841.14992>
- Avramova, M., Cibrario, A., Peltier, E., Coton, M., Coton, E., Schacherer, J., Spano, G., Capozzi, V., Blaiotta, G., Salin, F., Dols-Lafargue, M., Grbin, P., Curtin, C., Albertin, W., Masneuf-Pomarede, I., 2018. *Brettanomyces bruxellensis* population survey reveals a diploid-triploid complex structured according to substrate of isolation and geographical distribution. *Sci. Rep.* 8, 4136. <https://doi.org/10.1038/s41598-018-22580-7>
- Banasiewicz, J., Lisboa, B.B., da Costa, P.B., Schlindwein, G., Venter, S.N., Steenkamp, E.T., Vargas, L.K., Passaglia, L.M.P., Stepkowski, T., 2021. Culture-independent assessment of the diazotrophic *Bradyrhizobium* communities in the Pampa and Atlantic Forest Biomes localities in southern Brazil. *Syst. Appl. Microbiol.* 44, 126228. <https://doi.org/10.1016/j.syapm.2021.126228>
- Barakat, N., Beaufort, S., Rizk, Z., Bouajila, J., Taillandier, P., El Rayess, Y., 2022. Kombucha analogues around the world: A review. *Crit. Rev. Food Sci. Nutr.* 0, 1–25. <https://doi.org/10.1080/10408398.2022.2069673>
- Barreau, M., Pagnier, I., La Scola, B., 2013. Improving the identification of anaerobes in the clinical microbiology laboratory through MALDI-TOF mass spectrometry. *Anaerobe* 22, 123–125. <https://doi.org/10.1016/j.anaerobe.2013.04.011>

- Belda, I., Ruiz, J., Alonso, A., Marquina, D., Santos, A., 2017. The Biology of *Pichia membranifaciens* Killer Toxins. *Toxins* 9, E112. <https://doi.org/10.3390/toxins9040112>
- Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., Kauserud, H., 2010a. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol.* 10, 189. <https://doi.org/10.1186/1471-2180-10-189>
- Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., Kauserud, H., 2010b. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol.* 10, 189. <https://doi.org/10.1186/1471-2180-10-189>
- Blaalid, R., Kumar, S., Nilsson, R.H., Abarenkov, K., Kirk, P.M., Kauserud, H., 2013. ITS 1 versus ITS 2 as DNA metabarcodes for fungi. *Mol. Ecol. Resour.* 13, 218–224. <https://doi.org/10.1111/1755-0998.12065>
- Blanc, P.J., 1996. Characterization of the tea fungus metabolites. *Biotechnol. Lett.* 18, 139–142. <https://doi.org/10.1007/BF00128667>
- Blasche, S., Kim, Y., Mars, R.A.T., Machado, D., Maansson, M., Kafkia, E., Milanese, A., Zeller, G., Teusink, B., Nielsen, J., Benes, V., Neves, R., Sauer, U., Patil, K.R., 2021. Metabolic cooperation and spatiotemporal niche partitioning in a kefir microbial community. *Nat. Microbiol.* 6, 196–208. <https://doi.org/10.1038/s41564-020-00816-5>
- Bokulich NA, Mills DA., 2013. Improved selection of internal transcribed spacer-specific primers enables quantitative, ultra-high-throughput profiling of fungal communities. *Appl Environ Microbiol* 79, 19-26. <https://doi.org/10.1128/AEM.03870-12>.
- Bukin, Y.S., Galachyants, Y.P., Morozov, I.V., Bukin, S.V., Zakharenko, A.S., Zemskaya, T.I., 2019. The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Sci. Data* 6, 190007. <https://doi.org/10.1038/sdata.2019.7>
- Callahan, B.J., Wong, J., Heiner, C., Oh, S., Theriot, C.M., Gulati, A.S., McGill, S.K., Dougherty, M.K., 2019. High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Res.* 47, e103. <https://doi.org/10.1093/nar/gkz569>
- Callanan, M., Naranjo, D.R., McAuliffe, O., 2021. 1.10 - Genomic Insights Into Food Fermentations, in: Cifuentes, A. (Ed.), *Comprehensive Foodomics*. Elsevier, Oxford, pp. 160–170. <https://doi.org/10.1016/B978-0-08-100596-5.22698-X>
- Cariello, N.F., Swenberg, J.A., De Bellis, A., Skopek, T.R., 1991. Analysis of mutations using PCR and denaturing gradient gel electrophoresis. *Environ. Mol. Mutagen.* 18, 249–254. <https://doi.org/10.1002/em.2850180408>
- Chakravorty, S., Bhattacharya, S., Chatzinotas, A., Chakraborty, W., Bhattacharya, D., Gachhui, R., 2016. Kombucha tea fermentation: Microbial and biochemical dynamics. *Int. J. Food Microbiol.* 220, 63–72. <https://doi.org/10.1016/j.ijfoodmicro.2015.12.015>
- Code de la santé publique, 2016. Article L3321-1 - Code de la santé publique - Légifrance [WWW Document]. URL (accessed 11.8.22).
- Comitini, F., Gobbi, M., Domizio, P., Romani, C., Lencioni, L., Mannazzu, I., Ciani, M., 2011. Selected non-Saccharomyces wine yeasts in controlled multistarter

fermentations with *Saccharomyces cerevisiae*. *Food Microbiol.* 28, 873–882.  
<https://doi.org/10.1016/j.fm.2010.12.001>

- Comuzzo, P., Battistutta, F., 2019. Chapter 2 - Acidification and pH Control in Red Wines, in: Morata, A. (Ed.), *Red Wine Technology*. Academic Press, pp. 17–34. <https://doi.org/10.1016/B978-0-12-814399-5.00002-5>
- Corsetti, A., Perpetuini, G., Schirone, M., Tofalo, R., Suzzi, G., 2012. Application of starter cultures to table olive fermentation: an overview on the experimental studies. *Front. Microbiol.* 3.
- Coton, M., Pawtowski, A., Taminiau, B., Burgaud, G., Deniel, F., Coulloume-Labarthe, L., Fall, A., Daube, G., Coton, E., 2017. Unraveling microbial ecology of industrial-scale Kombucha fermentations by metabarcoding and culture-based methods. *FEMS Microbiol. Ecol.* 93. <https://doi.org/10.1093/femsec/fix048>
- Crum, H., LaGory, A., 2016. *The Big Book of Kombucha: Brewing, Flavoring, and Enjoying the Health Benefits of Fermented Tea*. Storey Publishing.
- Ćurčić, Ž., Kosovac, A., Stepanović, J., Rekanović, E., Kube, M., Duduk, B., 2021. Multilocus Genotyping of ‘Candidatus Phytoplasma solani’ Associated with Rubbery Taproot Disease of Sugar Beet in the Pannonian Plain. *Microorganisms* 9, 1950. <https://doi.org/10.3390/microorganisms9091950>
- Cvetković, D., Markov, S., Djurić, M., Savić, D., Velićanski, A., 2008. Specific interfacial area as a key variable in scaling-up Kombucha fermentation. *J. Food Eng.* 85, 387–392. <https://doi.org/10.1016/j.jfoodeng.2007.07.021>
- D’Andrea, S., Cuscó, A., Francino, O., 2021. Rapid and real-time identification of fungi up to species level with long amplicon nanopore sequencing from clinical samples, *Biology Methods and Protocols*, Volume 6, Issue 1, bpaa026, <https://doi.org/10.1093/biomethods/bpaa026>
- Dakal, T.C., Giudici, P., Solieri, L., 2016. Contrasting Patterns of rDNA Homogenization within the *Zygosaccharomyces rouxii* Species Complex. *PLOS ONE* 11, e0160744. <https://doi.org/10.1371/journal.pone.0160744>
- De Filippis, F., Parente, E., Ercolini, D., 2018a. Recent Past, Present, and Future of the Food Microbiome. *Annu. Rev. Food Sci. Technol.* 9, 589–608. <https://doi.org/10.1146/annurev-food-030117-012312>
- De Filippis, F., Troise, A.D., Vitaglione, P., Ercolini, D., 2018b. Different temperatures select distinctive acetic acid bacteria species and promotes organic acids production during Kombucha tea fermentation. *Food Microbiol.* 73, 11–16. <https://doi.org/10.1016/j.fm.2018.01.008>
- Derx, P.M., Janzen, T., Sørensen, K.I., Christensen, J.E., Stuer-Lauridsen, B., Johansen, E., 2014. The art of strain improvement of industrial lactic acid bacteria without the use of recombinant DNA technology. *Microb. Cell Factories* 13, S5. <https://doi.org/10.1186/1475-2859-13-S1-S5>
- Diez-Ozaeta, I., Astiazaran, O.J., 2022. Recent advances in Kombucha tea: Microbial consortium, chemical parameters, health implications and biocellulose production. *Int. J. Food Microbiol.* 377, 109783. <https://doi.org/10.1016/j.ijfoodmicro.2022.109783>
- Drewnowski, A., Gomez-Carneros, C., 2000. Bitter taste, phytonutrients, and the consumer: a review. *Am. J. Clin. Nutr.* 72, 1424–1435. <https://doi.org/10.1093/ajcn/72.6.1424>
- Dufresne, C., Farnworth, E., 2000. Tea, Kombucha, and health: a review. *Food Res. Int.* 33, 409–421. [https://doi.org/10.1016/S0963-9969\(00\)00067-3](https://doi.org/10.1016/S0963-9969(00)00067-3)

- Dutta, D., Gachhui, R. 2007, 2007. Nitrogen-fixing and cellulose-producing *Gluconacetobacter kombuchae* sp. nov., isolated from Kombucha tea. *Int. J. Syst. Evol. Microbiol.* 57, 353–357. <https://doi.org/10.1099/ijs.0.64638-0>
- Dutta, H., Paul, S.K., 2019. Kombucha Drink: Production, Quality, and Safety Aspects, in: *Production and Management of Beverages*. Elsevier, pp. 259–288. <https://doi.org/10.1016/B978-0-12-815260-7.00008-0>
- Eberlein, C., Saada, O.A., Friedrich, A., Albertin, W., Schacherer, J., 2021. Different trajectories of polyploidization shape the genomic landscape of the *Brettanomyces bruxellensis* yeast species. *Genome Res.* 31, 2316–2326. <https://doi.org/10.1101/gr.275380.121>
- Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. <https://doi.org/10.1093/bioinformatics/btr381>
- Edwards, R.A., Rodriguez-Brito, B., Wegley, L., Haynes, M., Breitbart, M., Peterson, D.M., Saar, M.O., Alexander, S., Alexander, E.C., Rohwer, F., 2006. Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genomics* 7, 57. <https://doi.org/10.1186/1471-2164-7-57>
- Edwards, R.A., Rohwer, F., 2005. Viral metagenomics. *Nat. Rev. Microbiol.* 3, 504–510. <https://doi.org/10.1038/nrmicro1163>
- Emiljanowicz, K.E., Malinowska-Pańczyk, E., 2020. Kombucha from alternative raw materials – The review. *Crit. Rev. Food Sci. Nutr.* 60, 3185–3194. <https://doi.org/10.1080/10408398.2019.1679714>
- Englezos, V., Giacosa, S., Rantsiou, K., Rolle, L., Cocolin, L., 2017. *Starmerella bacillaris* in winemaking: opportunities and risks. *Curr. Opin. Food Sci., Food Engineering and Processing • Food Mycology* 17, 30–35. <https://doi.org/10.1016/j.cofs.2017.08.007>
- Ercolini, D., 2004. PCR-DGGE fingerprinting: novel strategies for detection of microbes in food. *J. Microbiol. Methods* 56, 297–314. <https://doi.org/10.1016/j.mimet.2003.11.006>
- Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., Maman, S., Hernandez-Raquet, G., Combes, S., Pascal, G., 2018. FROGS: Find, Rapidly, OTUs with Galaxy Solution. *Bioinformatics* 34, 1287–1294. <https://doi.org/10.1093/bioinformatics/btx791>
- Espejo, R.T., Plaza, N., 2018. Multiple Ribosomal RNA Operons in Bacteria; Their Concerted Evolution and Potential Consequences on the Rate of Evolution of Their 16S rRNA. *Front. Microbiol.* 9, 1232. <https://doi.org/10.3389/fmicb.2018.01232>
- Evans, J.S., López-Legentil, S., Erwin, P.M., 2018. Comparing Two Common DNA Extraction Kits for the Characterization of Symbiotic Microbial Communities from Ascidian Tissue. *Microbes Environ.* 33, 435–439. <https://doi.org/10.1264/jsme2.ME18031>
- Fell, J.W., Boekhout, T., Fonseca, A., Scorzetti, G., Stätzell-Tallman, A., 2000. Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. *Int. J. Syst. Evol. Microbiol.* 50, 1351–1371. <https://doi.org/10.1099/00207713-50-3-1351>

- Fournier, T., Gounot, J.-S., Freel, K., Cruaud, C., Lemainque, A., Aury, J.-M., Wincker, P., Schacherer, J., Friedrich, A., 2017. High-Quality de Novo Genome Assembly of the *Dekkera bruxellensis* Yeast Using Nanopore MinION Sequencing. *G3 Bethesda Md* 7, 3243–3250. <https://doi.org/10.1534/g3.117.300128>
- Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A., Olsen, G.J., 2008. Critical Evaluation of Two Primers Commonly Used for Amplification of Bacterial 16S rRNA Genes. *Appl. Environ. Microbiol.* 74, 2461–2470. <https://doi.org/10.1128/AEM.02272-07>
- Gaggìa, F., Baffoni, L., Galiano, M., Nielsen, D.S., Jakobsen, R.R., Castro-Mejía, J.L., Bosi, S., Truzzi, F., Musumeci, F., Dinelli, G., Di Gioia, D., 2018. Kombucha Beverage from Green, Black and Rooibos Teas: A Comparative Study Looking at Microbiology, Chemistry and Antioxidant Activity. *Nutrients* 11, 1. <https://doi.org/10.3390/nu11010001>
- Gänzle, M., 2022. The periodic table of fermented foods: limitations and opportunities. *Appl. Microbiol. Biotechnol.* <https://doi.org/10.1007/s00253-022-11909-y>
- Gebremeskel, F.T., Alemayehu, T., Ali, M.M., 2022. Methicillin-resistant *Staphylococcus aureus* antibiotic susceptibility profile and associated factors among hospitalized patients at Hawassa University Comprehensive Specialized Hospital, Ethiopia. *IJID Reg. Online* 3, 129–134. <https://doi.org/10.1016/j.ijregi.2022.03.015>
- Gibbons, J.G., Rinker, D.C., 2015. The genomics of microbial domestication in the fermented food environment. *Curr. Opin. Genet. Dev., Genomes and evolution* 35, 1–8. <https://doi.org/10.1016/j.gde.2015.07.003>
- Gill, S.R., Pop, M., DeBoy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., Gordon, J.I., Relman, D.A., Fraser-Liggett, C.M., Nelson, K.E., 2006. Metagenomic Analysis of the Human Distal Gut Microbiome. *Science* 312, 1355–1359. <https://doi.org/10.1126/science.1124234>
- Góes-Neto, A., Kukharensko, O., Orlovska, I., Podolich, O., Imchen, M., Kumavath, R., Kato, R.B., de Carvalho, D.S., Tiwari, S., Brenig, B., Azevedo, V., Reva, O., de Vera, J.-P.P., Kozyrovska, N., Barh, D., 2021. Shotgun metagenomic analysis of kombucha mutualistic community exposed to Mars-like environment outside the International Space Station. *Environ. Microbiol.* 23, 3727–3742. <https://doi.org/10.1111/1462-2920.15405>
- Gounot, J.-S., Neuvéglise, C., Freel, K.C., Devillers, H., Piškur, J., Friedrich, A., Schacherer, J., 2020. High Complexity and Degree of Genetic Variation in *Brettanomyces bruxellensis* Population. *Genome Biol. Evol.* 12, 795–807. <https://doi.org/10.1093/gbe/evaa077>
- Gramza-Michałowska, A., Kulczyński, B., Xindi, Y., Gumienna, M., 2016. Research on the effect of culture time on the kombucha tea beverage's antiradical capacity and sensory value. *Acta Sci. Pol. Technol. Aliment.* 15, 447–457. <https://doi.org/10.17306/J.AFS.2016.4.43>
- Grassi, A., Cristani, C., Palla, M., Di Giorgi, R., Giovannetti, M., Agnolucci, M., 2022. Storage time and temperature affect microbial dynamics of yeasts and acetic acid bacteria in a kombucha beverage. *Int. J. Food Microbiol.* 382, 109934. <https://doi.org/10.1016/j.ijfoodmicro.2022.109934>

- Guiasu, R.C., Guiasu, S., 2010. The Rich-Gini-Simpson quadratic index of biodiversity. *Nat. Sci.* 02, 1130–1137. <https://doi.org/10.4236/ns.2010.210140>
- Gupte, Y., Kulkarni, A., Raut, B., Sarkar, P., Choudhury, R., Chawande, A., Kumar, G.R.K., Bhadra, B., Satapathy, A., Das, G., Vishnupriya, B., Dasgupta, S., 2021. Characterization of nanocellulose production by strains of *Komagataeibacter* sp. isolated from organic waste and Kombucha. *Carbohydr. Polym.* 266, 118176. <https://doi.org/10.1016/j.carbpol.2021.118176>
- Guzman, J., Poehlein, A., Glaeser, S.P., Schwengers, O., Blom, J., Hollensteiner, J., Kämpfer, P., Vilcinskis, A., 2022. *Pseudocitrobacter corydidari* sp. nov., isolated from the Asian emerald cockroach *Corydidarum magnifica*. *Int. J. Syst. Evol. Microbiol.* 72. <https://doi.org/10.1099/ijsem.0.005497>
- Hamady, M., Knight, R., 2009. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res.* 19, 1141–1152. <https://doi.org/10.1101/gr.085464.108>
- Harrison, K., Curtin, C., 2021. Microbial Composition of SCOBY Starter Cultures Used by Commercial Kombucha Brewers in North America. *Microorganisms* 9, 1060. <https://doi.org/10.3390/microorganisms9051060>
- Hebert, P.D.N., Ratnasingham, S., de Waard, J.R., 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B Biol. Sci.* 270, S96–S99. <https://doi.org/10.1098/rsbl.2003.0025>
- Huang, X., Xin, Y., Lu, T., 2022. A Systematic, Complexity-Reduction Approach to Dissect Microbiome: the Kombucha Tea Microbiome as an Example. <https://doi.org/10.1101/2022.01.12.475982>
- Huber, J.A., Mark Welch, D.B., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., Sogin, M.L., 2007. Microbial Population Structures in the Deep Marine Biosphere. *Science* 318, 97–100. <https://doi.org/10.1126/science.1146689>
- Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., FitzGerald, M.G., Fulton, R.S., Giglio, M.G., Hallsworth-Pepin, K., Lobos, E.A., Madupu, R., Magrini, V., Martin, J.C., Mitreva, M., Muzny, D.M., Sodergren, E.J., Versalovic, J., Wollam, A.M., Worley, K.C., Wortman, J.R., Young, S.K., Zeng, Q., Aagaard, K.M., Abolude, O.O., Allen-Vercoe, E., Alm, E.J., Alvarado, L., Andersen, G.L., Anderson, S., Appelbaum, E., Arachchi, H.M., Armitage, G., Arze, C.A., Ayvaz, T., Baker, C.C., Begg, L., Belachew, T., Bhonagiri, V., Bihan, M., Blaser, M.J., Bloom, T., Bonazzi, V., Paul Brooks, J., Buck, G.A., Buhay, C.J., Busam, D.A., Campbell, J.L., Canon, S.R., Cantarel, B.L., Chain, P.S.G., Chen, I-M.A., Chen, L., Chhibba, S., Chu, K., Ciulla, D.M., Clemente, J.C., Clifton, S.W., Conlan, S., Crabtree, J., Cutting, M.A., Davidovics, N.J., Davis, C.C., DeSantis, T.Z., Deal, C., Delehaunty, K.D., Dewhirst, F.E., Deych, E., Ding, Y., Dooling, D.J., Dugan, S.P., Michael Dunne, W., Scott Durkin, A., Edgar, R.C., Erlich, R.L., Farmer, C.N., Farrell, R.M., Faust, K., Feldgarden, M., Felix, V.M., Fisher, S., Fodor, A.A., Forney, L.J., Foster, L., Di Francesco, V., Friedman, J., Friedrich, D.C., Fronick, C.C., Fulton, L.L., Gao, H., Garcia, N., Giannoukos, G., Giblin, C., Giovanni, M.Y., Goldberg, J.M., Goll, J., Gonzalez, A., Griggs, A., Gujja, S., Kinder Haake, S., Haas, B.J., Hamilton, H.A., Harris, E.L., Hepburn, T.A., Herter, B., Hoffmann, D.E., Holder, M.E., Howarth, C., Huang, K.H., Huse,

- S.M., Izard, J., Jansson, J.K., Jiang, H., Jordan, C., Joshi, V., Katancik, J.A., Keitel, W.A., Kelley, S.T., Kells, C., King, N.B., Knights, D., Kong, H.H., Koren, O., Koren, S., Kota, K.C., Kovar, C.L., Kyrpides, N.C., La Rosa, P.S., Lee, S.L., Lemon, K.P., Lennon, N., Lewis, C.M., Lewis, L., Ley, R.E., Li, K., Liolios, K., Liu, B., Liu, Y., Lo, C.-C., Lozupone, C.A., Dwayne Lunsford, R., Madden, T., Mahurkar, A.A., Mannon, P.J., Mardis, E.R., Markowitz, V.M., Mavromatis, K., McCorrison, J.M., McDonald, D., McEwen, J., McGuire, A.L., McInnes, P., Mehta, T., Mihindukulasuriya, K.A., Miller, J.R., Minx, P.J., Newsham, I., Nusbaum, C., O’Laughlin, M., Orvis, J., Pagani, I., Palaniappan, K., Patel, S.M., Pearson, M., Peterson, J., Podar, M., Pohl, C., Pollard, K.S., Pop, M., Priest, M.E., Proctor, L.M., Qin, X., Raes, J., Ravel, J., Reid, J.G., Rho, M., Rhodes, R., Riehle, K.P., Rivera, M.C., Rodriguez-Mueller, B., Rogers, Y.-H., Ross, M.C., Russ, C., Sanka, R.K., Sankar, P., Fah Sathirapongsasuti, J., Schloss, J.A., Schloss, P.D., Schmidt, T.M., Scholz, M., Schriml, L., Schubert, A.M., Segata, N., Segre, J.A., Shannon, W.D., Sharp, R.R., Sharpton, T.J., Shenoy, N., Sheth, N.U., Simone, G.A., Singh, I., Smillie, C.S., Sobel, J.D., Sommer, D.D., Spicer, P., Sutton, G.G., Sykes, S.M., Tabbaa, D.G., Thiagarajan, M., Tomlinson, C.M., Torralba, M., Treangen, T.J., Truty, R.M., Vishnivetskaya, T.A., Walker, J., Wang, L., Wang, Z., Ward, D.V., Warren, W., Watson, M.A., Wellington, C., Wetterstrand, K.A., White, J.R., Wilczek-Boney, K., Wu, Y., Wylie, K.M., Wylie, T., Yandava, C., Ye, L., Ye, Y., Yooseph, S., Youmans, B.P., Zhang, L., Zhou, Y., Zhu, Y., Zoloth, L., Zucker, J.D., Birren, B.W., Gibbs, R.A., Highlander, S.K., Methé, B.A., Nelson, K.E., Petrosino, J.F., Weinstock, G.M., Wilson, R.K., White, O., The Human Microbiome Project Consortium, 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. <https://doi.org/10.1038/nature11234>
- Ivanišová, E., Meňhartová, K., Terentjeva, M., Godočíková, L., Árvay, J., Kačániová, M., 2019. Kombucha tea beverage: Microbiological characteristic, antioxidant activity, and phytochemical composition. *Acta Aliment.* 48, 324–331. <https://doi.org/10.1556/066.2019.48.3.7>
- Jarrell, J., Cal, T., Bennett, J.W., 2000. The Kombucha consortia of yeasts and bacteria. *Mycologist* 14, 166–170. [https://doi.org/10.1016/S0269-915X\(00\)80034-8](https://doi.org/10.1016/S0269-915X(00)80034-8)
- Jayabalan, R., Malbaša, R.V., Lončar, E.S., Vitas, J.S., Sathishkumar, M., 2014. A Review on Kombucha Tea—Microbiology, Composition, Fermentation, Beneficial Effects, Toxicity, and Tea Fungus. *Compr. Rev. Food Sci. Food Saf.* 13, 538–550. <https://doi.org/10.1111/1541-4337.12073>
- Jayabalan, R., Subathradevi, P., Marimuthu, S., Sathishkumar, M., Swaminathan, K., 2008. Changes in free-radical scavenging ability of kombucha tea during fermentation. *Food Chem.* 109, 227–234. <https://doi.org/10.1016/j.foodchem.2007.12.037>
- Jeong, J., Yun, K., Mun, S., Chung, W.-H., Choi, S.-Y., Nam, Y., Lim, M.Y., Hong, C.P., Park, C., Ahn, Y.J., Han, K., 2021. The effect of taxonomic classification by full-length 16S rRNA sequencing with a synthetic long-read technology. *Sci. Rep.* 11, 1727. <https://doi.org/10.1038/s41598-020-80826-9>
- Johnston-Monje, D., & Lopez Mejia, J., 2020. Botanical microbiomes on the cheap: Inexpensive molecular fingerprinting methods to study plant-associated communities of bacteria and fungi. *Applications in plant sciences*, 8(4), e11334.

<https://doi.org/10.1002/aps3.11334>

- Júnior, J.C. da S., Meireles Mafaldo, Í., de Lima Brito, I., Tribuzy de MagalhãesCordeiro, A.M., 2022. Kombucha: Formulation, chemical composition, and therapeutic potentialities. *Curr. Res. Food Sci.* 5, 360–365. <https://doi.org/10.1016/j.crfs.2022.01.023>
- Kapardar, R.K., Ranjan, R., Puri, M., Sharma, R., 2010. Sequence analysis of a salt tolerant metagenomic clone. *Indian J. Microbiol.* 50, 212–215. <https://doi.org/10.1007/s12088-010-0041-x>
- Kiangkitiwan, N., Wasanapiarnpong, T., Srikulkit, K., 2022. Multilayered bacterial cellulose/reduced graphene oxide composite films for self-standing and binder-free electrode application. *Heliyon* 8. <https://doi.org/10.1016/j.heliyon.2022.e10327>
- kombucha kamp, kamp, n.d. Origins and Legends of the Kombucha Mushroom Tea [WWW Document]. *Kombucha Kamp*. URL <https://www.kombuchakamp.com/what-is-kombucha/history-and-legends-of-kombucha> (accessed 9.10.21).
- Krause, L., Diaz, N.N., Bartels, D., Edwards, R.A., Pühler, A., Rohwer, F., Meyer, F., Stoye, J., 2006. Finding novel genes in bacterial communities isolated from the environment. *Bioinformatics* 22, e281–e289. <https://doi.org/10.1093/bioinformatics/bt1247>
- Kurtzman, C.P., 2011. Chapter 38 - Kregervanrija Kurtzman (2006), in: Kurtzman, C.P., Fell, J.W., Boekhout, T. (Eds.), *The Yeasts* (Fifth Edition). Elsevier, London, pp. 497–501. <https://doi.org/10.1016/B978-0-444-52149-1.00038-0>
- Kurtzman, C.P., 2006. New species and new combinations in the yeast genera *Kregervanrija* gen. nov., *Saturnispora* and *Candida*. *FEMS Yeast Res.* 6, 288–297. <https://doi.org/10.1111/j.1567-1364.2005.00019.x>
- Kurtzman, C.P., Fell, J.W., Boekhout, T., Robert, V., 2011. Methods for Isolation, Phenotypic Characterization and Maintenance of Yeasts, in: *The Yeasts*. Elsevier, pp. 87–110. <https://doi.org/10.1016/B978-0-444-52149-1.00007-0>
- Kurtzman, C.P., Robnett, C.J., 1998. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Van Leeuwenhoek* 73, 331–371. <https://doi.org/10.1023/A:1001761008817>
- Laavanya, D., Shirkole, S., Balasubramanian, P., 2021. Current challenges, applications and future perspectives of SCOBY cellulose of Kombucha fermentation. *J. Clean. Prod.* 295, 126454. <https://doi.org/10.1016/j.jclepro.2021.126454>
- Lachance, M.-A., 2011. Chapter 71 - *Starmerella Rosa & Lachance* (1998), in: Kurtzman, C.P., Fell, J.W., Boekhout, T. (Eds.), *The Yeasts* (Fifth Edition). Elsevier, London, pp. 811–815. <https://doi.org/10.1016/B978-0-444-52149-1.00071-9>
- Leal, J., Suárez, L., Jayabalan, R., Huerta, J., Escalante-Aburto, A., 2018. A review on health benefits of kombucha nutritional compounds and metabolites. *CyTA - J. Food* 16, 390–399. <https://doi.org/10.1080/19476337.2017.1410499>
- Lebleux, M., Abdo, H., Coelho, C., Basmaciyani, L., Albertin, W., Maupeu, J., Laurent, J., Roullier-Gall, C., Alexandre, H., Guilloux-Benatier, M., Weidmann, S., Rousseaux, S., 2020. New advances on the *Brettanomyces bruxellensis* biofilm mode of life. *Int. J. Food Microbiol.* 318, 108464. <https://doi.org/10.1016/j.ijfoodmicro.2019.108464>

- Ledormand, P., Desmasures, N., Dalmasso, M., 2021. Phage community involvement in fermented beverages: an open door to technological advances? *Crit. Rev. Food Sci. Nutr.* 61, 2911–2920. <https://doi.org/10.1080/10408398.2020.1790497>
- Leroy, F., De Vuyst, L., 2004. Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends Food Sci. Technol.* 15, 67–78. <https://doi.org/10.1016/j.tifs.2003.09.004>
- Li, L., Mendis, N., Trigui, H., Oliver, J.D., Faucher, S.P., 2014. The importance of the viable but non-culturable state in human bacterial pathogens. *Front. Microbiol.* 5.
- Liu, C.-H., Hsu, W.-H., Lee, F.-L., Liao, C.-C., 1996. The isolation and identification of microbes from a fermented tea beverage, Haipao, and their interactions during Haipao fermentation. *Food Microbiol.* 13, 407–415. <https://doi.org/10.1006/fmic.1996.0047>
- Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L., Law, M., 2012. Comparison of Next-Generation Sequencing Systems. *BioMed Res. Int.* 2012, e251364. <https://doi.org/10.1155/2012/251364>
- Liu, P.-Y., Wu, W.-K., Chen, C.-C., Panyod, S., Sheen, L.-Y., Wu, M.-S., 2020. Evaluation of Compatibility of 16S rRNA V3V4 and V4 Amplicon Libraries for Clinical Microbiome Profiling. <https://doi.org/10.1101/2020.08.18.256818>
- Lloyd, K.G., Steen, A.D., Ladau, J., Yin, J., Crosby, L., 2018. Phylogenetically Novel Uncultured Microbial Cells Dominate Earth Microbiomes. *mSystems* 3, e00055-18. <https://doi.org/10.1128/mSystems.00055-18>
- Louca, S., Doebeli, M., Parfrey, L.W., 2018. Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome* 6, 41. <https://doi.org/10.1186/s40168-018-0420-9>
- Magoc, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>
- Mahé, F., Rognes, T., Quince, C., Vargas, C. de, Dunthorn, M., 2015. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* 3, e1420. <https://doi.org/10.7717/peerj.1420>
- Malbaša, R., Lončar, E., Djurić, M., Došenović, I., 2008. Effect of sucrose concentration on the products of Kombucha fermentation on molasses. *Food Chem.* 108, 926–932. <https://doi.org/10.1016/j.foodchem.2007.11.069>
- Manuel, R.C., Esther, A.R., Guillermo, N.R., Mireles, R.L.D.C., Figueroa, L.Á., Orozco, J.J.I., Peña, E.R.D.L., 2014. Evaluation of the Properties of Healing of the Extract of Kombucha in Sheep in Growth with Malnutrition, Parasitocis and Respiratory Problems. *Open J. Vet. Med.* 4, 175–182. <https://doi.org/10.4236/ojvm.2014.48020>
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L.I., Jarvie, T.P., Jirage, K.B., Kim, J.-B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons,

- J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380. <https://doi.org/10.1038/nature03959>
- Marsh, A.J., O’Sullivan, O., Hill, C., Ross, R.P., Cotter, P.D., 2014. Sequence-based analysis of the bacterial and fungal compositions of multiple kombucha (tea fungus) samples. *Food Microbiol.* 38, 171–178. <https://doi.org/10.1016/j.fm.2013.09.003>
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17, 10–12. <https://doi.org/10.14806/ej.17.1.200>
- Matsuo, Y., Komiya, S., Yasumizu, Y., Yasuoka, Y., Mizushima, K., Takagi, T., Kryukov, K., Fukuda, A., Morimoto, Y., Naito, Y., Okada, H., Bono, H., Nakagawa, S., Hirota, K., 2021. Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution. *BMC Microbiol.* 21, 35. <https://doi.org/10.1186/s12866-021-02094-5>
- Matsushita, K., Toyama, H., Tonouchi, N., Okamoto-Kainuma, A. (Eds.), 2016. *Acetic Acid Bacteria*. Springer Japan, Tokyo. <https://doi.org/10.1007/978-4-431-55933-7>
- May, A., Narayanan, S., Alcock, J., Varsani, A., Maley, C., Aktipis, A., 2019. Kombucha: a novel model system for cooperation and conflict in a complex multi-species microbial ecosystem. *PeerJ* 7, e7565. <https://doi.org/10.7717/peerj.7565>
- May, A.N., Medina, J., Alcock, J., Maley, C., Aktipis, A., 2017. Kombucha as a model system for multispecies microbial cooperation: theoretical promise, methodological challenges and new solutions ‘in solution.’ <https://doi.org/10.1101/214478>
- McMurdie, P.J., Holmes, S., 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One* 8, e61217. <https://doi.org/10.1371/journal.pone.0061217>
- Miller CS, Handley KM, Wrighton KC, Frischkorn KR, Thomas BC, Banfield JF., 2013 Short-Read Assembly of Full-Length 16S Amplicons Reveals Bacterial Diversity in Subsurface Sediments. *PLoS ONE* 8(2): e56018. <https://doi.org/10.1371/journal.pone.0056018>
- Min, S., Kim, M.-J., Jeon, J., Kim, H.-Y., Han, E.S., 2022. Comparison of fermentation characteristics of kimchi made with fresh and stored spring kimchi cabbage. *Food Sci. Biotechnol.* 31, 221–229. <https://doi.org/10.1007/s10068-021-01019-2>
- Mishra, S., Singh, P.K., Pattnaik, R., Kumar, S., Ojha, S.K., Srichandan, H., Parhi, P.K., Jyothi, R.K., Sarangi, P.K., 2022. Biochemistry, Synthesis, and Applications of Bacterial Cellulose: A Review. *Front. Bioeng. Biotechnol.* 10.
- Mossion, A., Potin-Gautier, M., Delerue, S., Le Hécho, I., Behra, P., 2008. Effect of water composition on aluminium, calcium and organic carbon extraction in tea infusions. *Food Chem.*, 4th International Workshop on Water in Foods 106, 1467–1475. <https://doi.org/10.1016/j.foodchem.2007.05.098>

- Nagori, R., Purohit, S.D., 2012. An improved method for DNA isolation from polyphenols rich leaf samples of *Annona squamosa* L. *J. Plant Biochem. Biotechnol.* 21, 140–143. <https://doi.org/10.1007/s13562-011-0075-0>
- Nagy, L.G., Kocsubé, S., Csanádi, Z., Kovács, G.M., Petkovits, T., Vágvölgyi, C., Papp, T., 2012. Re-Mind the Gap! Insertion – Deletion Data Reveal Neglected Phylogenetic Potential of the Nuclear Ribosomal Internal Transcribed Spacer (ITS) of Fungi. *PLOS ONE* 7, e49794. <https://doi.org/10.1371/journal.pone.0049794>
- Niraj, K., Shardendu, 2020. Application of 16 Rrna Gene of V3-V4 Region for Meta Barcoding of Bacterial Community in High Density Population of Eastern India | Bioscience Biotechnology Research Communications. URL <https://bbrc.in/application-of-16-rrna-gene-of-v3-v4-region-for-meta-barcoding-of-bacterial-community-in-high-density-population-of-eastern-india/> (accessed 11.11.22).
- O'Donnell, K., Cigelnik, E., 1997. Two Divergent Intragenomic rDNA ITS2 Types within a Monophyletic Lineage of the Fungus *Fusarium* Are Nonorthologous. *Mol. Phylogenet. Evol.* 7, 103–116. <https://doi.org/10.1006/mpev.1996.0376>
- Ogier, J.-C., Pagès, S., Galan, M., Barret, M., Gaudriault, S., 2019. rpoB, a promising marker for analyzing the diversity of bacterial communities by amplicon sequencing. *BMC Microbiol.* 19, 171. <https://doi.org/10.1186/s12866-019-1546-z>
- Ojeda-Linares, C.I., Solís-García, I.A., Casas, A., 2022. Constructing Micro-Landscapes: Management and Selection Practices on Microbial Communities in a Traditional Fermented Beverage. *Front. Ecol. Evol.* 10.
- Oulas, A., Pavludi, C., Polymenakou, P., Pavlopoulos, G.A., Papanikolaou, N., Kotoulas, G., Arvanitidis, C., Iliopoulos, Ioannis, 2015. Metagenomics: Tools and Insights for Analyzing Next-Generation Sequencing Data Derived from Biodiversity Studies. *Bioinforma. Biol. Insights* 9, BBIS12462. <https://doi.org/10.4137/BBIS12462>
- Parente, E., Cogan, T.M., Powell, I.B., 2017. Chapter 8 - Starter Cultures: General Aspects, in: McSweeney, P.L.H., Fox, P.F., Cotter, P.D., Everett, D.W. (Eds.), *Cheese* (Fourth Edition). Academic Press, San Diego, pp. 201–226. <https://doi.org/10.1016/B978-0-12-417012-4.00008-9>
- Pelletier, E., Kreimeyer, A., Bocs, S., Rouy, Z., Gyapay, G., Chouari, R., Rivière, D., Ganesan, A., Daegelen, P., Sghir, A., Cohen, G.N., Médigue, C., Weissenbach, J., Le Paslier, D., 2008. “Candidatus *Cloacamonas Acidaminovorans*”: Genome Sequence Reconstruction Provides a First Glimpse of a New Bacterial Division. *J. Bacteriol.* 190, 2572–2579. <https://doi.org/10.1128/JB.01248-07>
- Peter, J., De Chiara, M., Friedrich, A., Yue, J.-X., Pflieger, D., Bergström, A., Sigwalt, A., Barre, B., Freel, K., Llored, A., Cruaud, C., Labadie, K., Aury, J.-M., Istace, B., Lebrigand, K., Barbry, P., Engelen, S., Lemainque, A., Wincker, P., Liti, G., Schacherer, J., 2018. Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* 556, 339–344. <https://doi.org/10.1038/s41586-018-0030-5>
- Petruzzello, Melissa, 2021. kombucha | Description, History, SCOBY, & Nutrition | Britannica [WWW Document]. URL <https://www.britannica.com/topic/kombucha> (accessed 11.11.22).

- Podolich, O., Zaets, I., Kukhareno, O., Orlovska, I., Reva, O., Khirunenko, L., Sosnin, M., Haidak, A., Shpylova, S., Rabbow, E., Skoryk, M., Kremenskoy, M., Demets, R., Kozyrovska, N., de Vera, J.-P., 2017. Kombucha Multimicrobial Community under Simulated Spaceflight and Martian Conditions. *Astrobiology* 17, 459–469. <https://doi.org/10.1089/ast.2016.1480>
- Pop, M., Walker, A.W., Paulson, J., Lindsay, B., Antonio, M., Hossain, M.A., Oundo, J., Tamboura, B., Mai, V., Astrovska, I., Bravo, H.C., Rance, R., Stares, M., Levine, M.M., Panchalingam, S., Kotloff, K., Ikumapayi, U.N., Ebruke, C., Adeyemi, M., Ahmed, D., Ahmed, F., Alam, M.T., Amin, R., Siddiqui, S., Ochieng, J.B., Ouma, E., Juma, J., Mailu, E., Omore, R., Morris, J.G., Breiman, R.F., Saha, D., Parkhill, J., Nataro, J.P., Stine, O.C., 2014. Diarrhea in young children from low-income countries leads to large-scale alterations in intestinal microbiota composition. *Genome Biol.* 15, R76. <https://doi.org/10.1186/gb-2014-15-6-r76>
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. <https://doi.org/10.1093/nar/gks1219>
- Ramadai, Abulreesh, H., 2010. Isolation and identification of yeast flora in local kombucha sample: Al Nabhah. *Umm Al-Qura Univ. J. Appl. Sci.* 2, 42–51.
- Reiss, J., 1994. Influence of different sugars on the metabolism of the tea fungus. *Z. Für Lebensm.-Unters. -Forsch.* 198, 258–261. <https://doi.org/10.1007/BF01192606>
- Reva, O.N., Zaets, I.E., Ovcharenko, L.P., Kukhareno, O.E., Shpylova, S.P., Podolich, O.V., de Vera, J.-P., Kozyrovska, N.O., 2015. Metabarcoding of the kombucha microbial community grown in different microenvironments. *AMB Express* 5, 35. <https://doi.org/10.1186/s13568-015-0124-5>
- Ricciardi, A., Ianniello, R.G., Tramutola, A., Parente, E., Zotta, T., 2014. Rapid detection assay for oxygen consumption in the *Lactobacillus casei* group. *Ann. Microbiol.* 64, 1861–1864. <https://doi.org/10.1007/s13213-014-0819-x>
- Rodríguez-Pérez, H., Ciuffreda, L., Flores, C., 2021. NanoCLUST: a species-level analysis of 16S rRNA nanopore sequencing data. *Bioinformatics* 37, 1600–1601. <https://doi.org/10.1093/bioinformatics/btaa900>
- Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584. <https://doi.org/10.7717/peerj.2584>
- Rothberg, J.M., Leamon, J.H., 2008. The development and impact of 454 sequencing. *Nat. Biotechnol.* 26, 1117–1124. <https://doi.org/10.1038/nbt1485>
- Saimaiti, A., Huang, S.-Y., Xiong, R.-G., Wu, S.-X., Zhou, D.-D., Yang, Z.-J., Luo, M., Gan, R.-Y., Li, H.-B., 2022. Antioxidant Capacities and Polyphenol Contents of Kombucha Beverages Based on Vine Tea and Sweet Tea. *Antioxid. Basel Switz.* 11, 1655. <https://doi.org/10.3390/antiox11091655>
- Santos, A., Mauro, M.S., Bravo, E., Marquina, D., 2009. PMKT2, a new killer toxin from *Pichia membranifaciens*, and its promising biotechnological properties for control of the spoilage yeast *Brettanomyces bruxellensis*. *Microbiology* 155, 624–634. <https://doi.org/10.1099/mic.0.023663-0>

- Savary, O., Mounier, J., Thierry, A., Poirier, E., Jourden, J., Maillard, M.-B., Penland, M., Decamps, C., Coton, E., Coton, M., 2021. Tailor-made microbial consortium for Kombucha fermentation: Microbiota-induced biochemical changes and biofilm formation. *Food Res. Int.* 147, 110549. <https://doi.org/10.1016/j.foodres.2021.110549>
- Schmeisser, C., Stöckigt, C., Raasch, C., Wingender, J., Timmis, K.N., Wenderoth, D.F., Flemming, H.-C., Liesegang, H., Schmitz, R.A., Jaeger, K.-E., Streit, W.R., 2003. Metagenome Survey of Biofilms in Drinking-Water Networks. *Appl. Environ. Microbiol.* 69, 7298–7309. <https://doi.org/10.1128/AEM.69.12.7298-7309.2003>
- Seifert, K.A., 2009. Progress towards DNA barcoding of fungi. *Mol. Ecol. Resour.* 9, 83–89. <https://doi.org/10.1111/j.1755-0998.2009.02635.x>
- Sharma, A., Lee, S., Park, Y.-S., 2020. Molecular typing tools for identifying and characterizing lactic acid bacteria: a review. *Food Sci. Biotechnol.* 29, 1301–1318. <https://doi.org/10.1007/s10068-020-00802-x>
- Shendure, J., Balasubramanian, S., Church, G.M., Gilbert, W., Rogers, J., Schloss, J.A., Waterston, R.H., 2017. DNA sequencing at 40: past, present and future. *Nature* 550, 345–353. <https://doi.org/10.1038/nature24286>
- Sievers, M., Lanini, C., Weber, A., Schuler-Schmid, U., Teuber, M., 1995. Microbiology and Fermentation Balance in a Kombucha Beverage Obtained from a Tea Fungus Fermentation. *Syst. Appl. Microbiol.* 18, 590–594. [https://doi.org/10.1016/S0723-2020\(11\)80420-0](https://doi.org/10.1016/S0723-2020(11)80420-0)
- Sipiczki, M., 2022. When barcoding fails: Genome chimerization (admixing) and reticulation obscure phylogenetic and taxonomic relationships. *Mol. Ecol. Resour.* 22, 1762–1785. <https://doi.org/10.1111/1755-0998.13586>
- Soares, M.G., de Lima, M., Reolon Schmidt, V.C., 2021. Technological aspects of kombucha, its applications and the symbiotic culture (SCOBY), and extraction of compounds of interest: A literature review. *Trends Food Sci. Technol.* 110, 539–550. <https://doi.org/10.1016/j.tifs.2021.02.017>
- Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J.M., Herndl, G.J., 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere.” *Proc. Natl. Acad. Sci.* 103, 12115–12120. <https://doi.org/10.1073/pnas.0605127103>
- Somerfield, P., 2008. Identification of the Bray-Curtis similarity index: Comment on Yoshioka (2008). *Mar. Ecol.-Prog. Ser. - MAR ECOL-PROGR SER* 372, 303–306. <https://doi.org/10.3354/meps07841>
- Steels, H., Bond, C.J., Collins, M.D., Roberts, I.N., Stratford, M., James, S.A., 1999. *Zygosaccharomyces lentus* sp. nov., a new member of the yeast genus *Zygosaccharomyces* Barker. *Int. J. Syst. Evol. Microbiol.* 49, 319–327. <https://doi.org/10.1099/00207713-49-1-319>
- Suárez, R., Suárez-Lepe, J.A., Morata, A., Calderón, F., 2007. The production of ethylphenols in wine by yeasts of the genera *Brettanomyces* and *Dekkera*: A review. *Food Chem.* 102, 10–21. <https://doi.org/10.1016/j.foodchem.2006.03.030>
- Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermet, T., Corthier, G., Brochmann, C., Willerslev, E., 2007. Power and limitations of the

- chloroplast trn L (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 35, e14. <https://doi.org/10.1093/nar/gk1938>
- Tadeusz, 2021. History of Kombucha - part 2. Baerbucha-Kombucha.
- Tadeusz, 2020. History of Kombucha - Part 1 [WWW Document]. Baerbucha-Kombucha. URL <https://www.baerbucha-kombucha.com/post/history-of-kombucha-part-1> (accessed 9.10.21).
- Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., Nishijima, M., 2014. Development of a Prokaryotic Universal Primer for Simultaneous Analysis of Bacteria and Archaea Using Next-Generation Sequencing. *PLOS ONE* 9, e105592. <https://doi.org/10.1371/journal.pone.0105592>
- Tartary, M., Vidaud, D., Piao, Y., Costa, J.-M., Bahnak, B.R., Fressinaud, E., Congard, B., Laurian, Y., Meyer, D., Lavergne, J.-M., Vidaud, M., 1993. Detection of a molecular defect in 40 of 44 patients with haemophilia B by PCR and denaturing gradient gel electrophoresis. *Br. J. Haematol.* 84, 662–669. <https://doi.org/10.1111/j.1365-2141.1993.tb03143.x>
- Taylor, D.L., Booth, M.G., Mcfarland, J.W., Herriott, I.C., Lennon, N.J., Nusbaum, C., Marr, T.G., 2008. Increasing ecological inference from high throughput sequencing of fungi in the environment through a tagging approach. *Mol. Ecol. Resour.* 8, 742–752. <https://doi.org/10.1111/j.1755-0998.2008.02094.x>
- Teoh, A.L., Heard, G., Cox, J., 2004. Yeast ecology of Kombucha fermentation. *Int. J. Food Microbiol.* 95, 119–126. <https://doi.org/10.1016/j.ijfoodmicro.2003.12.020>
- Tran, T., Billet, K., Torres-Cobos, B., Vichi, S., Verdier, F., Martin, A., Alexandre, H., Grandvalet, C., Tourdot-Maréchal, R., 2022a. Use of a Minimal Microbial Consortium to Determine the Origin of Kombucha Flavor. *Front. Microbiol.* 13, 836617. <https://doi.org/10.3389/fmicb.2022.836617>
- Tran, T., Grandvalet, C., Verdier, F., Martin, A., Alexandre, H., Tourdot-Maréchal, R., 2020a. Microbiological and technological parameters impacting the chemical composition and sensory quality of kombucha. *Compr. Rev. Food Sci. Food Saf.* 19, 2050–2070. <https://doi.org/10.1111/1541-4337.12574>
- Tran, T., Grandvalet, C., Verdier, F., Martin, A., Alexandre, H., Tourdot-Maréchal, R., 2020b. Microbial Dynamics between Yeasts and Acetic Acid Bacteria in Kombucha: Impacts on the Chemical Composition of the Beverage. *Foods* 9, 963. <https://doi.org/10.3390/foods9070963>
- Tran, T., Grandvalet, C., Winckler, P., Verdier, F., Martin, A., Alexandre, H., Tourdot-Maréchal, R., 2021. Shedding Light on the Formation and Structure of Kombucha Biofilm Using Two-Photon Fluorescence Microscopy. *Front. Microbiol.* 12, 725379. <https://doi.org/10.3389/fmicb.2021.725379>
- Tran, T., Roullier-Gall, C., Verdier, F., Martin, A., Schmitt-Kopplin, P., Alexandre, H., Grandvalet, C., Tourdot-Maréchal, R., 2022b. Microbial Interactions in Kombucha through the Lens of Metabolomics. *Metabolites* 12, 235. <https://doi.org/10.3390/metabo12030235>
- Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.-H., Smith, H.O., 2004.

- Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* 304, 66–74. <https://doi.org/10.1126/science.1093857>
- Verachtert, H., Derdelinckx, G., 2014. BELGIAN ACIDIC BEERS Daily Reminiscences of the Past. *Cerevisia* 38, 121–128. <https://doi.org/10.1016/j.cervis.2014.04.002>
- Villarreal-Soto, S.A., Beaufort, S., Bouajila, J., Souchard, J.-P., Taillandier, P., 2018. Understanding Kombucha Tea Fermentation: A Review. *J. Food Sci.* 83, 580–588. <https://doi.org/10.1111/1750-3841.14068>
- Villarreal-Soto, S.A., Bouajila, J., Pace, M., Leech, J., Cotter, P.D., Souchard, J.-P., Taillandier, P., Beaufort, S., 2020. Metabolome-microbiome signatures in the fermented beverage, Kombucha. *Int. J. Food Microbiol.* 333, 108778. <https://doi.org/10.1016/j.ijfoodmicro.2020.108778>
- von Gastrow, L., Michel, E., Legrand, J., Amelot, R., Segond, D., Guezenec, S., Rué, O., Chable, V., Goldringer, I., Dousset, X., Serpolay-Bessoni, E., Taupier-Letage, B., Vindras-Fouillet, C., Onno, B., Valence, F., Sicard, D., 2022. Microbial community dispersal from wheat grains to sourdoughs: A contribution of participatory research. *Mol. Ecol.*, special issue n/a. <https://doi.org/10.1111/mec.16630>
- Vu, N.T., Adel, H., Gupta, P., Schütze, H., 2016. Combining Recurrent and Convolutional Neural Networks for Relation Classification, in: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Presented at the NAACL-HLT 2016, Association for Computational Linguistics, San Diego, California, pp. 534–539. <https://doi.org/10.18653/v1/N16-1065>
- Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J.A., Jansson, J.K., Caporaso, J.G., Fuhrman, J.A., Apprill, A., Knight, R., 2015. Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* 1, e00009-15. <https://doi.org/10.1128/mSystems.00009-15>
- Wilmes, P., Simmons, S.L., Denef, V.J., Banfield, J.F., 2009. The dynamic genetic repertoire of microbial communities. *FEMS Microbiol. Rev.* 33, 109–132. <https://doi.org/10.1111/j.1574-6976.2008.00144.x>
- Wolfe, B.E., Dutton, R.J., 2015. Fermented foods as experimentally tractable microbial ecosystems. *Cell* 161, 49–55. <https://doi.org/10.1016/j.cell.2015.02.034>
- Xu, Y., Zhao, G.A., Wang, L.P., 2006. Controlled formation of volatile components in cider making using a combination of *Saccharomyces cerevisiae* and *Hanseniaspora valbyensis* yeast species. *J. Ind. Microbiol. Biotechnol.* 33, 192–196. <https://doi.org/10.1007/s10295-005-0051-6>
- Yamada, Y., Yukphan, P., Lan Vu, H.T., Muramatsu, Y., Ochaikul, D., Tanasupawat, S., Nakagawa, Y., 2012. Description of *Komagataeibacter* gen. nov., with proposals of new combinations (Acetobacteraceae). *J. Gen. Appl. Microbiol.* 58, 397–404. <https://doi.org/10.2323/jgam.58.397>
- Zabat, M.A., Sano, W.H., Cabral, D.J., Wurster, J.I., Belenky, P., 2018. The impact of vegan production on the kimchi microbiome. *Food Microbiol.* 74, 171–178. <https://doi.org/10.1016/j.fm.2018.04.001>
- Zampieri, A., Babbucci, M., Carraro, L., Milan, M., Fasolato, L., Cardazzo, B., 2021. Combining Culture-Dependent and Culture-Independent Methods: New

- Methodology Insight on the *Vibrio* Community of *Ruditapes philippinarum*. *Foods Basel Switz.* 10, 1271. <https://doi.org/10.3390/foods10061271>
- Zhang, W., Fan, X., Zhu, S., Zhao, H., Fu, L., 2013. Species-Specific Identification from Incomplete Sampling: Applying DNA Barcodes to Monitoring Invasive *Solanum* Plants. *PLOS ONE* 8, e55927. <https://doi.org/10.1371/journal.pone.0055927>
- Zheng, J., Wittouck, S., Salvetti, E., Franz, C.M.A.P., Harris, H.M.B., Mattarelli, P., O'Toole, P.W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G.E., Gänzle, M.G., Lebeer, S., 2020. A taxonomic note on the genus *Lactobacillus*: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol.* 70, 2782–2858. <https://doi.org/10.1099/ijsem.0.004107>
- Zubaidah, E., Dewantari, F.J., Novitasari, F.R., Srianta, I., Blanc, P.J., 2018. Potential of snake fruit (*Salacca zalacca* (Gaerth.) Voss) for the development of a beverage through fermentation with the *Kombucha* consortium. *Biocatal. Agric. Biotechnol.* 13, 198–203. <https://doi.org/10.1016/j.cbab.2017.12.012>

# **Annexes**



**Publication en cours de rédaction :**

Diversity in the microbial communities involved in kombucha fermentations performed with SCOBY from different origins.

Ben Saâd E, Bleykasten-Grosshans C, Friedrich A, Schacherer J.

Evolution in the microbial composition of kombucha SCOBYs along successive back-slopping

Ben Saâd E, Friedrich A, Schacherer J, Bleykasten-Grosshans C.

**Communications scientifiques :**

**Séminaire de Microbiologie de Strasbourg** (10e édition), 2022, Strasbourg (France)

24/03/2022, communication orale

Deciphering the microbial communities involved in the kombucha fermentation by metabarcoding.

Ben Saâd E, Friedrich A, Schacherer J, Bleykasten-Grosshans C.

**International congress on yeasts (ICY15)**, 2021, Vienne (Autriche)

23\_27 /08/2021, communication par affiche

Deciphering the microbial communities involved in the kombucha fermentation.

Ben Saâd E, Friedrich A, Schacherer J, Bleykasten-Grosshans C.

## **Activités chez le producteur de kombucha BioBrasseurs**

En 2016, j'ai participé à l'installation de l'unité de production de kombucha de la société BioBrasseurs. Un lien a été gardé pour permettre de valoriser mes compétences acquises au sein de l'entreprise et de réfléchir aux futurs possibilités de développements. C'est dans ce contexte que ma convention de thèse CIFRE a été mise en place, avec une répartition du temps de travail estimée à 70 % du temps au sein de l'équipe Haploteam et 30 % de temps au sein de l'entreprise. Ainsi, pendant mes 3.5 ans de thèse, une journée par semaine a été dédiée à mes activités chez BioBrasseurs. Mon activité principale est de veiller au bon déroulement des fermentations, dans des conditions appliquant les bonnes pratiques de fabrication et d'hygiène. J'ai mis en place le système de contrôle en interne des paramètres physico-chimiques (mesures du pH, du degré de Brix et du taux d'éthanol). J'ai défini la fréquence des prélèvements et j'ai mis en place le système documentaire qui assure la traçabilité de la production et la mise en place des actions correctives face à des non-conformités. Et j'ai formé les agents Brasseurs à l'obligation du respect des règles d'hygiène en manipulant le SCOBY et l'importance de la séparation des SCOBY pour grader la traçabilité

BioBrasseurs doit faire appel à des laboratoires externes accrédités, afin de valider ces contrôles réalisés en interne et d'assurer la conformité et la sécurité des produits finis commercialisés (conformité législative d'une boisson nonalcoolisée avec une teneur en éthanol inférieure à 1.2 % v/v ; recherche des pathogènes alimentaires). Les analyses des prélèvements faits sur la ligne de production pour tester l'efficacité du système de nettoyage et de désinfection, ainsi la détermination des valeurs nutritionnelles pour les nouvelles recettes sont également traitées en externe. Pour mener ces analyses, j'assure la gestion et l'envoi des prélèvements, l'archivage des résultats .

Le système d'archivage et de suivi documentaire que j'ai mis en place est indispensable lors des audits, dont ceux réalisés par l'organisme certificateur ECOCERT passant notamment un test de traçabilité ascendante et descendante.

Mon implication dans le démarrage de la brasserie m'a permis de participer à la résolution des problématiques liées à des outils et matériels de production en collaboration avec l'équipe de production gérée par Mr COUROT. Mon rôle est de chercher les causes potentielles, de proposer un plan d'action et de mettre en place un plan de contrôle.

Dans le cadre de travaux de recherche pour la thèse, afin d'étudier la composition des communautés microbiennes sur le site de production, j'ai mis en place une démarche de prélèvement systématique à partir des différentes cuves de fermentation. L'agent brasseur a été formé sur l'importance de ces prélèvements et de la modalité d'échantillonnage pour qu'ils soient les plus représentatifs de la communauté impliquée dans le brassin. Les outils de prélèvement ont été mis à disposition (des bocaux de 400 ml nettoyés et couverts, des pipettes, des tubes falcons pour le transfert). L'étiquetage et la formalisation de tous les renseignements ont été assurés pour la traçabilité et pour la disponibilité des données lors des analyses.

## Résumé

Elucider les interactions entre les espèces d'une niche écologique est une thématique majeure en biologie. Grâce aux avancées technologiques des méthodes de séquençage haut débit, cette thématique connaît un essor considérable dans le contexte des communautés microbiennes. Elles permettent d'identifier les taxons présents, de connaître leurs abondances et de mesurer l'impact des conditions environnementales, avec une très grande précision. Les communautés microbiennes appelées SCOBY (Symbiotic Culture of Bacteria and Yeast) ont représenté les modèles d'étude de ma thèse. Ces communautés de levures et de bactéries, dispersées entre une phase liquide et un biofilm de cellulose bactérienne, permettent d'élaborer le kombucha, une boisson fermentée ancestrale, pétillante et acidulée. Dans un premier temps, la caractérisation par métabarcoding d'une collection de 18 SCOBY d'origines variées a constitué une ressource pour décrire la diversité taxonomique de ces communautés. Dans un second temps, des séries temporelles, au laboratoire et en brasserie, ont permis de caractériser leur stabilité au cours des fermentations successives.

**Mots-clé :** fermentation, kombucha, métagénomique, diversité taxonomique, SCOBY, microbiologie.

## Abstract

Elucidating the interactions between the species of an ecological niche is a major theme in biology. Thanks to the recent advances in high-throughput sequencing methods, this theme is experiencing considerable growth in the context of microbial communities. They make it possible to identify the taxa present, to assess their abundance, and to measure the impact of environmental conditions, with great precision. The microbial communities called SCOBY (Symbiotic Culture of Bacteria and Yeast) represented the models of my thesis. These communities of yeasts and bacteria, dispersed between a liquid phase and a biofilm of bacterial cellulose, are involved in kombucha production, an ancestral fermented soft drink. First, the characterization by metabarcoding of a collection of 18 SCOBYs of various origins constituted a resource to describe the taxonomic diversity of these communities. Secondly, the study of time series, in the laboratory and in the brewery, made it possible to characterize their stability during successive fermentations.

**Keywords:** fermentation, kombucha, metagenomics, taxonomic diversity, SCOBY, microbiology.