



**HAL**  
open science

# Analyse de la communauté bactérienne et de la diversité inter et intra spécifique des bactéries acétiques et lactiques impliquées dans la fermentation de cacao selon trois origines géographiques

Lucie Farrera

## ► To cite this version:

Lucie Farrera. Analyse de la communauté bactérienne et de la diversité inter et intra spécifique des bactéries acétiques et lactiques impliquées dans la fermentation de cacao selon trois origines géographiques. Médecine humaine et pathologie. Université Montpellier, 2019. Français. NNT : 2019MONTG091 . tel-02866731v2

**HAL Id: tel-02866731**

**<https://theses.hal.science/tel-02866731v2>**

Submitted on 15 Jun 2020

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

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

# THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Biotechnologie, Microbiologie

École doctorale GAIA

Unité mixte de recherche QualiSud- CIRAD

**Analyse de la communauté bactérienne et de la diversité  
inter et intra spécifique des bactéries acétiques et  
lactiques impliquées dans la fermentation de cacao selon  
trois origines géographiques.**

Présentée par Lucie FARRERA

Le 19 décembre 2019

Sous la direction de Corinne TEYSSIER et Didier MONTET

## THÈSE POUR OBTENIR LE GRADE DE DOCTEUR

Devant le jury composé de

Jean-Jacques GODON, Président de jury, Directeur de recherche

Sévastianos ROUSSOS, Membre du jury, Directeur de recherche émérite

Rémy CACHON, Rapporteur, Professeur des universités

Didier MONTET, Co-Directeur de thèse, Cadre scientifique des EPIC

Béatrice SEGURENS, Rapporteur, Cadre scientifique des EPIC

Corinne TEYSSIER, Directeur de thèse, Maître de conférences



UNIVERSITÉ  
DE MONTPELLIER

## Remerciements

Au cours de ces trois années de thèse les rencontres ont été nombreuses et enrichissantes. Je tiens à remercier toutes les personnes dont j'ai croisé le chemin, qui ont contribué de près ou de loin au bon déroulement de la thèse.

J'adresse mes remerciements aux membres du jury d'avoir accepté de juger ce travail, à Mr Rémy Cachon et Mme Ségurens Béatrice d'avoir accepté d'en être les rapporteurs. Et merci à Mr Godon et Mr Roussos d'être présent en tant qu'examineurs.

Je tiens tout d'abord à remercier mes directeurs de thèse Corinne Teyssier et Didier Montet, pour m'avoir permis de réaliser cette thèse et de leur patience dans les corrections.

De manière générale je tiens à remercier les membres de l'équipe 2, « Maitrise des contaminants de la chaîne alimentaire » que j'ai eu la chance de rencontrer : Angélique Fontana, Nathalie Paillusson, Catherine Brabet, Isabelle Piro-Metayer...

Je souhaite remercier Florentin Constancias, ancien membre de l'équipe, pour son aide dans l'élaboration de la campagne de prélèvements au Mexique et le traitement des données de metabarcoding.

Je tiens à remercier plus particulièrement Alexandre Colas de La Noue et Caroline Strub, pour leur aide dans les manipes de biochimie, dans l'interprétation des données et les corrections, mais surtout pour votre soutien, merci à vous deux.

Je tiens également à remercier Benjamin Guibert, membre de l'équipe 2, de m'avoir soutenu ou plutôt supporté tout au long de la thèse, merci pour tes conseils, merci pour ton soutien infaillible et ce même à 6 h du matin ! Nos petites pauses clopes/chocolat vont beaucoup me manquer...

Je remercie également Dr Mirna Suarez Quiroz et toute son équipe à l'institut technologique de Veracruz au Mexique pour leur accueil chaleureux au sein de leur laboratoire et leur aide précieuse dans l'élaboration de prélèvements d'ADN et de fèves de cacao fraîches.

Je remercie également la fondation DUFRENOY et le crédit Agricole d'Ile de France pour m'avoir permis de financer ce séjour au Mexique.

Ensuite, je souhaite remercier tous les stagiaires que j'ai pu rencontrer, la team Guayule (Flo, Kevin, Ilona, Nicolas), Ines, Charles, Anais, Victoria, Jacques... mais surtout Thomas et Ismérie devenus de vrais amis maintenant. Je souhaite remercier la team Ivoirienne, pour leur joie de vivre, leurs bons repas, leurs sourires : Didier-Ange, Michel Olivier, Antonin, Corinne, Christian, Joelle, Richard...

Je souhaite remercier toutes ces personnes autour de moi qui m'ont soutenu tout du long : Camille, la team des affamés : Rémy, Nicolas, Laura, Laure-Anne, Eric, merci de votre soutien les amis <3.

Je souhaite également remercier mon premier collègue de bureau et ami : Abderahim.

Merci à mon binôme du CIRAD : Sophie, pour ta bonne humeur, ta joie de vivre et ton sourire qui illuminait mes journées.

Mes remerciements vont plus particulièrement à mes deux amis, mes deux camarades, devenus ma seconde famille : Magali & Moez. Votre amitié est la plus belle chose qui me soit arrivée au cours de cette thèse, je vous remercie d'être de si belles personnes.

Je souhaite également remercier tous mes amis qui m'ont soutenu : Aurélien, Aly, Noor, Thomas, Lucas Ellois... tous les membres de la team salsa et ceux que j'oublie. Je remercie mes colocataires de leur soutien et leurs conseils des plus précieux.

Finalement, mes remerciements vont à ma famille pour leur amour et leur soutien qui m'ont été des plus vitaux vers la fin de la thèse. Merci de votre fierté, merci de votre amour <3.

# Tables des matières

Tables des matières.....	4
Listes des abbréviations .....	8
Liste des tableaux.....	10
Liste des figures.....	12
Introduction.....	16
Chapitre I : Etat de l'art .....	19
<b>I.1 Généralités sur le cacao</b> .....	20
<b>I.1.1 Production mondiale de cacao</b> .....	20
<b>I.1.2 Le cacaoyer <i>Theobroma cacao</i> L.</b> .....	22
<i>I.1.2.1 Généralités sur le cacaoyer</i> .....	22
<i>I.1.2.2 Variétés de cacaoyer</i> .....	24
<i>I.1.2.3 Composition chimique de la pulpe mucilagineuse et des fèves de cacao</i> .....	25
<b>I.2 De la récolte au cacao marchand</b> .....	27
<b>I.2.1 De la récolte à l'écabossage des cabosses</b> .....	27
<b>I.2.2 Les traitements post récolte</b> .....	29
<i>I.2.2.1 La fermentation</i> .....	29
<i>I.2.2.2 le séchage et le stockage des fèves</i> .....	31
<b>I.3 Composition des communautés microbiennes au cours du processus de fermentation du cacao</b> .....	32
<b>I.3.1 Méthodes d'analyse des communautés microbiennes</b> .....	33
<i>I.3.1.1 Marqueurs moléculaires universels</i> .....	33
<i>I.3.1.2 Techniques d'empreinte moléculaire</i> .....	35
<i>I.3.1.3 La métagénomique</i> .....	37
<i>I.3.1.4 Les nouvelles techniques de séquençage (NGS)</i> .....	38
<i>I.3.1.5 Les techniques de séquençage de troisième génération</i> .....	41
<b>I.3.3 Succession et rôle des microorganismes au cours de la fermentation</b> .....	45
<b>I.4 Les bactéries lactiques et acétiques : principales actrices de la fermentation du cacao</b> .....	50
<b>I.4.1 Méthodes d'analyse de la diversité intra spécifique</b> .....	51
<i>I.4.1.1 Notions d'espèce et de souche</i> .....	51
<i>I.4.1.2 Méthodes d'analyse de la diversité intra spécifique</i> .....	52
I.4.1.2.1 Comparaison de profils génomiques.....	53
<i>Méthodes basées sur la digestion enzymatique</i> .....	54
I.4.1.2.2 Méthodes basées sur le séquençage d'ADN .....	56
<b>I.4.2 Les bactéries lactiques (BAL)</b> .....	57

<b>I.4.3 Les bactéries acétiques (BAA)</b> .....	63
<i>I.4.3.1 Taxonomie</i> .....	63
<i>I.4.3.2 Isolement des bactéries acétiques</i> .....	63
<i>I.4.3.3 Les bactéries acétiques impliquées dans la fermentation du cacao</i> .....	64
<i>I.4.3.4 Métabolisme des bactéries acétiques</i> .....	68
<i>I.4.3.5 Organisation génomique des Acetobacter pasteurianus</i> .....	71
<b>I.4.4 Starter de fermentations</b> .....	72
<b>I.5 Conclusion</b> .....	74
<b>II.1 Echantillonnage des fèves de cacao</b> .....	77
<b>II.1.1 Fèves de cacao séchées</b> .....	78
<b>II.1.2 Fèves de cacao fraîches avec la pulpe</b> .....	80
<b>II.2 Echantillonnage environnemental</b> .....	83
<b>II.3 Analyse des communautés microbiennes associées aux fèves en fermentation et à l'environnement de la fermentation</b> .....	85
<b>II.3.1 Extraction, mesure de la qualité et de la quantité d'ADN total</b> .....	85
<i>II.3.1.1 Extraction des ADN par la méthode de « phénol/chloroforme »</i> .....	85
<i>II.3.1.2 Détermination de la quantité et de la qualité des ADN</i> .....	87
<b>II.3.2 Analyse des communautés bactériennes associées aux fèves de cacao par PCR-DGGE</b> .....	88
<i>II.3.2.1 Amplification partielle du gène de l'ARNr 16S par PCR</i> .....	88
<i>II.3.2.2 Electrophorèse en gel d'acrylamide avec gradient dénaturant (DGGE)</i> .....	89
<i>II.3.2.3 Traitement des images</i> .....	90
<i>II.3.2.4 Identification des bandes</i> .....	91
<b>II.3.3 Analyse des communautés bactériennes par metabarcoding (Miseq, Illumina) sur le gène de l'ARNr 16S</b> .....	91
<i>II.3.3.1 Amplification partielle par PCR du gène de l'ARNr 16S (PCR1)</i> .....	92
<i>II.3.3.2 Préparation de la banque d'amplicons : purification, PCR d'indexation (PCR2) et séquençage</i> .....	93
<i>II.3.3.3 Analyse des données brutes</i> .....	95
<b>II.3.4 Analyses statistiques</b> .....	96
<b>II.4 Isolement de souches de BAL et BAA impliquées dans la fermentation du cacao</b> .....	97
<b>II.4.1 Protocole d'isolement</b> .....	97
<b>II.4.2 Caractérisations phénotypiques des isolats</b> .....	99
<b>II.4.3 Identification moléculaire des isolats au niveau de l'espèce</b> .....	100
<i>II.4.3.1 Extraction de l'ADN par choc thermique</i> .....	100
<i>II.4.3.2 Pré-identification des BAL par PCR</i> .....	100
<i>II.4.3.3 Identification des BAL et BAA par amplification et séquençage du gène de l'ARNr 16S</i> .....	101

<b>II.5 Analyse de la diversité intra spécifique des BAA isolées lors des fermentations de cacao.....</b>	<b>102</b>
<b>II.5.1 Analyse du polymorphisme génomique : PCR sur séquences répétées.....</b>	<b>102</b>
<i>II.5.1.1 PCR sur séquences répétées : ERIC-PCR, REP-PCR, Box-PCR, (GTG)<sub>5</sub>-PCR.....</i>	<i>102</i>
<i>II.5.1.2 PCR Multiplexe sur séquences répétées (Box1AR, ERIC, REP).....</i>	<i>103</i>
<i>II.5.1.3 Analyse des profils génomiques .....</i>	<i>103</i>
<b>II.5.2 Caractérisation biochimique des souches.....</b>	<b>104</b>
<i>II.5.2.1 Suivi de cinétique de croissance à l'aide d'un spectrofluorimètre .....</i>	<i>104</i>
<i>II.5.2.2 Analyse des substrats consommés et du produit formé .....</i>	<i>105</i>
<i>II.5.2.3 Analyse statistique .....</i>	<i>106</i>
<b>Chapitre III : Résultats et discussion.....</b>	<b>107</b>
<b>III.1 Influence de l'origine géographique et de la période de récolte sur la structure des communautés bactériennes impliquées dans la fermentation de fèves de cacao.....</b>	<b>108</b>
<b>III.1.1 Etude comparative des communautés bactériennes des fèves de cacao en fermentation en fonction de leurs origines géographiques.....</b>	<b>109</b>
<i>III.1.1.1 Comparaison des empreintes globales bactériennes en fonction de 2 pays.....</i>	<i>109</i>
<i>III.1.1.1.1 Comparaison de la diversité <math>\alpha</math> des échantillons de fèves de cacao en fermentation en provenance de 2 pays .....</i>	<i>111</i>
<i>III.1.1.1.2 Analyse de la diversité <math>\beta</math> des échantillons de fèves de cacao en fermentation en provenance de 2 pays .....</i>	<i>112</i>
<i>III.1.1.1.3 Identification des OTUs bactériennes .....</i>	<i>115</i>
<i>III.1.1.1.4 Conclusion partielle .....</i>	<i>117</i>
<i>III.1.1.2 Analyse par metabarcoding des communautés bactériennes associées aux fèves de cacao en fermentation de 3 pays .....</i>	<i>118</i>
<i>III.1.1.2.1 Comparaison de la diversité <math>\alpha</math> des échantillons de fèves de cacao en fermentation en provenance de 3 pays .....</i>	<i>119</i>
<i>III.1.1.2.2 Analyse de la diversité <math>\beta</math> des échantillons de fèves de cacao en fermentation en provenance de 3 pays .....</i>	<i>121</i>
<i>III.1.1.2.3 Comparaison de la composition des communautés bactériennes associées aux échantillons de fèves de cacao en fermentation en provenance de 3 pays.....</i>	<i>125</i>
<b>III.1.2 Etude comparative des communautés bactériennes des fèves de cacao en fermentation en fonction de la grande et de la petite récolte au Mexique.....</b>	<b>130</b>
<i>III.1.2.1 Comparaison de la diversité <math>\alpha</math> des échantillons des fèves de cacao en fermentation en fonction de la grande et de la petite récolte au Mexique.....</i>	<i>130</i>
<i>III.1.2.2 Analyse de la diversité <math>\beta</math> des échantillons des fèves de cacao en fermentation en fonction de la grande et de la petite récolte au Mexique.....</i>	<i>132</i>
<i>III.1.2.3 Comparaison de la composition des communautés bactériennes des échantillons des fèves de cacao en fermentation en fonction de la grande et de la petite récolte au Mexique .</i>	<i>135</i>
<b>III.1.3 Conclusion partielle .....</b>	<b>138</b>

III.2 Détermination de l'origine du pool de bactéries impliqué dans les fermentations de fèves de cacao : au niveau de la plantation ou de l'environnement de la fermentation du Mexique.....	141
III.2.1 Comparaison de la diversité $\alpha$ des prélèvements environnementaux et des fèves de cacao en fermentation.....	142
III.2.2 Analyse de la diversité $\beta$ des prélèvements environnementaux et des fèves de cacao en fermentation.....	146
III.2.3 Analyse de la composition des communautés bactériennes des fèves de cacao en fermentation et des surfaces de l'environnement de la fermentation.....	150
III.2.4 Conclusion partielle .....	154
III.3– Diversité intraspécifique de souches de bactéries acétiques et lactiques en fonction de leurs origines géographiques.....	155
III.3.1 Elaboration d'une collection de souches de BAA et BAL.....	156
III.3.1.1 Stratégies d'échantillonnage et de conservation des échantillons .....	156
III.3.1.2 Isolement des BAA et des BAL .....	157
III.3.1.3 Identification des BAA et BAL .....	159
III.3.1.4 Conclusion partielle .....	162
III.3.2 Analyse de la diversité intraspécifique des souches d' <i>A. pasteurianus</i> .....	164
III.3.2.1 Analyse du polymorphisme génomique.....	164
III.3.2.1.1 Mise au point de la méthode.....	164
III.3.2.1.2 Analyse du polymorphisme génomique à l'aide des amorces (GTG) <sub>5</sub> .....	170
III.3.2.2 Etude du comportement des souches d' <i>A. pasteurianus</i> dans le milieu simulant la pulpe de cacao.....	174
III.3.2.2.1 Détermination du taux de croissance spécifique des souches .....	174
III.3.2.2.2 Courbes d'étalonnage et analyse des chromatogrammes à l'HPLC .....	177
III.3.2.2.3 Analyse comparative des souches et regroupement en dendrogramme.....	178
III.3.2.3 Conclusion partielle .....	188
III.3.3 Analyse du polymorphisme génomique des souches d' <i>A. tropicalis</i> .....	189
III.3.4 Conclusion partielle .....	191
Conclusion et perspectives.....	194
Références bibliographiques.....	203
Annexes .....	215



# Listes des abréviations

<b>ACP</b>	Analyse en Composantes Principales
<b>ADH</b>	Alcool Déshydrogénase
<b>ALDH</b>	Aldéhyde Déshydrogénase
<b>ANOVA</b>	Analyse de la Variance
<b>BAA</b>	Bactérie acétique
<b>BAL</b>	Bactérie lactique
<b>CaCO<sub>3</sub></b>	Carbonate de Calcium
<b>CI</b>	Côte d'Ivoire
<b>CIRAD</b>	Centre de International de Recherche Agronomique et pour le Développement
<b>CPSM</b>	Cocoa Pulp Simulation Medium
<b>DGGE</b>	Denaturing Gradient Gel Electrophoresis
<b>DO</b>	Densité optique
<b>EDTA</b>	Ethylenediamine tetra acetic acid
<b>ERIC</b>	Enterobacterial Repetitive Intergenic Consensus
<b>GF</b>	Guyane
<b>HPLC</b>	High Performance liquid chromatography
<b>ICCO</b>	International Cocoa Organization
<b>MLST</b>	Multi Locus Sequencing Typing
<b>MRS</b>	Man Rogosa Sharp
<b>MX</b>	Mexique
<b>NCBI</b>	National Center for Biotechnology Information
<b>NDMS</b>	Non-metric Multidimensional Scaling
<b>NGS</b>	Next generation sequencing
<b>OTU</b>	Operational Taxonomic Unit
<b>R<sub>AL/ETOH</sub></b>	Ratio de la consommation d'acide lactique sur celle de l'éthanol
<b>REP</b>	Repetitive Element-Palindromic
<b>SDS</b>	Sodium dodecyl sulphate
<b>TAE</b>	Tris / Ammonium / EDTA
<b>TCA</b>	Tricarboxylic acid cycle

**TE** Tris / EDTA

**TEMED** Tétraméthyléthylènediamine

**UGPMA** Unweighted pair group method with arithmetic mean

**UV** Ultra Violet

**VNBC** Viable but non-culturable

**WGS** Whole genome sequencing

**$Y_{AA/ETOH}$**  Efficacité de conversion de l'éthanol en acide acétique

**$Y_{X/AA}$**  Rendement en biomasse produite par mole d'acide acétique produite

**$Y_{X/S}$**  Rendement en biomasse à partir des substrats

# Liste des tableaux

<b>Tableau 1:</b> Composition de la pulpe de cacao (en g pour 100 g de matière fraîche). .....	26
<b>Tableau 2:</b> Inventaire de la diversité des principales espèces de levures isolées lors de fermentation de cacao selon différents pays producteurs. ....	47
<b>Tableau 3 :</b> Distribution de la fréquence des principales espèces de BAL présentes dans les fèves de cacao fermentées selon différents pays producteurs.....	60
<b>Tableau 4:</b> Distribution de la fréquence des principales espèces de BAA présentes dans les fèves de cacao fermentées selon différents pays producteurs.....	65
<b>Tableau 5:</b> Liste détaillée des échantillons de fèves séchées. ....	78
<b>Tableau 6:</b> Liste détaillée des échantillons de fèves fraîches avec pulpe. ....	80
<b>Tableau 7:</b> Liste et code des prélèvements environnementaux réalisés au Mexique en Octobre 2018. ....	85
<b>Tableau 8:</b> Composition du mélange réactionnel pour l'amplification partielle du gène de l'ARNr 16S par PCR. ....	89
<b>Tableau 9:</b> Composition des solutions d'agents dénaturants.....	89
<b>Tableau 10:</b> Liste des amorces utilisées pour l'amplification partielle du gène de l'ARNr 16S. ....	92
<b>Tableau 11:</b> Composition du mélange réactionnel pour l'amplification partielle du gène de l'ARNr 16S par PCR. ....	93
<b>Tableau 12:</b> Composition des milieux de culture liquides utilisés pour la croissance des BAL et des BAA. ....	98
<b>Tableau 13:</b> Composition du milieu de culture gélosé pour l'isolement des BAA.....	99
<b>Tableau 14 :</b> Composition du volume réactionnel de la PCR pour la pré-identification des BAL. ....	101
<b>Tableau 15 :</b> Listes des amorces utilisées pour l'analyse par PCR sur séquences répétées (Versalovic et al. 1994). ....	102
<b>Tableau 16 :</b> Composition du mélange réactionnel pour l'analyse par PCR sur séquences répétées.....	103
<b>Tableau 17 :</b> Identification par séquençage de la région V3 du gène de l'ARNr 16S des bandes d'ADN bactériens des profils DGGE communes aux deux pays (Mexique et Guyane). ....	116
<b>Tableau 18 :</b> Résultats du test statistique PERMANOVA sur la matrice de distance réalisée avec les indices de dissimilarité de Bray-Curtis obtenus pour les échantillons de fèves de cacao de 3 pays, Mexique, Côte d'Ivoire et Guyane à différents temps de fermentation.....	123
<b>Tableau 19 :</b> Résultats du test statistique PERMANOVA sur la matrice de distance réalisée avec les indices de similarité de Jaccard obtenus pour les échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane à différents temps de fermentation.....	125
<b>Tableau 20 :</b> Résultats du test statistique PERMANOVA sur la matrice de distance réalisée avec les indices de dissimilarité de Bray-Curtis des échantillons de fèves de cacao à différents temps de fermentation en provenance du Mexique, grande et petite récolte.....	133
<b>Tableau 21 :</b> Résultats du test statistique PERMANOVA sur la matrice de distance réalisée avec les indices de dissimilarité de Jaccard des échantillons de fèves de cacao à différents temps de fermentation en provenance du Mexique, grande et petite récolte.....	135
<b>Tableau 22 :</b> Résultats du test statistique PERMANOVA sur la matrice de distance réalisées avec les indices de Bray-Curtis pour les échantillons de fèves fraîches à différents temps de fermentation et des prélèvements environnementaux en provenance du Mexique.....	148
<b>Tableau 23:</b> Résultats du test statistique PERMANOVA sur la matrice de distance réalisées avec les indices de Jaccard pour les échantillons de fèves fraîches à différents temps de fermentation et des prélèvements environnementaux en provenance du Mexique.....	149

<b>Tableau 24</b> : Mesure de l'activité de l'eau des fèves de cacao séchées en provenance de Côte d'Ivoire, du Pérou et du Nicaragua.....	158
<b>Tableau 25</b> : Répartition des 149 isolats de BAA selon les trois pays (Côte d'Ivoire, Guyane et Mexique). .....	160
<b>Tableau 26</b> : Répartition des 63 isolats de BAL selon les trois pays (Côte d'Ivoire, Guyane française et Mexique). .....	162
<b>Tableau 27</b> : Comparaison de la quantité d'ADN extraite de 6 souches d' <i>A. pasteurianus</i> et la souche type LMG12062 <sup>T</sup> d' <i>A.pasteurianus</i> par les deux méthodes d'extraction : par choc thermique et avec le kit FastDNA spin .....	168
<b>Tableau 28</b> : Répartition des souches dans les deux grands groupes du dendrogramme et de la heatmap <b>Figure 66</b> .....	181

# Liste des figures

<b>Figure 1:</b> Carte des principaux pays producteurs de cacao situés dans la zone équatoriale.....	20
<b>Figure 2:</b> Production de cacao en millions de tonnes des 10 premiers pays producteurs pour la saison 2017/2018 (ICCO 2019).....	21
<b>Figure 3:</b> Evolution morphologique de la fleur à la cabosse.....	23
<b>Figure 4:</b> Cabosses des trois principales variétés de cacaoyer utilisées dans la production mondiale de cacao.....	24
<b>Figure 5:</b> Coupe transversale d'une cabosse de cacao montrant les fèves entourées de la pulpe à l'intérieur.....	26
<b>Figure 6:</b> Ecabossage des cabosses à l'aide d'une machette au Mexique, Octobre 2018.....	28
<b>Figure 7:</b> Extraction manuelle des fèves de cacao au Mexique, Octobre 2018.....	28
<b>Figure 8:</b> Caisses de fermentation en bois utilisées au Mexique.....	29
<b>Figure 9:</b> Fermentation en tas sur des feuilles de bananiers en Côte d'Ivoire.....	29
<b>Figure 10:</b> Dynamique des communautés microbiennes impliquées au cours de la fermentation des fèves de cacao (De Vuyst et Weckx 2016).....	30
<b>Figure 11:</b> Séchage solaire des fèves de cacao après fermentation.....	31
<b>Figure 12:</b> Stockage des fèves avant l'envoi aux industriels pour transformation.....	32
<b>Figure 13:</b> Cinétique de la dégradation et de la production des métabolites pendant la fermentation de fèves de cacao (De Vuyst et Weckx 2016).....	32
<b>Figure 14:</b> Organisation des gènes des ARN ribosomiques chez les Eucaryotes (A) et chez les Procaryotes (B) (White et al. 1990).....	34
<b>Figure 15:</b> le coût du séquençage d'un génome humain (3,2 milliards de pb).....	38
<b>Figure 16:</b> Principe de l'approche Illumina (Bonnin et al. 2017).....	40
<b>Figure 17:</b> Principe de la technologie Oxford Nanopore (Montel 2018).....	42
<b>Figure 18:</b> Principe du SMRT séquençage utilisé dans les technologies Pacific biosciences.....	43
<b>Figure 19:</b> Classification des méthodes de génotypage de souches bactériennes basées sur l'obtention de profils d'ADN (Li <i>et al.</i> 2009).....	53
<b>Figure 20:</b> Classification des méthodes de génotypage basées sur les séquences d'ADN pour les souches bactériennes (Li et al. 2009).....	56
<b>Figure 21:</b> Equation de la réaction d'oxydation de l'éthanol en acide acétique.....	68
<b>Figure 22:</b> (A) Flux métaboliques de la souche <i>A. pasteurianus</i> en milieu simulant les conditions de la pulpe de cacao lors de la fermentation. (B) Contribution en pourcentage du lactate et de l'éthanol dans la formation de l'acide acétique et de l'acétoïne (Adler et al. 2014).....	71
<b>Figure 23 :</b> Carte du monde indiquant les pays d'origine des échantillons des fèves fraîches (nom du pays en noir) et des fèves séchées (nom en rouge).....	77
<b>Figure 24:</b> Carte de la Côte d'Ivoire indiquant les bassins producteurs et le site de fermentation.....	79
<b>Figure 25:</b> Méthode d'échantillonnage avec des écouvillons stériles.....	83
<b>Figure 26:</b> Prélèvement des caisses avant la mise en fermentation.....	84
<b>Figure 27 :</b> Profils DGGE des ADNr 16S présents à la surface des fèves de cacao en provenance du Mexique et de Guyane échantillonnées à différents temps de fermentation (2, 4 ,5 et 6 jours (uniquement Guyane)).....	110
<b>Figure 28 :</b> Indices de diversité $\alpha$ exprimés sous forme de boîtes à moustaches : (A) la richesse (nombre de bandes), (B) Shannon et (C) Inverse Simpson sur les profils obtenus après analyse par PCR-DGGE des communautés bactériennes associées aux fèves de cacao fermentées en provenance de Guyane et du Mexique.....	111

<b>Figure 29</b> : Analyse en composantes principales (ACP) des profils des communautés bactériennes obtenus après analyse par PCR-DGGE des communautés bactériennes associées aux fèves de cacao fermentées en provenance de Guyane et du Mexique. ....	113
<b>Figure 30</b> : (A) Positionnement multidimensionnel non-métrique (NMDS) et (B) dendrogramme pour le regroupement des profils obtenus après analyse par PCR-DGGE des communautés bactériennes associées aux fèves de cacao fermentées en provenance de Guyane et du Mexique. ....	114
<b>Figure 31</b> : Exemple de chromatogramme issue de l'analyse d'une séquence de bande ADN issue du gel DGGE. Visualisation avec le logiciel Sequence Scanner v1.0 (Applied Biosystems). ....	115
<b>Figure 32</b> : Courbes de raréfaction pour les échantillons de fèves de cacao de 3 pays, Mexique, Côte d'Ivoire et Guyane à différents de fermentation. ....	119
<b>Figure 33</b> : Indices de diversité $\alpha$ des échantillons de fèves de cacao à différents temps de fermentations de de 3 pays (Mexique, Côte d'Ivoire et Guyane). ....	120
<b>Figure 34</b> : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de dissimilarité de Bray-Curtis obtenus sur les échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane) à différents temps de fermentation. ....	122
<b>Figure 35</b> : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de similarité Jaccard issus des échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane) à différents temps de fermentation. ....	124
<b>Figure 36</b> : Composition des échantillons en phylum bactériens pour les échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane) à différents temps de fermentation. ....	125
<b>Figure 37</b> : Composition des échantillons en genre bactériens pour les échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane) à différents de fermentation. ....	127
<b>Figure 38</b> : Indices de diversité $\alpha$ des échantillons de fèves de cacao à différents temps de fermentation en provenance du Mexique. ....	131
<b>Figure 39</b> : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de dissimilarité de Bray-Curtis pour les échantillons de fèves de cacao à différents temps de fermentation en provenance du Mexique, grande et petite récolte. ....	132
<b>Figure 40</b> : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de dissimilarité de Jaccard pour les échantillons en provenance du Mexique, grande et petite récolte. ....	134
<b>Figure 41</b> : Composition en phylum des échantillons en provenance des deux récoltes réalisées au Mexique, grande (fevr-18) et petite récolte Oct-18). ....	136
<b>Figure 42</b> : Composition en genre des échantillons en provenance des deux récoltes réalisées au Mexique, grande (fevr-18) et petite récolte Oct-18). ....	137
<b>Figure 43</b> : Diagramme de Venn montrant la composition des communautés bactériennes en genre en début de fermentation selon les 3 pays (Mexique, Côte d'Ivoire et Guyane) ....	139
<b>Figure 44</b> : Courbes de raréfaction pour chaque échantillon des prélèvements environnementaux et des fèves de cacao en fermentation. ....	143
<b>Figure 45</b> : Indices de diversité $\alpha$ des échantillons de fèves de cacao en fermentation en provenance de la plantation, des fèves fraîches en fermentation et de l'environnement de la fermentation. ....	144
<b>Figure 46</b> : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de dissimilarité de Bray-Curtis pour les échantillons de fèves fraîche à différents temps de fermentation et des prélèvements environnementaux en provenance du Mexique. ....	146
<b>Figure 47</b> : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de similarité de Jaccard pour les échantillons de fèves fraîches à différents temps de fermentation et des prélèvements environnementaux en provenance du Mexique. ....	148

<b>Figure 48</b> : Composition des échantillons en phylum bactériens pour les prélèvements des fèves de cacao en fermentation et ceux issus de l'environnement de la fermentation .....	150
<b>Figure 49</b> : Composition des échantillons en genre bactériens pour les prélèvements des fèves de cacao en fermentation et ceux issus de l'environnement de la fermentation. ....	152
<b>Figure 50</b> : Graphique résumant la distribution des isolats au sein des espèces appartenant au genre <i>Acetobacter</i> . ....	160
<b>Figure 51</b> : Graphique résumant la distribution des isolats au sein des différents genres et espèces de BAL.....	162
<b>Figure 52</b> : Gels d'électrophorèse et regroupement des souches en dendrogramme à partir des profils génomiques des souches de BAA après amplification par PCR à l'aide des amorces de séquences répétées (REP, Box1AR, ERIC et (GTG) <sub>5</sub> . L'indice de Sørensen-Dice a été utilisé pour analyser la similarité et un algorithme de regroupement UGPMa a été appliqué pour regrouper les souches. M = marqueur de poids moléculaire 1 kb. T- = témoin négatif.....	165
<b>Figure 53</b> : Gel d'électrophorèse et regroupement des souches en dendrogramme sur les profils génomiques des souches de BAA après amplification par PCR par la méthode de multiplexe (amorces : Box1AR, REP et ERIC). L'indice de Sørensen-Dice a été utilisé pour analyser la similarité et un algorithme de regroupement UGPMa a été appliqué pour regrouper les souches. M = marqueur de poids moléculaire 1 kb. T- = témoin négatif.....	166
<b>Figure 54</b> : Boîtes à moustache des nombres de bandes des profils génomiques selon les différentes amorces (Box1AR, ERIC, (GTG) <sub>5</sub> , Multiplexe et REP). ....	167
<b>Figure 55</b> : Boîtes à moustaches des quantités d'ADN des extractions réalisées par les deux méthodes (choc thermique et le kit Fast DNA Spin) des 7 souches d' <i>A. pasteurianus</i> . ....	168
<b>Figure 56</b> : (A) Profils génomiques obtenus à l'aide de l'amorce (GTG) <sub>5</sub> et (B) Boîtes à moustache des nombres de bandes obtenus sur les profils génomique des souches d' <i>A. pasteurianus</i> . Durée de migration 85 min .....	169
<b>Figure 57</b> : (A) Gels d'électrophorèse des profils (GTG) <sub>5</sub> (B) Positionnement multidimensionnel non-métrique (NMDS) et (C) dendrogramme pour le regroupement des souches selon leur profil génomique. Les extraits d'ADN ont été réalisés par choc thermique.....	170
<b>Figure 58</b> : Dendrogramme généré après analyse statistique de la similarité des profils génomiques obtenus par PCR (GTG) <sub>5</sub> . ....	173
<b>Figure 59</b> : Exemple de courbe de croissance obtenue après le suivi cinétique au spectrophotomètre. ....	175
<b>Figure 60</b> : Exemple de la détermination du taux de croissance spécifique à l'aide de moyenne mobile sur 10 points.....	175
<b>Figure 61</b> : Histogramme des taux de croissance spécifique des 95 souches d' <i>A. pasteurianus</i> , de la souche type LMG01262 <sup>T</sup> et des trois BA : <i>A. fabarum</i> (CI25), <i>A. tropicalis</i> (MX92) et <i>G. oxydans</i> (GF51) .....	176
<b>Figure 62</b> : Courbes d'étalonnage de l'acide lactique, de l'acide acétique, de l'éthanol et du mannitol. ....	177
<b>Figure 63</b> : Chromatogramme obtenu après l'analyse à l'HPLC pour la solution étalon à 10 g/L.....	178
<b>Figure 64</b> : Droite de corrélation entre la DO et la consommation d'éthanol (A) et la production d'acide acétique (B). ....	179
<b>Figure 65</b> : Consommation en mannitol des 89 souches d' <i>A. pasteurianus</i> , et 3 BAA : <i>G. oxydans</i> , <i>A. fabarum</i> et <i>A. tropicalis</i> .....	181
<b>Figure 66</b> : Relation biochimique et regroupement en dendrogramme des 89 souches d' <i>Acetobacter pasteurianus</i> , la souche de LMG01262 <sup>T</sup> et les 3 BAA. ....	183
<b>Figure 67</b> : Boîtes à moustaches des rendements de biomasse en mol par unité de DO pour les deux groupes obtenus dans le dendrogramme et heatmap <b>Figure 66</b> .....	184

<b>Figure 68</b> : Boîtes à moustaches des rendements de biomasse, en mol par unité de DO, produite par mole d'AA produite pour les deux groupes obtenus dans le dendrogramme et heatmap <b>Figure 66</b>	184
<b>Figure 69</b> : Boîtes à moustaches pour l'efficacité de conversion de l'éthanol en acide acétique pour les deux groupes obtenus dans le dendrogramme et heatmap <b>Figure 66</b> .....	184
<b>Figure 70</b> : Boîtes à moustaches des ratios de consommation de l'acide lactique sur la consommation de l'éthanol pour les deux groupes obtenus dans le dendrogramme et heatmap <b>Figure 66</b> . ....	185
<b>Figure 71</b> : Histogramme de la consommation en acide lactique selon les 89 souches d' <i>Acetobacter pasteurianus</i> et la souche type LMG 01262 <sup>T</sup> .....	186
<b>Figure 72</b> : Histogramme de la consommation en éthanol selon les 89 souches d' <i>Acetobacter pasteurianus</i> et la souche type LMG01262 <sup>T</sup> .....	186
<b>Figure 73</b> : Boîtes à moustaches pour l'efficacité de conversion de l'éthanol en acide acétique pour les deux sous-groupes de GII. ....	187
<b>Figure 74</b> : Boîtes à moustaches des ratios de consommation de l'acide lactique sur la consommation de l'éthanol pour les deux sous-groupes de GII.....	187
<b>Figure 75</b> : Boîtes à moustaches pour les taux de croissance spécifique pour les deux groupes obtenus dans le dendrogramme et heatmap <b>Figure 66</b> . ....	188
<b>Figure 76</b> : Dendrogramme généré après analyse statistique de la similarité des profils génomiques obtenus par PCR (GTG) <sub>5</sub> . ....	190



# Introduction

Le cacao est produit dans 50 pays situés en région tropicale dans une zone bien définie appelée la ceinture de cacao où le climat chaud et humide est adéquat pour la culture des cacaoyers. Pour la saison 2017/2018, la production mondiale de cacao a été estimée à 4, 652 millions de tonnes, produits en majorité par 21 pays. Les pays Africains produisent 75 % de la production mondiale, suivis de l'Amérique latine (18 %) et des pays d'Asie (7 %). La Côte d'Ivoire est le leader mondial pour la production et l'exportation du cacao (ICCO 2019). Les fruits du cacaoyer sont des cabosses qui contiennent les graines, appelées fèves. Les fèves de cacao sont entourées d'une pulpe mucilagineuse riche en eau et en sucres. Elles sont transformées en cacao par des procédés post-récolte impliquant la fermentation, le séchage, et les étapes industrielles (torréfaction, broyage, conchage, etc) (De Vuyst et Weckx 2016). L'étape de fermentation est une étape cruciale et spontanée qui repose sur la succession de trois groupes de microorganismes : les levures, les bactéries lactiques et les bactéries acétiques qui réalisent les fermentations alcoolique, lactique et acétique respectivement (Schwan et Wheals 2004). Les levures contribuent de façon importante à la formation de composés aromatiques spécifiques du cacao (Koné *et al.* 2016; Samagaci *et al.* 2016; Pereira *et al.* 2017; Visintin *et al.* 2017). La production d'acide lactique n'est pas désirée en fermentation du cacao car cet acide n'est pas volatil ; Il peut pénétrer dans la fève et provoquer des saveurs non appréciées par le consommateur. Quant à l'acide acétique produit, il va se diffuser dans les fèves, tuer l'embryon et catalyser des réactions chimiques qui permettent de produire des précurseurs aromatiques du cacao. De nombreuses études ont été réalisées afin d'identifier les communautés microbiennes impliquées dans la fermentation du cacao principalement à l'échelle d'un pays. Le consortium microbien est composé de souches de levures (*Saccharomyces cerevisiae*, *Pichia kluyveri*, *Pichia kudriavzevii*, *Torulaspota delbruekii*, *Hanseniaspora opuntiae*, *Candida tropicalis*), de bactéries lactiques (*Lactobacillus fermentum*, *Lactobacillus plantarum*, *Leuconostoc pseudomesenteroides*) et de bactéries acétiques avec *Acetobacter pasteurianus* (De Vuyst and Weckx, 2016; Figueroa-Hernández *et al.* 2019; Ozturk and Young, 2017). Cependant, en fonction du producteur, différents paramètres peuvent varier et influencer l'établissement de ce consortium microbien. En effet, les différentes variétés de cacaoyer, les pratiques locales de récolte, de

mise en fermentation et de déroulement de la fermentation peuvent avoir un impact sur la diversité microbienne (Barel 2013; Visintin *et al.* 2016; Hamdouche *et al.* 2019).

Au début de la fermentation, une diversité plus importante a pu être observée. Des genres appartenant à la famille *Enterobacteriaceae* ont été isolés dans certaines fermentations (Garcia-Armisen *et al.* 2010; Illegghems *et al.* 2015). Ces bactéries participent à la dépectinisation de la pulpe, mais leur rôle n'est pas tout à fait déterminé. A partir du second jour de fermentation, les bactéries lactiques et acétiques dominent les fermentations.

Les fermentations sont réalisées de manière spontanée bénéficiant d'une inoculation par l'environnement. Entre l'ouverture de la cabosse et la mise en fermentation, les fèves sont mises en contact avec différentes surfaces liées à l'environnement ainsi qu'avec les mains des opérateurs. La contribution de chaque élément environnemental à la colonisation de ces fèves a été peu étudiée.

Le contrôle des fermentations afin d'améliorer la qualité organoleptique du cacao marchand peut être possible en inoculant les fèves de cacao avec des souches de levures et bactéries présélectionnées. Pour cela, des starters de cultures ont été développés. L'analyse du cacao final a démontré que sa qualité organoleptique est meilleure et appréciée par les consommateurs. Cependant, ces starters de culture ont été réalisés en se basant sur l'analyse de la diversité interspécifique des bactéries, sans s'intéresser à la diversité intra spécifique des bactéries lactiques et acétiques qui sont les principales actrices des fermentations. L'analyse de la diversité intraspécifique de souches isolées de fèves de cacao originaires de différents pays pourrait être judicieuse pour la mise au point d'un starter de culture préservant la diversité de chaque pays producteur ce qui permettrait d'éviter la standardisation de la fermentation des fèves de cacao dans le monde.

Au cours de cette étude, nous avons voulu répondre à des questions de recherches afin d'améliorer la connaissance sur les communautés bactériennes impliquées dans la fermentation du cacao pour proposer des pistes potentielles pour l'élaboration de starter de culture.

Les questions de recherche associées à ces travaux sont les suivantes :

- Quel est l'impact de l'origine géographique sur les communautés bactériennes impliquées dans la fermentation des fèves de cacao ?

- Quel est l'impact de la période de récolte des cabosses sur les communautés bactériennes associées à la fermentation des fèves ?
- Quel est l'origine du pool de bactéries impliqué dans la fermentation du cacao ?
- Existe-t-il des taxons spécifiques d'une origine géographique ?
- Existe-t-il des souches de bactéries lactiques et acétiques spécifiques à d'une origine géographique ?
- Est-ce que les souches de bactéries acétiques qui appartiennent à ~~d'~~une même espèce présentent les mêmes caractéristiques biochimiques lors de la fermentation des fèves de cacao ?

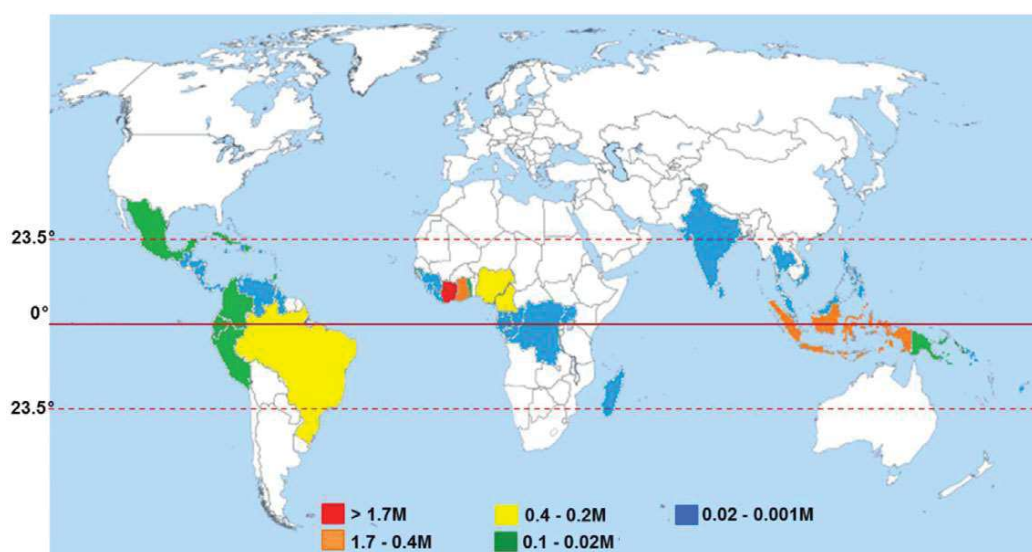
La première partie de ce manuscrit sera consacrée à l'état de l'art sur des généralités concernant le cacao, les processus de transformation des fèves et plus particulièrement la fermentation, la composition des communautés microbiennes associées à la fermentation. La seconde partie concernera le matériels et méthodes utilisés au cours de la thèse. Ensuite les résultats obtenus durant la thèse seront présentés sous la forme de trois chapitres : le premier chapitre présentera les résultats issus de la comparaison des communautés bactériennes associées aux fèves de cacao en fermentation selon l'origine géographique et la période de récolte des cabosses de cacao. Le second chapitre portera sur la discrimination de l'origine du pool de bactéries impliquées dans les fermentations de cacao au niveau d'une fermentation réalisée au Mexique. Le dernier chapitre portera sur la discrimination de souches de bactéries lactiques et acétiques selon leurs origines géographiques. Enfin, l'ensemble des résultats sera discuté dans la dernière partie.

# Chapitre I : Etat de l'art

## I.1 Généralités sur le cacao

### I.1.1 Production mondiale de cacao

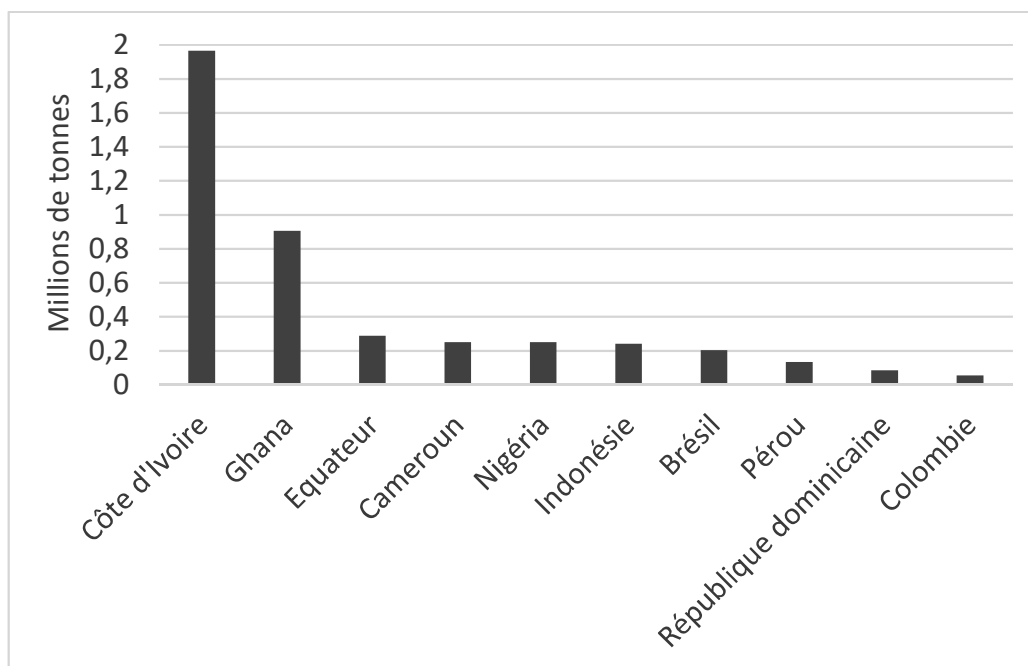
Le cacao est produit depuis des milliers d'années et est originaire d'Amérique centrale dans les civilisations Mésoaméricaines. En se basant sur la linguistique du cacao, les plus vieilles preuves de domestication du cacao proviennent du golfe du Mexique au sud de l'état de Veracruz (Gilet 2006). De nos jours, le cacao est produit dans 50 pays situés en région tropicale : en Afrique, en Amérique latine et en Asie, dans une zone bien définie appelée la ceinture de cacao, entre 23.5 ° de latitude au Nord et 23.5 ° au Sud de l'équateur. Les principaux pays producteurs en Afrique sont la Côte d'Ivoire, le Ghana, le Nigeria, le Cameroun, l'Ouganda, le Togo, la Sierra Leone, et Madagascar. En ce qui concerne l'Asie : l'Indonésie, la Malaisie, la Papouasie nouvelle Guinée, l'Inde sont les plus grands producteurs. En Amérique latine : le Brésil, l'Equateur, le Mexique, le Pérou, la République Dominicaine, la Colombie, le Venezuela, le Guatemala et Haïti (**Figure 1**) (Ozturk et Young 2017; ICCO 2019).



**Figure 1:** Carte des principaux pays producteurs de cacao situés dans la zone équatoriale.

L'organisation mondiale du cacao (ICCO) a été créée afin de renforcer le secteur mondial du cacao et de favoriser son développement durable. Elle regroupe les pays producteurs et consommateurs de cacao.

Selon le rapport de l'ICCO pour la saison 2017 / 2018, la production mondiale de cacao s'est élevée à 4, 6 millions de tonnes par 21 pays producteurs principalement. Soixante-quinze pourcents de la production mondiale a été réalisée par les pays Africains. Les pays d'Amérique latine, représentent seulement 18 % de la production mondiale suivi des pays d'Asie qui produisent 7 % (ICCO 2019).



**Figure 2:** Production de cacao en millions de tonnes des 10 premiers pays producteurs pour la saison 2017/2018 (ICCO 2019).

Sept pays produisent plus de 90 % de la production mondiale, la Côte d'Ivoire étant le premier pays producteur, représente plus de 42 % de la production avec 1,9 millions de tonnes de cacao récoltées, suivi du Ghana (905 000 de tonnes) puis de l'Equateur, du Cameroun, du Nigéria, de l'Indonésie et du Brésil (entre 204 000 et 287 000 tonnes). Les autres pays ont produit moins de 100 000 tonnes (**Figure 2**) (ICCO 2019).

Dans ces pays, la production de cacao est une grande source de revenus pour de nombreuses familles. Elle se fait de manière artisanale dans des petites plantations familiales. La majorité de leur production de cacao est exportée principalement en Union Européenne, suivi de l'Asie,

l'Océanie et l'Amérique du Nord. Ces derniers transforment les fèves en liqueur, beurre de cacao, poudres ou tourteaux et commercialisent les produits finaux (chocolat, confiserie, produits de cosmétiques et de santé (ICCO 2019)

### 1.1.2 Le cacaoyer *Theobroma cacao* L

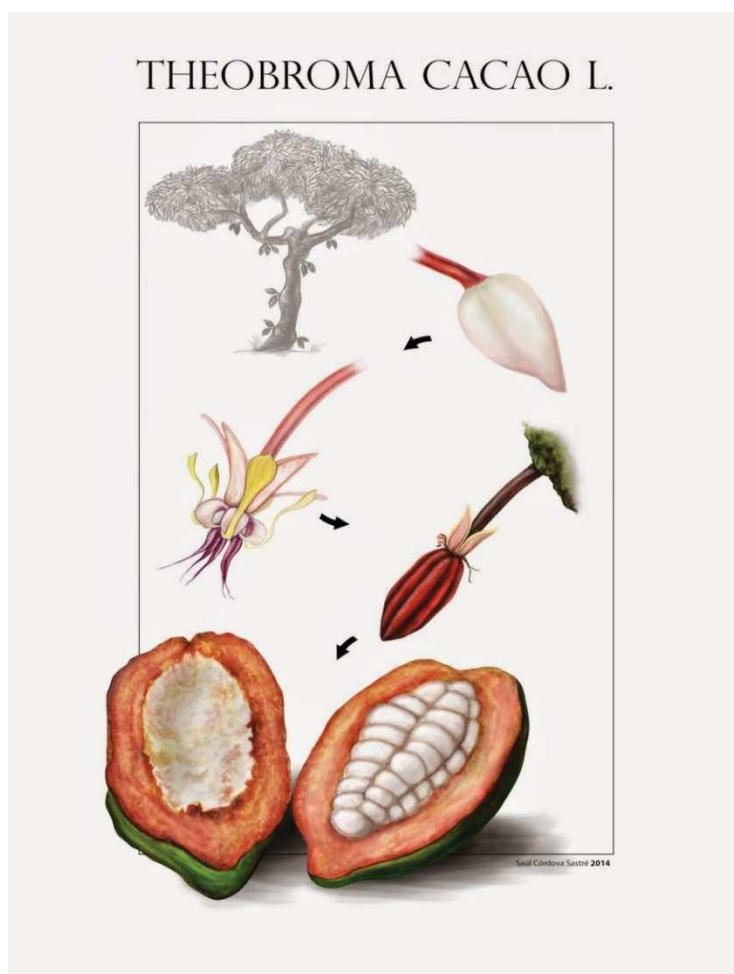
#### 1.1.2.1 Généralités sur le cacaoyer

L'espèce *Theobroma cacao* L appartient au sous embranchement des Angiospermes, ordre des *Malvales* et à la famille des *Sterculiacées*, selon la classification de Linné basée sur l'observation morphologique. Elle a été reclassée dans la famille des *Malvaceae* selon la classification phylogénétique, basée sur l'analyse du gène codant pour l'ARNr 18S et des gènes chloroplastiques (*atpB* et *rbcL*). Elle est la seule parmi les 22 espèces du genre *Theobroma* à être utilisée pour la production de cacao.

Le cacaoyer est un petit arbre à feuilles persistantes pouvant atteindre de 5 à 15 m de haut selon qu'il soit cultivé dans les plantations ou trouvé à l'état sauvage.

À partir de sa troisième année, l'arbre commence à être productif et il atteint son rendement maximal à partir de 6/7 ans.

Les fleurs poussent directement sur les troncs (cauliflores) tout au long de l'année et une fleur sur 500 donne un jeune fruit, appelé chérelle qui deviendra la cabosse à maturité. A l'intérieur de la cabosse se trouvent les graines, appelées fèves (**Figure 3**). La couleur de la cabosse est une indication pour estimer la maturité, car elle passe du vert au jaune, orangé, rouge. Elle est de forme ovoïde de 15 à 30 cm de long et contient 30 à 60 fèves, ovoïdes ou aplaties, rangées longitudinalement, en épi (Gallais *et al.* 1991).



**Figure 3:** Evolution morphologique de la fleur à la cabosse.

Le cacaoyer est planté dans les zones tropicales, à la ceinture de cacao (23,5 ° au Nord et au Sud de l'Equateur) où les conditions de température, d'humidité et de pluviométrie sont les plus adéquates. Le cacaoyer a besoin d'une température entre 18 et 32°C pour croître et un pourcentage d'humidité de 80 %. Des bananiers sont souvent cultivés dans la même plantation pour apporter l'ombrage nécessaire aux jeunes cacaoyers (Pontillon 1998, Schwan et Fleet 2015). Les cacaoyers ont besoin d'un sol riche en matière organique, légèrement acide, qui assure une bonne rétention de l'eau, selon la pluviométrie un sol argileux ou sableux peut convenir (Barel 1998).



### 1.1.2.2 Variétés de cacaoyer

*Theobroma cacao* L présente une importante diversité de variétés nommées selon leurs origines géographiques. Cependant trois variétés se distinguent dans la fabrication du cacao : *Criollo*, *Forastero* et *Trinitario*.



**Figure 4:** Cabosses des trois principales variétés de cacaoyer utilisées dans la production mondiale de cacao.

La variété *Criollo*, originaire d'Amérique centrale et du Sud, est cultivée principalement au Mexique, en Colombie et au Venezuela. Ses cabosses sont de forme allongée, jaune orangé à maturité et les fèves arrondies avec des cotylédons blancs à rose. C'est une variété très fragile et sensible aux maladies ; elle ne représente que 5 % de la production mondiale. Cependant, sa faible amertume et sa qualité aromatique lui permet d'être très convoitée par les chocolatiers pour faire du chocolat de luxe (Kongor *et al.* 2016).

La variété *Forastero*, est originaire de l'Amazonie, cultivée principalement en Afrique de l'ouest (Côte d'Ivoire, Ghana), à l'Equateur et au Venezuela. Elle est la plus cultivée dans le monde, 80 % de la production mondiale proviennent des cacaoyers de cette variété. Les cabosses sont lisses, vertes puis jaunes à maturité, leurs fèves sont aplaties et leurs cotylédons

violet foncé. Une forte amertume est obtenue dans le chocolat final. Par leur importante diversité génétique, les cacaoyers *Forastero* sont plus résistants aux maladies et plus robustes. Les cacaos obtenus à partir de *Forastero* sont considérés de qualité standard (Kongor *et al.* 2016).

La variété *Trinitario*, originaire de Trinidad qui est issue du croisement entre la variété *Criollo* et *Forastero*, présente des caractéristiques similaires avec ses parents, avec la taille des arbres, la forme des cabosses et la couleur des fèves qui sont variables selon le « cultivar ou sous-variété » de *Trinitario*. Son rendement très élevé lui permet de participer de 10 à 15 % à la production mondiale. Les cacaoyers *Trinitario* sont principalement cultivés dans les pays d'où sont originaires les *Criollo* (Gilet 2006, Schwan et Fleet 2015).

Au cours des dernières années, de nombreux hybrides de cacaoyer ont été créés afin d'augmenter leur résistance à certains champignons pathogènes (*Moniliophthora perniciosa* responsable de la maladie du balai de la sorcière (moniliose) et *Phytophthora megakarya* responsable de la pourriture brune de la cabosse), mais également pour améliorer leur vitesse de croissance ou leur rendement. C'est le cas du CCN-51, hybride venant de l'Equateur, issu du croisement entre les variétés Iquitos (Equateur-Pérou, 45,4%), *Criollo* (Amazonie, 22,2%) et Amelonado (Ghana et Amérique centrale, 21,5%) qui est résistant notamment à *Moniliophthora perniciosa* ou de la variété Mercedes, qui a une précocité de croissance et un rendement final très important.

#### 1.1.2.3 Composition chimique de la pulpe mucilagineuse et des fèves de cacao

La couleur des fèves dépend de la variété de cacao, allant de blanc à violet (**Figure 5**). Elles sont entourées d'une pulpe mucilagineuse, blanche, visqueuse avec un pH faible (3 à 4). La pulpe est principalement composée de 80 à 90 % d'eau, de 10 à 15 % de sucre (glucose, fructose et sucrose), de 2 % d'acide citrique et de 1 % de pectine et autres polysaccharides. La composition détaillée de pulpe en provenance de 3 pays est donnée dans le **Tableau 1**.



**Figure 5:** Coupe transversale d'une cabosse de cacao montrant les fèves entourées de la pulpe à l'intérieur.

**Tableau 1:** Composition de la pulpe de cacao (en g pour 100 g de matière fraîche).

Composition	Côte d'Ivoire <sup>a</sup>	Nigéria <sup>a</sup>	Malaisie <sup>a</sup>
Eau	82.60	82.50	85.90
Glucose	3.00	5.06	4.90
Fructose	3.80	6.07	5.35
Saccharose	4.35	1.92	1.35
Plante et polymères de paroi cellulaire	2.81	-	1.48
Acide citrique	1.31	0.79	0.29
Protéines/peptides	0.57	0.51	0.43
Acide aminés libres	0.15	0.11	0.21
Matières grasses	0.45	0.75	0.35
Métaux	0.24	0.22	-
Vitamines	0.05	0.05	-
Ethanol	0	0.1	0.2

<sup>a</sup> (Pettipher 1986), inspiré de (Lima *et al.* 2011)

C'est la présence d'acide citrique (44-45 mg/g de matière sèche) qui confère à la pulpe ce faible pH initial. D'autres acides organiques sont présents dans la pulpe mais en de plus faibles proportions, comme l'acide malique (7-8 mg/g de matière sèche), l'acide succinique (5-6 mg/g de matière sèche) (Ho *et al.* 2014).

Concernant les fèves de cacao non fermentées, celles-ci sont composées en majorité de matières grasses (55,2 % de matière fraîche) puis de protéines (21,6 % de matière fraîche), glucides (15,5 % de matière fraîche) et d'eau (4,2 % de matière fraîche) (Afoakwa *et al.* 2013). Concernant la composition en glucides des fèves, elle est principalement composée de 17 mg/g de matière fraîche de saccharose, de 0.5 mg/g matière fraîche de glucose et de 1mg/g de fructose. Pour la composition en acides organiques, les mesures ont été réalisées sur les matières sèches, les fèves sont composées de 16-17 mg/g d'acide citrique, 13-14 mg/g d'acide succinique, 7-8 mg/g d'acide oxalique et de 1-2 mg/g d'acide malique (Ho *et al.* 2014). Cependant la composition des fèves évolue peu au cours de la fermentation contrairement à celle de la pulpe, car c'est principalement dans la pulpe que la fermentation a lieu.

## I.2 De la récolte au cacao marchand

### I.2.1 De la récolte à l'écabossage des cabosses

La maturation des cabosses se fait tout au long de l'année et elles sont généralement récoltées lors de deux campagnes, la « grande » campagne et la « petite » campagne. Par exemple, en Côte d'Ivoire, la grande campagne débute en octobre et se termine en février et la petite campagne débute en mars et se termine en juin. C'est lors de la « grande » campagne que la majorité des cabosses est récoltée, car elles ont atteint la maturité adéquate et contiennent une quantité plus importante de fèves, plus charnues avec plus de pulpe (Ozturk et Young 2017).

L'ouverture de la cabosse, dit l'écabossage se réalise à l'ombre des cacaoyers ou dans un lieu spécialement désigné à cet effet. Le délai d'écabossage correspond à l'intervalle de temps entre la récolte de la cabosse et son ouverture. Les cabosses récoltées sont ouvertes en deux à l'aide d'instruments tranchants ou contendants, comme un gourdin ou une machette, perpendiculairement sans endommager les fèves contenues à l'intérieur (**Figure 6**).

Le but est de réaliser une coupe facilitant le détachement des fèves de la coque interne de la cabosse. Une fois ouverte, les fèves sont extraites manuellement (**Figure 7**), stockées dans des sacs en polypropylène et/ou toile de jute pour être transférées sur le site de fermentation.



**Figure 6:** Ecabossage des cabosses à l'aide d'une machette au Mexique, Octobre 2018.



**Figure 7:** Extraction manuelle des fèves de cacao au Mexique, Octobre 2018.

## I.2.2 Les traitements post récolte

Les fèves de cacao sont transformées par une succession de traitements post-récolte : fermentation, séchage, le stockage et toutes les étapes industrielles.

### I.2.2.1 La fermentation

La fermentation des fèves de cacao s'effectue dans des fermes de manière artisanale, dans des caisses en bois (**Figure 8**) ou en tas sur des feuilles de bananiers (**Figure 9**) selon le pays producteur.



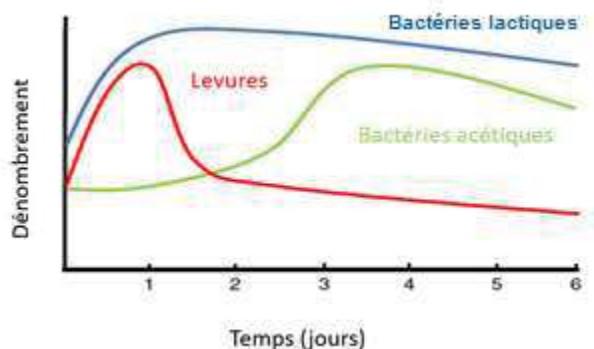
**Figure 8:** Caisses de fermentation en bois utilisées au Mexique.



**Figure 9:** Fermentation en tas sur des feuilles de bananiers en Côte d'Ivoire.

Elle dure de 4 à 8 jours, selon la variété de cacaoyer, les conditions climatiques, la masse de fèves mise en fermentation, le délai d'écabossage, la qualité et la quantité de la pulpe mucilagineuse (Barel 2013). La couleur brunâtre des fèves, l'odeur forte d'acide acétique nous permet de déterminer la fin de fermentation. La quantité de fèves mises en fermentation peut être de 20 Kg pour les fermentations expérimentales à 200 à 1000 Kg pour la production de cacao industrielle. Selon la variété de cacao, les cabosses ne contiennent pas la même quantité de pulpe autour des fèves et cela a un impact sur le temps de fermentation.

La fermentation des fèves de cacao repose principalement sur la succession de trois groupes de microorganismes : les levures, les bactéries lactiques et les bactéries acétiques (**Figure 10**) qui réalisent les fermentations alcoolique, lactique et enfin acétique respectivement.



**Figure 10:** Dynamique des communautés microbiennes impliquées au cours de la fermentation des fèves de cacao (De Vuyst et Weckx 2016).

Ces trois étapes fermentaires sont nécessaires afin d'obtenir les caractéristiques organoleptiques du cacao final.

De la récolte à la mise en fermentation, les différentes pratiques locales selon les fermiers ou le pays producteur influencent le déroulement de la fermentation. Par exemple la sélection de cabosses matures et saines ou le délai de la mise en fermentation après écabossage peuvent favoriser la croissance des bactéries lactiques à défaut des levures.

Hamdouche *et al.* (2019) ont démontré que le temps de stockage des cabosses avant l'écabossage et la mise en fermentation (2 jours et 8 jours) n'ont qu'un faible impact sur les communautés microbiennes. Cependant, si le délai entre l'écabossage et la mise en fermentation des fèves est trop long cela peut favoriser la croissance des bactéries lactiques à défaut des levures et aboutir à des défauts de la fermentation alcoolique. C'est pourquoi il est préconisé de ne pas dépasser un délai de 6 h (Barel 2013).

La réalisation du brassage des fèves au cours de la fermentation est une pratique couramment employée pour les fermentations en caisse. Il est réalisé manuellement ou à l'aide de pales en bois prévues à cet effet. Il est recommandé de réaliser 2 à 3 brassages de la masse de fèves à

partir du second jour de la fermentation afin d'homogénéiser les fermentations (Camu *et al.* 2008; Papalexandratou *et al.* 2011a; Barel 2013; Saltini *et al.* 2013; Hamdouche *et al.* 2019).

#### 1.2.2.2 le séchage et le stockage des fèves

Lorsque la fermentation est finie (environ 4 à 8 jours après écabossage), l'étape de séchage des fèves est initiée, elle aide à stopper définitivement la fermentation. Cette étape est importante d'une part car elle permet de réduire la quantité d'eau présente dans les fèves après la fermentation (de 50 % à 8 % d'eau) évitant ainsi le développement des moisissures. D'autre part elle permet l'évaporation de l'acide acétique produit lors de la fermentation. Dans la majorité des fermes, le séchage s'effectue au soleil. Les fèves sont étalées sur des toiles de jute, des bâches ou des claies à même le sol, pendant 8 à 15 jours (Djedjro *et al.* 2008) **(Figure 11)**.



**Figure 11:** Séchage solaire des fèves de cacao après fermentation.

Des séchages artificiels peuvent aussi être réalisés à l'aide de fours avec ventilation à air chaud et peuvent être utilisés lorsque les conditions climatiques ne sont pas adaptées à un séchage à l'air libre (Hii *et al.* 2009; Barel 2013). Cependant, de cette manière l'acide acétique n'est pas totalement éliminé des fèves de cacao et peut provoquer son acidité.



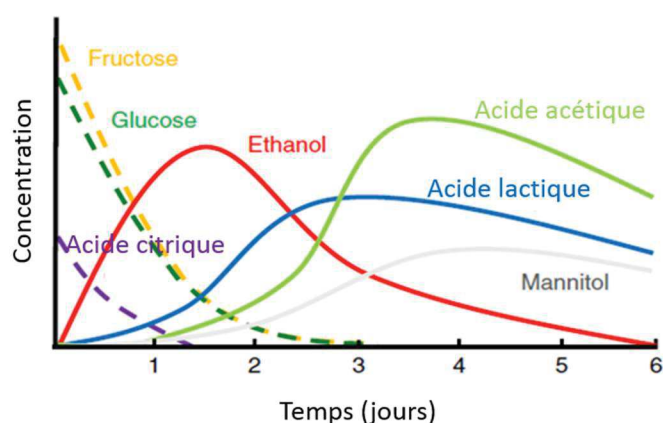
Les fèves de cacao sont stockées dans des sacs en toile de jute ou polypropylène jusqu'à l'envoi dans les pays importateurs (**Figure 12**) où les fèves subiront les étapes de transformation industrielle (torréfaction, broyage, conchage).



**Figure 12:** Stockage des fèves avant l'envoi aux industriels pour transformation.

### I.3 Composition des communautés microbiennes au cours du processus de fermentation du cacao

Durant la fermentation, seuls certains microorganismes ont la capacité de croître au sein des fèves de cacao et de la pulpe. La composition chimique de la masse des fèves, la température, le pH, la quantité de métabolites (sucres et acides principalement) évoluent au



**Figure 13:** Cinétique de la dégradation et de la production des métabolites pendant la fermentation de fèves de cacao (De Vuyst et Weckx 2016).

cours de la fermentation créant ainsi des conditions propices pour le développement d'un groupe de microorganismes ou d'un autre (**Figure 13**).

### 1.3.1 Méthodes d'analyse des communautés microbiennes

Les premières méthodes d'analyse des communautés microbiennes correspondaient à des méthodes dépendantes de la culture. Les échantillons alimentaires ou leurs dilutions sont étalés sur des milieux de culture sélectifs ou non, pour estimer le nombre de microorganismes cultivables et / ou pour réaliser l'isolement clonal. Ensuite, les isolats sont identifiés par des méthodes phénotypiques classiques basés sur des critères morphologiques, biochimiques, physiologiques et enzymatiques associées ou non à des méthodes moléculaires. Ces méthodes comportent des limites puisqu'elles ne permettent pas de détecter les microorganismes viables mais non cultivables (VNBC, Viable But Non-Culturable). En effet, seul 1% des microorganismes de la plupart des environnements sont dits cultivables, sous-estimant la biodiversité microbienne (Staley & Konopka 1985; Amann *et al.* 1995; Kirk *et al.* 2004). Ces organismes non cultivables nécessitent des conditions difficiles à reproduire en laboratoire et des substances nécessaires à leur développement présentes dans leur environnement qui sont pour la plupart encore inconnues de l'homme.

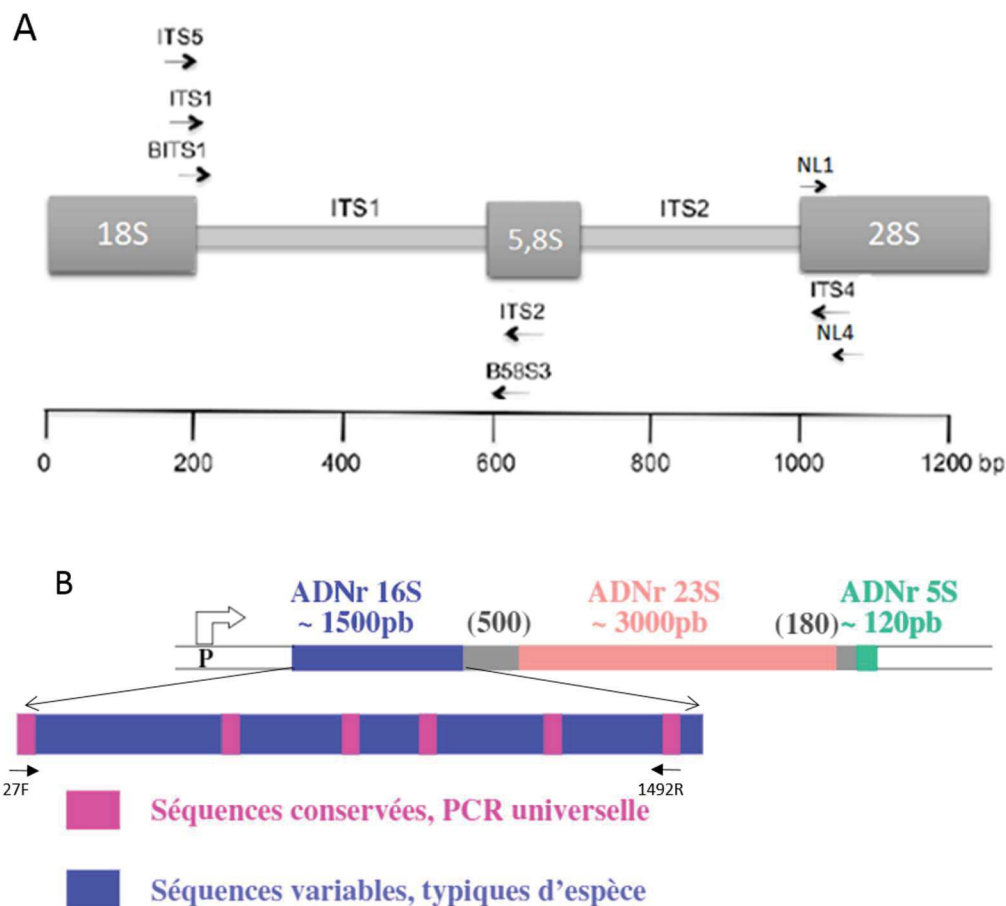
C'est pourquoi, des méthodes moléculaires, indépendantes de la culture, basées sur l'analyse de l'ADN total d'un échantillon ont été développées en complément des méthodes dépendantes de la culture.

En ce qui concerne l'étude des communautés microbiennes associées à la fermentation du cacao, les deux méthodes ont souvent été associées afin de déterminer les communautés microbiennes impliquées (Nielsen *et al.* 2007a; Lefeber *et al.* 2011; Arana-Sánchez *et al.* 2015).

#### 1.3.1.1 Marqueurs moléculaires universels

Les méthodes indépendantes de la culture nécessitent la plupart du temps l'utilisation de marqueurs moléculaires universels. Un marqueur moléculaire est une portion de l'ADN, qui doit être ubiquiste, avoir une évolution lente, posséder des régions hautement conservées quel que soit l'espèce. Ces régions permettent de mettre au point des amorces universelles

pour l'amplification par PCR des séquences variables encadrées par ces régions. En 1977, Carl Woese a proposé comme marqueurs universels de taxonomie l'utilisation des gènes codant pour les ARN ribosomiques (Woese and Fox 1977).



**Figure 14:** Organisation des gènes des ARN ribosomiques chez les Eucaryotes (A) et chez les Procaryotes (B) (White et al. 1990).

Les ARNr sont retrouvés dans toutes les cellules vivantes. Pour l'analyse des communautés bactériennes, c'est le gène de l'ARNr 16S qui est utilisé comme marqueur moléculaire (**Figure 14** (White *et al.* 1990)). En ce qui concerne l'étude des communautés fongiques se sont les espaces internes transcrits: ITS (Internal Transcribed spacer) qui sont les plus utilisés comme marqueurs moléculaires. Ces espaces sont situés entre les régions des gènes de l'ARNr 18S et 28S et encadrent le gène de l'ARNr 5.8S. Ce sont des régions non codantes de l'ARN ribosomique et hautement polymorphes selon les espèces. La région D1/D2 de la grande sous

unité 26S de l'ARN ribosomique est aussi utilisée comme code-barre universel pour les levures.

Cependant, l'opéron ribosomique est présent en un nombre de copies variables selon les génomes. Cette variation du nombre de copies introduit un biais d'estimation de l'abondance relative des différents organismes présents dans le mélange d'ADN provenant de l'échantillon analysé.

#### *1.3.1.2 Techniques d'empreinte moléculaire*

Différentes techniques ont été développées afin d'obtenir une vision globale de la diversité des communautés microbiennes ou l'identification des microorganismes présents dans un environnement complexe. Elles sont appelées techniques d'empreinte moléculaire ou de « fingerprinting » et permettent d'obtenir un profil spécifique pour l'étude des communautés présentes dans l'échantillon comme les empreintes digitales ou génétiques peuvent l'être chez l'homme. L'ensemble de la communauté microbienne ciblée au sein d'un échantillon est étudié et peut être aussi comparé à d'autres échantillons. Ces techniques sont pertinentes pour étudier rapidement leur structure et leur évolution selon différents échantillons.

Différentes techniques de fingerprinting peuvent être employées pour l'étude des communautés microbiennes. Pour ces méthodes c'est le polymorphisme au niveau des séquences nucléotidiques qui est analysé. Une première étape d'amplification par PCR est réalisée en ciblant les marqueurs moléculaires universels. Puis, la seconde étape consiste en l'analyse des produits PCR soit par des techniques d'électrophorèse basées sur l'analyse de la séquence des amplicons générés (DGGE - électrophorèse sur gel en gradient dénaturant chimique - TGGE électrophorèse sur gel en gradient de température (Muyzer 1999), SSCP (polymorphisme de conformation des simples-brin), ARISA, l'analyse automatisée des espaces intergéniques (Ghosh *et al.* 2015), soit des techniques faisant appel à des enzymes de restrictions : ARDRA, (analyse des fragments de restriction de l'ADN ribosomal, T-RFLP, - polymorphisme de longueur des fragments de restriction).

La PCR-DGGE fait partie des méthodes les plus couramment utilisées parmi les techniques d'empreinte moléculaire pour l'étude des communautés levuriennes et bactériennes notamment pour celles impliquées dans la fermentation du cacao. Le principe de l'électrophorèse sur gel d'agarose avec gradient dénaturant (DGGE) repose sur le fait qu'au

cours de la migration; les fragments d'ADN double brins sont soumis à un environnement dénaturant croissant. En fonction de leur composition en AT (Adénine-Thymine) et GC (Guanine- Cytosine), les fragments n'ont pas le même point de fusion et sont dénaturés à une vitesse différente et ainsi ne migrent pas de la même manière. Donc cette méthode permet une séparation des fragments d'ADN amplifiés de même taille selon leur composition en bases azotés. Nielsen *et al.* (2001) ont mis en évidence la dynamique des populations de levures impliquées dans la fermentation du cacao au Ghana. L'analyse des profils obtenus a montré que la dynamique des populations de levures est dépendante de la méthode, du site et de l'état d'avancement de la fermentation. Papalexandratou *et al.* (2011d) ont comparé les communautés levuriennes selon différentes zones géographiques et ont mis en évidence que selon la zone, trois espèces ont été retrouvées majoritairement (*S. cerevisiae*, *Hanseniaspora sp.* et *P. kudriavzevii*). Arana-Sánchez *et al.* (2015) ont utilisé la PCR-DGGE afin d'identifier les espèces levuriennes dominantes des fermentations au Mexique. Hamdouche *et al.* (2015) ont analysé la dynamique des populations microbiennes au cours du traitement post-récolte (fermentation et séchage au soleil) afin de déterminer des marqueurs microbiens moléculaires associés à la fermentation du cacao et à l'étape de séchage des fèves. Hamdouche *et al.* (2019), à l'aide de la PCR-DGGE a pu étudier l'impact du traitement post-récolte du cacao sur les communautés microbiennes associées aux fèves de cacao. Le principal paramètre qui influence la structure des communautés microbiennes est le temps de fermentation, suivi du brassage, alors que la durée de stockage des cabosses n'a pas d'impact significatif.

Enfin, la DGGE peut être réalisée après l'amplification de portions de gènes spécifiques aux BAL ou aux BAA afin de suivre la dynamique de ces microorganismes. Cela permet également de confirmer les résultats obtenus à l'aide des méthodes dépendante des cultures, obtenus à l'aide de milieux de culture spécifiques à ces microorganismes (Papalexandratou *et al.* 2011c, 2013).

La PCR-DGGE comporte certaines limites qu'il est nécessaire de prendre en compte. Elle est dépendante des conditions expérimentales comme les conditions d'extraction de l'ADN, de l'amplification par PCR, etc. Les taxons, dits rares, présents en de faibles abondances dans les échantillons, ne sont pas détectés par cette méthode sous estimant donc la diversité des communautés microbiennes. La longueur des fragments d'ADN amplifiés et analysés ne dépasse pas 500 pb rendant parfois l'identification au niveau de l'espèce impossible. Elle ne

peut fournir qu'une évaluation qualitative ou pseudo-quantitative de la communauté microbienne. Cependant, malgré ces limites cette méthode reste très pertinente et contribue à l'amélioration des connaissances sur les communautés microbiennes.

### 1.3.1.3 La métagénomique

En 1998, Handelsmann a introduit le concept de métagénomique, dérivé du concept statistique de méta-analyse et de génomique (analyse du matériel génétique d'un microorganisme). La métagénomique permet l'étude du microbiome d'intérêt, sa composition en termes de diversité d'espèces mais aussi d'abondance. Cette méthode est basée sur l'analyse de l'ADN et est couplée à des outils de séquençage haut débit et des analyses bio-informatiques.

Au cours des dernières années, les méthodes moléculaires ont bénéficié de grandes avancées technologiques et continuent à être en perpétuelle évolution grâce à l'avancement des technologies de séquençage à haut débit et de la bio-informatique.

Ces méthodes permettent d'étudier les communautés microbiennes mixtes dans les matrices complexes, de détecter les microorganismes à un niveau faible d'abondance et de discriminer les populations minoritaires présentes dans les échantillons (Handelsman *et al.* 1998; Ivey et Phister 2011). Il existe deux grands types de métagénomique : la métagénomique ciblée, appelée metabarcoding et la métagénomique globale.

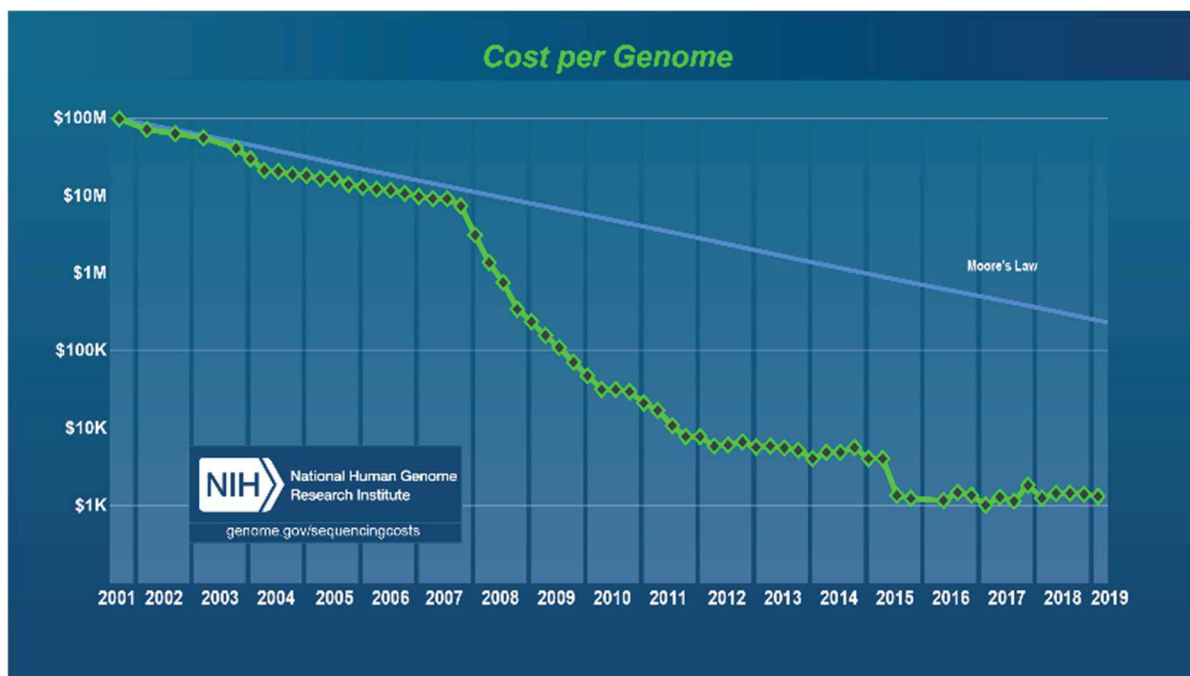
Le metabarcoding n'est pas précisément de la métagénomique mais plutôt de la métagénétique puisqu'elle consiste à amplifier un gène présent chez les microorganismes étudiés. Pour l'analyse des communautés bactériennes, c'est le gène de l'ARNr 16S qui est utilisé, cependant l'utilisation de ce marqueur apporte un biais, par sa présence en de nombreuses copies dans les génomes bactériens, surestimant ainsi l'abondance.

Alors que la métagénomique globale est basée sur le séquençage complet du génome qui se fait sans *a priori*. Cette méthode appelée aussi Whole Genome Sequencing (WGS), consiste à fragmenter tous les ADN présents dans un échantillon et à séquencer les fragments puis les ré assembler bio informatiquement pour reconstruire les génomes.

Ces deux types de métagénomique permettent à terme d'obtenir un regroupement des échantillons analysés selon la composition des communautés microbiennes associées.

#### 1.3.1.4 Les nouvelles techniques de séquençage (NGS)

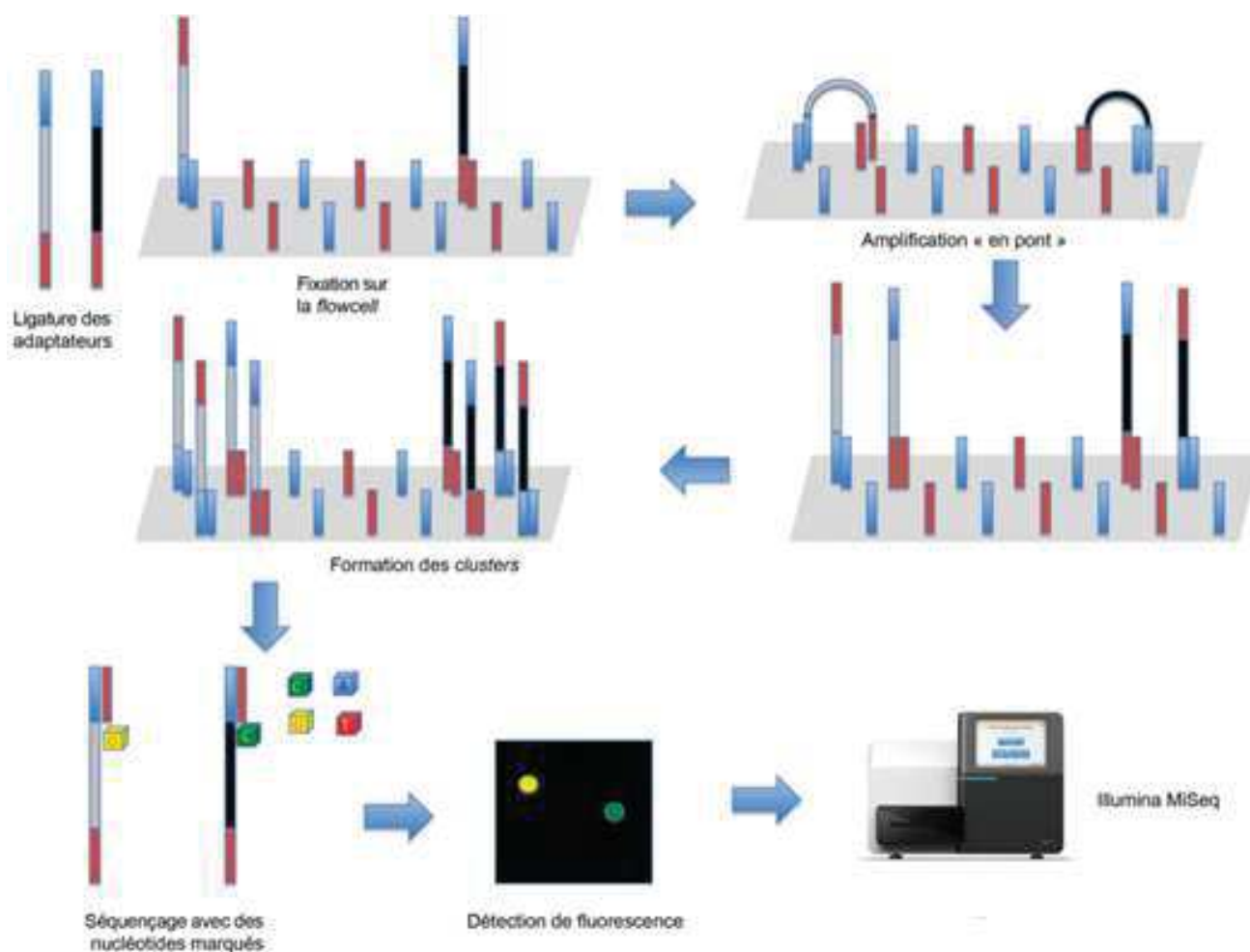
Les nouvelles techniques de séquençage, dites de seconde génération sont basées sur le séquençage à haut débit de fragments d'ADN. C'est le séquençage ciblé de marqueurs génétiques informatifs sur le plan taxonomique, qui permet de mesurer la biodiversité, à moindre coût, de manière complète et répétable. Elles sont dites de haut débit car elles permettent de générer des millions de séquences/ fragments d'ADN en quelques heures pour un coût de plus en plus dégressif. Les objectifs d'améliorations des méthodes de séquençage sont surtout basés sur l'augmentation de la longueur de lecture et de la profondeur du séquençage. Les coûts réduisent avec l'amélioration et la miniaturisation des automates pour une utilisation en routine, par exemple le coût pour le séquençage complet d'un génome humain est passé de 100 millions de dollars en 2011 à 1000 dollars en 2015-2019 (**Figure 15**). Les technologies les plus utilisées dans l'analyse NGS sont le pyroséquençage 454 (Roche) et le séquence avec terminateurs réversibles Illumina (Solexa).



**Figure 15:** le coût du séquençage d'un génome humain (3,2 milliards de pb).

Le pyroséquençage est la première technique de séquençage nouvelle génération développée par Roche, employée dans l'analyse par metabarcoding des communautés microbiennes. Lors du processus de séquençage les dNTP ne sont pas introduits tous ensemble dans le mélange réactionnel mais un à un. Si le nucléotide introduit correspond à celui que la polymérase a besoin pour synthétiser le brin d'ADN complémentaire à la matrice alors celui-ci est incorporé. Cette incorporation permet la libération d'un pyrophosphate inorganique (PPi) à la suite d'une réaction enzymatique et le couplage avec une luciférine. Cette libération émet un signal lumineux qui est détecté par l'automate et qui le reproduit sous forme de pic sur un pyrogramme. Cette méthode est limitée par la courte longueur de lecture (25-250 bp) et le nombre réduit d'échantillons que l'on peut analyser à la fois. Elle a été abandonnée en 2012 pour être remplacée par le séquençage avec terminateurs réversibles Illumina (Solexa). Après une étape préalable de constitution d'une banque d'ADN, la technique de séquençage Illumina repose sur une amplification en ponts de fragments d'ADN sur un support solide, appelé flow-cell suivi d'un séquençage par synthèse chimique (SBS) des amplicons. L'ancrage des amplicons sur la flow-cell se fait à l'aide d'extensions nucléotidiques (adaptateur) ajoutés à l'extrémité 5' des brins d'ADN lors de la préparation de la banque. Le brin complémentaire est ainsi synthétisé. L'amplification par pont permet de produire des millions de copies d'un brin fixé à la flow-cell, ce qui va former un groupement d'ADN clonal, appelé « cluster ». Les brins anti-sens sont ensuite clivés pour être séquencés à l'aide d'ajout de nucléotides marqués, cette incorporation émet un signal lumineux propre à chaque nucléotide qui est détecté par l'automate (**Figure 16**). Le MiSeq est l'automate le plus utilisé en routine dans les laboratoires. Différents types de kits peuvent être employés pour une analyse de fragments d'ADN allant de 2\* 250 à 300 pb. Le kit v2 génère 15 millions de lecture en single-read (30 millions de lectures en paired-end) avec un score de qualité (Q-score, est une prédiction de la probabilité d'une erreur dans l'incorporation d'une base) supérieur à 75 % dans le cas du séquençage 2\*250 pb. Un second kit, le v3, génère 25 millions de single-reads (50 millions de lectures en paired-end) avec un Q-score supérieur à 70 % pour un séquençage de fragment de 2\*300 pb.





**Figure 16:** Principe de l'approche Illumina (Bonnin et al. 2017).

L'avancée de la métagénomique en écologie microbienne est liée au coût dégressif de ces méthodes de séquençage. Ces outils sont de plus en plus appliqués pour l'étude des communautés microbiennes associées à des matrices alimentaires, dans le domaine médical également afin d'élucider la diversité et la structure microbienne, mais également pour réaliser l'annotation fonctionnelle et la cartographie des voies métaboliques. Des études récentes ont été réalisées sur les communautés microbiennes impliquées dans la fermentation du cacao. Illegheims *et al.* 2012, ont été les premiers à analyser les communautés microbiennes par une approche de métagénomique par la méthode de pyroséquençage. Cette approche a permis de mettre en évidence une plus grande diversité bactérienne (principalement  $\gamma$ -Proteobacteria) et fongique qu'auparavant, avec les microorganismes du consortium microbien dominant les fermentations (*H.uvarum*, *H.*

*opuntiae*, *S. cerevisiae*, *L. fermentum* et *A. pasteurianus*). Agyirifo *et al.* (2019), ont analysé la structure des communautés microbiennes associée à la fermentation de cacao au Ghana par WGS et ont démontré que les bactéries sont majoritaires (90 %) suivies des levures (9 %) et des bactériophages (1 %). Ils ont également démontré que 4 genres bactériens dominent les fermentations : *Lactobacillus*, *Gluconacetobacter*, *Acetobacter* et *Gluconobacter*. Ils ont également déterminé les voies métaboliques codant pour les composés aromatiques nécessaires à la production des composés d'arômes et des saveurs présents dans le cacao final. Serra *et al.* (2019) ont analysé des communautés microbiennes associées à des fermentations de cacao dans différentes localités du Brésil par métabarcoding. Ils ont démontré la présence d'un consortium microbien au sein de chaque fermentation, composé de levures (*Pichia* et *Candida*), de champignons filamenteux (*Aspergillus*, *Penicillium* et *Fusarium*), de BAL (*Leuconostoc*, *Lactococcus* et *Lactobacillus*), de BAA (*Acetobacter* et *Gluconobacter*) et d'entérobactéries et de bacilli. L'espèce *A. pasteurianus* est retrouvée en majorité dans les fermentations. Ils ont également démontré que certaines espèces de levures et de BAL ont été retrouvées uniquement dans une localité et pas l'autre, montrant une variabilité microbienne.

Papalexandratou *et al.* 2019 ont étudié le lien entre des variétés de cacaoyers et leur diversité microbienne impliquée lors de fermentation de cacao au Nicaragua par métabarcoding. Ils ont démontré que le traitement post-récolte du cacao surpasse l'impact de la variété de cacao sur la diversité des communautés microbiennes.

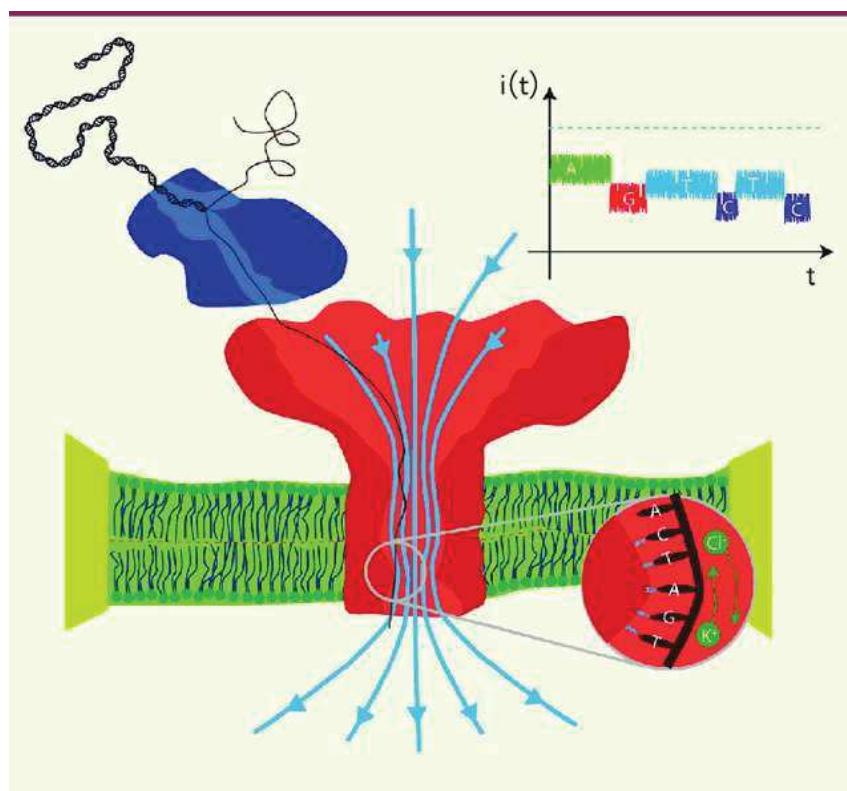
Ces méthodes indépendantes de la culture présentent également des limites, l'amplification de l'ADN total des microorganismes présents dans les échantillons est réalisée et ces méthodes ne permettent pas de discriminer les microorganismes vivants des morts. C'est pourquoi les méthodes indépendantes de la culture sont parfois couplées à des méthodes dépendantes de la culture. Ces deux méthodes sont complémentaires et selon le type de matrice alimentaire ou environnementale et la question de l'étude, différentes approches peuvent être appliquées utilisant des outils variés.

#### 1.3.1.5 Les techniques de séquençage de troisième génération

De nouvelles méthodes de séquençage de troisièmes générations sont en développement mais pas encore utilisées en routine dans les laboratoires. Elles s'affranchissent de l'étape

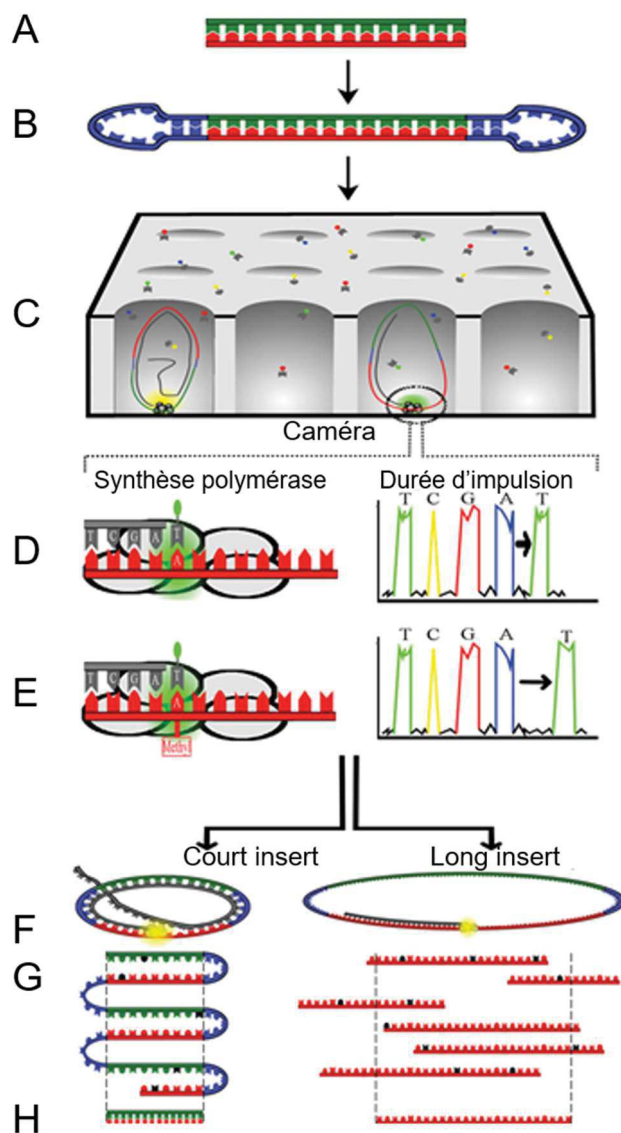
d'amplification de l'ADN lors de la préparation de la banque d'amplicons et permettent donc le séquençage direct des molécules d'ADN. Sans étape d'amplification PCR, les modifications de base sont directement détectées pendant le séquençage (détection de l'épigénie). Il existe deux technologies :

- La technologie Oxford Nanopore : les brins d'ADN sont passés au travers de nanopores. Ce sont soit des protéines lipidiques soit des ouvertures nanométriques réalisées par gravure dans des membranes solides ultrafines déposées sur un substrat. Les bases du fragment qui traversent le nanopore sont lues en temps réel et l'identité de celle-ci est obtenue grâce à la mesure de l'intensité du courant électrique traversant le nanopore (**Figure 17**). Cette méthode détecte des fragments longs (supérieurs à 10 kb)



**Figure 17:** Principe de la technologie Oxford, séquençage par Nanopore biologique. L'ADN à séquençer est capté par une enzyme (hélicase ou ADN polymérase, en bleu) qui extrude l'ADN simple brin à une vitesse réduite. Le nanopore lui-même (en rouge) constitue une ouverture dans la membrane lipidique (en vert) dans laquelle le champ électrique (flèches bleues) force le passage de l'ADN simple brin (en marron). Ce dernier est ralenti au passage par la constriction du nanopore (encart du bas) et permet l'identification des bases individuelles par mesure de l'intensité du courant électrique traversant le pore à un instant donné (encart du haut). Le trait pointillé bleu représente le courant à vide du nanopore (Montel 2018).

- La technologie Pacific Biosciences : elle est basée sur la méthode SMRT (Single Molecule, Real-Time Sequencing) qui permet le séquençage de longs fragments (10 kb). L'appareil de séquençage dispose de plusieurs cellules SMRT, contenant chacune des détecteurs ZMW (Zero-Mode Waveguide). Chacun de ces ZMW contient une polymérase immobilisée qui incorpore des nucléotides couplés à des fluorochromes, libérant ce dernier à l'intérieur de chaque ZMW. C'est par l'émission du pic de fluorescence émit que l'identification de la base est obtenue (**Figure 18**) (Montel 2018, Ardui *et al.* 2018).



**Figure 18:** Principe du SMRT séquençage utilisé dans les technologies Pacific biosciences.

Le séquençage commence par la préparation d'une bibliothèque à partir d'ADN double brin (A) auquel sont ligaturés des adaptateurs en épingle à cheveux (B). Cette bibliothèque est ensuite chargée sur une cellule SMRT composée de chambres d'observation à l'échelle nanométrique (ZMWs). Les molécules d'ADN de la bibliothèque seront tirées au fond des ZMW où la polymérase incorporera des nucléotides marqués par fluorescence (C). Notez que tous les ZMW ne contiendront pas une molécule d'ADN car la bibliothèque est chargée par diffusion. La fluorescence émise par les nucléotides est enregistrée par une caméra en temps réel. Ainsi, non seulement la couleur de la fluorescence peut être enregistrée, mais aussi le temps entre l'incorporation des nucléotides, appelé durée d'impulsion (IPD) (D, panneau de droite). Lorsqu'une L'ADN polymérase rencontre des nucléotides sur le brin d'ADN contenant une modification (épigénétique), comme par exemple une modification de la 6-méthyl-adénosine (E, panneau gauche), alors la DPI sera retardée (E, panneau droit) par rapport à l'ADN non méthylé (D, panneau droit). En raison de la structure circulaire de la bibliothèque, un insert court sera couvert plusieurs fois par la longue lecture continue (CLR). Chaque passage de la

molécule d'ADN originale est appelé une sous-lecture, qui peut être combinée en une séquence consensus très précise appelée séquence consensus circulaire (CCS) ou lectures d'insert (ROI) (F-H, panneau de gauche). Bien que le séquençage SMRT utilise toujours un modèle circulaire, les longues bibliothèques d'insertion n'ont généralement qu'un seul passage et génèrent donc une séquence linéaire avec un taux d'erreur de passage unique (nucléotides noirs) (FG, panneau de droite). Ensuite, les passes uniques qui se chevauchent peuvent être combinées en une séquence consensuelle de haute qualité (H, panneau de droite). Dans l'ensemble, les lectures CCS ont l'avantage d'être très précises, tandis que les passes uniques se distinguent par leur grande longueur de lecture (>20 kb) (Ardui *et al.* 2018).

### 1.3.2 Le consortium microbien

De nombreuses études ont été réalisées afin de mieux comprendre la structuration des communautés microbiennes impliquées dans la fermentation du cacao en fonction des différents pays producteurs et de la variété de cacao. La conclusion importante de ces études est qu'il existe un consortium microbien qui domine les fermentations qui comprend des espèces de levures, de bactéries lactiques (BAL) et bactéries acétiques (BAA).

Certains microorganismes sont retrouvés dans les fermentations à travers tous les pays producteurs formant ce consortium microbien, qui est composé principalement de levures (*Saccharomyces cerevisiae*, *Pichia kluyveri*, *Pichia kudriavzevii*, *Torulaspota delbruekii*, *Hanseniaspora opuntiae*, *Candida tropicalis*), de BAL (*Lactobacillus fermentum*, *Lactobacillus plantarum*, *Leuconostoc pseudomesenteroides*) et de BAA avec *Acetobacter pasteurianus*. Cependant, selon le pays producteur, d'autres espèces de BAA peuvent être incluses dans le consortium, notamment *Acetobacter tropicalis*, *Acetobacter ghanensis*, *Acetobacter senegalensis* retrouvées au Ghana, en Malaisie et à l'Equateur (Camu *et al.* 2008; Papalexandratou *et al.* 2011a, 2013, 2019; Pereira *et al.* 2013a; Mota-Gutierrez *et al.* 2018; Aguirifo *et al.* 2019; Serra *et al.* 2019).

D'autres microorganismes sont détectés de manière sporadique au cours de la fermentation ou ne sont trouvés que dans certains pays. Les entérobactéries sont souvent décelées en début de fermentation et des bactéries sporulantes, les *Bacillus*, en fin de fermentation. Ces bactéries peuvent avoir un rôle dans la fermentation (Illegghems *et al.* 2012; De Vuyst et Weckx 2016; Pereira *et al.* 2016; Ozturk et Young, 2017).

### I.3.3 Succession et rôle des microorganismes au cours de la fermentation

#### Les levures

La fermentation débute dès l'ouverture de la cabosse, le faible pH (3,6) et l'environnement anaérobie créé par la pulpe autour des fèves sont favorables au développement des levures qui sont donc les premiers microorganismes à coloniser les fèves de cacao. Elles interviennent durant les premières 24 h-48 h de la fermentation. Au bout d'un jour les levures atteignent leur abondance maximale puis déclinent jusqu'à la fin de la fermentation (De Vuyst et Weckx 2016). Pour leur croissance, les levures consomment les sucres (glucose, fructose et saccharose) présents dans la pulpe et les transforment en éthanol. Cette réaction est légèrement exothermique; la température au sein des fèves passe de 25-30°C à 35-40°C (Schwan et Wheals 2004). Les pectines présentes dans la pulpe sont dégradées grâce aux enzymes pectinolytiques des levures. Ainsi, la viscosité de la pulpe diminue et celle-ci est transformée en jus, ce qui permet une entrée d'air au sein des fèves et favorise un autre groupe de microorganismes : les bactéries acétiques.

Les levures présentent une importante biodiversité et varient d'un pays à l'autre, d'une région à une autre et selon le processus de fermentation (fermentation en « caisse » ou « en tas »). Cependant trois genres sont majoritairement retrouvés : *Saccharomyces*, *Hanseniaspora* et *Pichia*.

*Saccharomyces cerevisiae* est l'espèce la plus fréquemment isolée dans les fermentations à travers le monde, en raison de sa croissance rapide, de son activité pectinolytique et de sa tolérance à l'éthanol (Daniel *et al.* 2009); Ludlow *et al.* 2016 ; Pereira *et al.* 2016). Selon le pays producteur de cacao, les espèces dominantes dans les fermentations peuvent être différentes. Au Ghana, ce sont les espèces *S. cerevisiae*, *Hanseniaspora opuntiae* et *Pichia kudriavzevii* et qui sont trouvées en grande majorité (Daniel *et al.* 2009). Au Mexique, on retrouve : *S. cerevisiae*, *Hanseniaspora* et *P. kudriavzevii*, *Saccharomycopsis crataegensis* et *Hanseniaspora guilliermondii* (Arana-Sánchez *et al.* 2015), en Equateur, *S. cerevisiae*, *P. kudriavzevii* et *Pichia manshurica* (Papalexandratou *et al.* 2011c) en Malaisie, *S. cerevisiae*, *Hanseniaspora thailandica*, *Hanseniaspora opuntiae* and *P. kudriavzevii* (Meersman *et al.* 2013; Papalexandratou *et al.* 2013). En ce qui concerne l'Indonésie ce sont les levures du genre *Saccharomyces* et *Candida* et l'espèce *Hanseniaspora uvarum* (Ardhana, 2003; Jamili, 2016), à Cuba, *P. manshurica*, *H. opuntiae* et *P. kudriavzevii*, et en Côte d'Ivoire *S. cerevisiae*, *H.*

*opuntiae*, *P. kudriavzevii* et *P. kluyveri* qui sont les espèces dominantes (Hamdouche *et al.* 2015; Koff *et al.* 2017). Au Nicaragua, les levures qui dominent sont *S. cerevisiae*, *Hanseniaspora uvarum/opuntiae* et *P. kudriavzevii* (Papalexandratou *et al.* 2019) (**Tableau 2**).

Selon leur capacité fermentaire, leur tolérance à l'éthanol et à la chaleur, la succession des espèces de levures évolue au cours de la fermentation. Ce sont les levures du genre *Hanseniaspora* qui dominent principalement au début de la fermentation et *Saccharomyces* et *Pichia* du milieu à la fin de la fermentation. Dans la phase initiale, l'espèce *Hanseniaspora opuntiae* est majoritairement dominante mais sa population décline au cours de la fermentation car elle a une faible tolérance à la chaleur et à l'éthanol. *Saccharomyces cerevisiae*, par sa rapide croissance à un pH acide, une forte tolérance à l'éthanol et à la chaleur est décelée pratiquement tout le long de la fermentation et domine la fin de la fermentation. De la même manière l'espèce *P. kudriavzevii* est retrouvée au long de la fermentation mais en moindre proportion. D'autres espèces de levures non fermentaires sont également détectées au cours de la fermentation, comme les levures du genre *Candida* (*Candida sp.* et *Candida ethanolica*) connues pour leur rôle de lutte biologique. Elles inhibent la croissance de pathogènes (*Bacillus cereus*, *Escherichia coli*, *Salmonella sp.*, *Staphylococcus aureus* et *Pseudomonas sp.*), de moisissures toxigènes des genres *Aspergillus* et *Penicillium* et ont donc un rôle dans la qualité et la sûreté du cacao (Mahazar *et al.* 2015; Ruggirello *et al.* 2019).

L'impact des levures sur la qualité aromatique du cacao final a été largement étudié au cours de ces dernières années. Notamment les souches de levure *P. kudriavzevii* contribuent de façon importante à la formation de composés aromatiques spécifiques du cacao. Elles ont un métabolisme des sucres et une production d'éthanol efficace et leur activité métabolique influence la qualité finale du cacao, avec un meilleur développement de la couleur des fèves et une composition aromatique plus riche (Koné *et al.* 2016; Samagaci *et al.* 2016 ; Pereira *et al.* 2017). Les souches de *S. cerevisiae* jouent un rôle important pour la qualité aromatique finale du cacao (Visintin *et al.* 2017).





Selon le mode de fermentation employée, la structure des communautés levuriennes est différente. Visintin *et al.* (2016) ont comparé deux modes de fermentation en Côte d'Ivoire, la fermentation en caisse et celle en tas sur des feuilles de bananier. En termes de dominance d'espèces, en caisse *S. cerevisiae* est présente dans presque toutes les fermentations ainsi que *Candida ethanolica* présente les deux derniers jours de la fermentation. Trois espèces de levures ne sont détectées que dans les fermentations en tas : *Hanseniaspora guilliermondii*, *Pichia mashurica* et *Schizosaccharomyces pombe* et deux seulement en caisse : *Torulasporea delbrueckii* et *Pichia kudriavzevii*. Pour la fermentation en tas, *S. cerevisiae* et *Schizosaccharomyces pombe* sont détectées tout au cours de la fermentation avec *H. guilliermondii*, *P. mashurica* présentes également en de fortes proportions (Visintin *et al.* 2016). La même comparaison a été faite sur des fermentations réalisées au Cameroun (Mota-Gutierrez *et al.* 2018) et il a été mis en évidence une plus grande diversité levurienne dans les fermentations réalisées en caisse qu'en tas.

### Les entérobactéries

L'activité de dépectinisation des levures permet une diminution de la viscosité de la pulpe et une entrée d'air progressive au sein des fèves, créant ainsi des conditions de microaérobiose favorables aux *Enterobacteriaceae*. Les entérobactéries appartenant aux genres *Tatumella*, *Pantoea* et *Erwinia* sont détectées au début des fermentations en de fortes abondances. Il a été démontré qu'elles sont capables de métaboliser les glucides ainsi que le citrate présent dans la pulpe (Garcia-Armisen *et al.* 2010; Illegheems *et al.* 2015). Ces dernières peuvent jouer un rôle aussi de dépectinisation de la pulpe dès le début d'entrée d'air.

Les espèces du genre *Tatumella* ne sont pas désirables pour la fermentation du cacao car elles produisent de l'acide gluconique à partir du glucose, créant ainsi une forte acidité dans les fèves et consommant le glucose disponible pour la croissance des levures et des BAL (Papalexandratou *et al.* 2011b, 2011c; De Vuyst & Weckx 2016). Certaines espèces de *Tatumella* ont été répertoriées comme des phytopathogènes de plantes, notamment d'Ananas au Mexique (Marin-Cevada *et al.* 2007) et pourraient donc être pathogènes pour le cacaoyer.

### Les bactéries lactiques (BAL)

En même temps que les levures et les entérobactéries, les BAL interviennent dans les premiers jours de la fermentation et sont détectées jusqu'à la fin. Elles fermentent les sucres présents au sein de la pulpe principalement en acide lactique mais selon l'espèce elles produisent aussi l'éthanol et l'acide acétique.

Elles consomment également l'acide citrique pour le transformer en acide lactique, acide acétique et autres composés (Ho *et al.* 2015) et réduisent le fructose en mannitol. Ce sont des bactéries lactiques anaérobies facultatives, homofermentaires ou hétérofermentaires selon l'espèce. Parmi les bactéries lactiques, *Lactobacillus fermentum* et *Lactobacillus plantarum* sont les espèces dominantes en fin de fermentation ; Elles sont hétérofermentaires car ne produisent pas uniquement de l'acide lactique (De Vuyst et Weckx 2016). Par leur production d'acides, les bactéries lactiques participent à la transformation de la pulpe et des fèves de cacao et peuvent participer aussi à l'activation des enzymes précurseuses d'arôme du cacao.

Cependant, l'acide lactique produit n'est pas volatil, s'il pénètre dans les cotylédons en grande quantité il ne pourra pas être éliminé après les étapes de séchage au soleil. Cela peut amener à des défauts aromatiques du cacao final (Barel 2013).

### Les bactéries acétiques (BAA)

Les BAA représentent un groupe important de bactéries intervenant dans les étapes de fermentation. Ce sont des bactéries aérobies et sont donc favorisées par l'entrée d'air qui est de plus en plus importante au cours de la fermentation. Les pratiques agricoles qui consistent à réaliser des brassages toutes les 24 h favorisent leur croissance. Elles dominent les fermentations à partir de 48 h. Leur domination à la fin des fermentations est expliquée par leur résistance aux hautes températures (40-45°C) et à une forte acidité du milieu (pH=4,5) (Illegheems *et al.* 2013; Sato *et al.* 2015). Elles utilisent l'alcool produit par les levures afin de le transformer en acide acétique; Cette réaction est exothermique puisque la température peut atteindre 50°C. Les BAA consomment également l'acide lactique produit par les BAL.

L'acide acétique produit est très volatil et va pénétrer dans les cotylédons, tuer l'embryon et ainsi initier la production de précurseurs d'arômes nécessaires à la qualité organoleptique du cacao. Grâce à l'étape de séchage, l'acide acétique sera éliminé des fèves de cacao par évaporation (Schwan et Wheals 2004; Barel 2013; De Vuyst et Weckx 2016). Cependant, leur

diversité est très limitée et est représentée en majorité par une seule espèce : *Acetobacter pasteurianus*.

### Les *Bacillus*

A la fin de la fermentation, l'augmentation du pH (de 3,5 à 5), de l'aération et de la température (jusqu'à 50°C) dans la masse de cacao permettent le développement de bactéries aérobies sporulantes, thermotolérantes du genre *Bacillus* qui deviennent les espèces bactériennes dominant à la fin de la fermentation. Les différentes espèces isolées durant les fermentations de cacao sont principalement : *Bacillus pumilus*, *Bacillus licheniformis*, *Bacillus subtilis*, *Bacillus cereus*, *Bacillus megaterium*, et *Bacillus flexus* (Ardhana 2003; Moreira *et al.* 2013; Pereira *et al.* 2013a; Illeghems *et al.* 2015). Une étude réalisée par (Ouattara *et al.* 2008) a démontré que les *Bacillus* pourraient être impliqués dans la dégradation de la pulpe de cacao, car les espèces de *Bacillus* sont capables de produire une enzyme pectinolytique. Cependant, en fin de fermentation, elles peuvent avoir un impact sur la qualité organoleptique du cacao final par la production d'acides organiques et d'arômes tels que les pyrazines et le 2,3-butanediol (Ardhana 2003; Schwan et Wheals 2004).

Pour les BAL et les BAA un chapitre sera consacré à l'analyse de leur diversité inter et intra spécifique dans la suite de l'état de l'art.

## 1.4 Les bactéries lactiques et acétiques : principales actrices de la fermentation du cacao

Les bactéries lactiques et les bactéries acétiques ont un rôle important dans les fermentations de cacao. Elles permettent d'obtenir les conditions nécessaires pour initier la production des enzymes précurseuses d'arômes. Les principales espèces impliquées dans les fermentations de fèves de cacao sont largement connues dans la littérature scientifique ; de nombreuses études ont été menées dans les différents pays producteurs de cacao afin de les identifier. Différentes études ont été réalisées pour analyser la diversité interspécifique des bactéries lactiques et acétiques impliquées dans les fermentations du cacao au Ghana à l'aide de méthodes d'empreintes moléculaires comme les PCR sur séquences répétées (GTG)<sub>5</sub> (Camu *et al.* 2007; De Vuyst *et al.* 2008; Papalexandratou *et al.* 2009). Ils ont démontré la pertinence de

la méthode pour identifier les isolats au niveau de l'espèce. Cependant, jusqu' à présent, la diversité intra spécifique des espèces bactériennes majoritaires n'a été étudiée que sur 4 souches isolées au Ghana (Camu *et al.* 2007; Lefeber *et al.* 2010).

#### I.4.1 Méthodes d'analyse de la diversité intra spécifique

##### I.4.1.1 Notions d'espèce et de souche

La notion d'espèce, proposée par Mayr en 1942 ne s'applique pas aux bactéries : « Les espèces sont des groupes de populations naturelles, effectivement ou potentiellement interfécondes, qui sont génétiquement isolées d'autres groupes similaires et engendrant une progéniture viable et féconde » (Mayr, 1964). En effet, les bactéries ne sont pas capables de reproduction sexuée. Elles ont une reproduction asexuée par multiplication (bourgeonnement, scissiparité, mitose...) et échangent leur matériel génétique par transfert horizontaux. Les bactériologistes ont dû trouver une définition adaptée : « une espèce est constituée par sa souche type et par l'ensemble des souches considérées comme suffisamment proches de la souche type pour être incluses au sein de la même espèce ». Selon, le critère choisi, cette notion d'espèce varie par l'évolution des technologies employées ou l'évolution de ces mêmes critères. Deux comités ont été chargés de définir les critères et une définition de l'espèce : l'International Committee on Systematic Bacteriology (ICSB) en 1987 (Moore *et al.* 1987) et l'International Committee on Systematic of Procaryotes (ICSP) en 2002 (Stackebrandt 2002). D'après ces comités, une espèce est définie (genomospecies) comme l'ensemble des souches dont les ADN totaux s'hybrident à plus de 70 % et que le pourcentage d'identité des séquences des gènes de l'ARNr 16S est à 97% d'homologie. Cette définition ne se base sur aucuns critères phénotypiques (Konstantinidis *et al.* 2006). De plus, Stackebrandt *et al.* 2006 ont suggéré que le pourcentage de similitude entre les séquences du gène de l'ARNr 16S des souches devrait être de 98.7/99%. Une autre méthode, peut être utilisée pour définir une espèce bactérienne : l'ANI (Average Nucleotide Identity). C'est l'identité nucléotidique moyen entre des gènes ou des séquences d'ADN communs à 2 génomes. Cet indice décrit pour la première fois par Konstantinidis et Tiedje (2005) est calculé en comparant des gènes conservés selon plusieurs souches. Avec l'explosion des données génomiques et le développement rapide de la technologie de séquençage, le séquençage complet du génome de souches d'intérêt peut être

réalisé et cet indice calculé à l'échelle du génome entier. Il permet d'évaluer la distance génétique entre deux génomes. Par exemple, l'étude de l'ANI moyen de 30 souches appartenant à l'espèce *Escherichia coli*, a démontré qu'elles ont une valeur élevée de similarité (0.934). L'ANI moyen calculé sur des souches d'espèces différentes appartenant au genre *Escherichia* a une valeur plus faible (0.872) que lors de la comparaison de souches appartenant à la même espèce (Zhang *et al.* 2014).

On peut définir aussi les espèces en se basant sur leur contenu génétique. On distingue dans les génomes bactériens : le « core » génome, le génome accessoire et le pangénome. Le « core » génome représente l'ensemble des gènes partagés et conservés à toutes les souches d'une même espèce. Le génome accessoire, représente uniquement l'ensemble des gènes présents chez la souche d'intérêt, ou ceux présents uniquement dans deux ou plusieurs souches et le pangénome qui est l'ensemble des gènes du « core » génome et du génome accessoire.

#### *1.4.1.2 Méthodes d'analyse de la diversité intra spécifique*

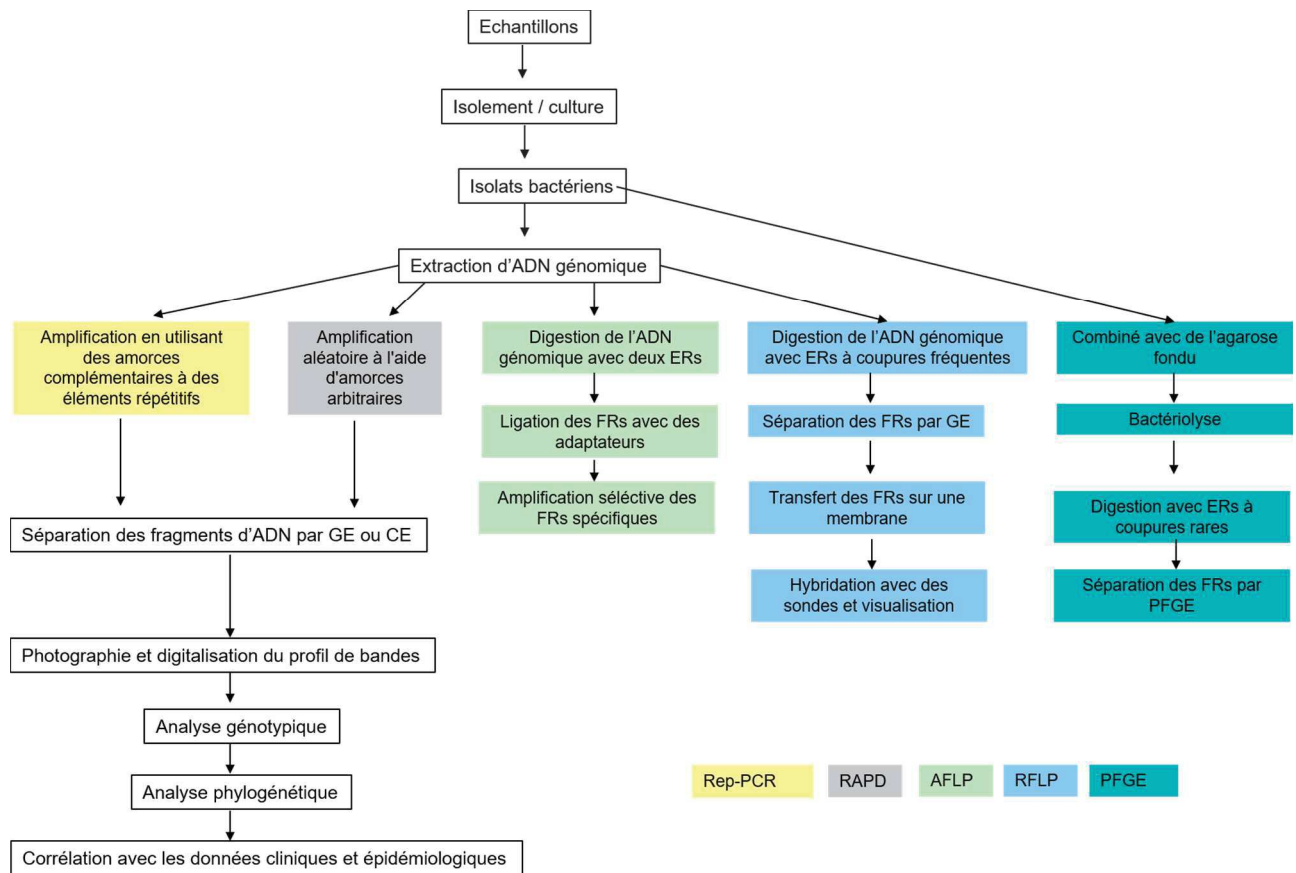
Il est difficile de typer les souches uniquement d'un point de vue phénotypique. En effet, il n'y a pas assez de variabilité entre deux souches pour les discriminer, que ce soit par la morphologie des colonies, par leur sensibilité aux antibiotiques, etc. C'est pourquoi l'apport de la génomique a beaucoup contribué au génotypage des souches. Différentes méthodes ont été établies et reposent sur l'étude du génome des microorganismes. Ces méthodes permettent d'obtenir des profils d'ADN spécifiques à chaque souche, nous permettant d'obtenir leur code barre d'identification, leur empreinte moléculaire (« fingerprinting ») (Li *et al.* 2009).

Ces méthodes peuvent être classées en deux catégories :

- (1) Les méthodes basées sur l'obtention de profils de bandes d'ADN après amplification et/ou digestion enzymatique de l'ADN génomique (PFGE, RFLP, PCR-RFLP, AFLP, RAPD).
- (2) Celles basées sur le séquençage d'ADN étudiant le polymorphisme des séquences (MLST, NGS).

## I.4.1.2.1 Comparaison de profils génomiques

Ces méthodes basées sur l'obtention de profils de bandes permettent une discrimination des souches selon la taille et nombre de fragments obtenus après l'amplification d'ADN et/ou de restriction (**Figure 19**).



**Figure 19:** Classification des méthodes de génotypage de souches bactériennes basées sur l'obtention de profils d'ADN. Avec ERs: enzymes de restriction, FRs: fragments de restriction, GE: gel d'électrophorèse, CE: Electrophorèse capillaire, RAPD: random amplified polymorphic DNA (amplification aléatoire d'ADN polymorphe), AFLP: amplified fragment-length polymorphism (polymorphisme de longueur des fragments amplifiés, RFLP: Restriction Fragment Length Polymorphism (profil de restriction de l'ADN génomique total), PFGE: Pulsed Field Gel Electrophoresis (électrophorèse en champs pulsé) (Li *et al.* 2009).

### Méthodes basées sur la digestion enzymatique

Il y a donc des techniques ne nécessitant pas d'amplification d'ADN par PCR comme l'électrophorèse en champs pulsé (ECP ou PFGE en anglais) et le profil de restriction de l'ADN génomique total (RFLP).

(1) Le principe de l'ECP, est une macro-restriction de l'ADN total, grâce à des enzymes de restriction à sites de coupure rares, ce qui permet d'obtenir un nombre de fragments réduit de hauts poids moléculaire. Puis, la migration par électrophorèse en champ pulsé permet donc d'obtenir un profil pour chaque souche et de détecter le polymorphisme génomique entre elles. Cette méthode a été utilisée pour le typage de souches de bactéries acétiques issues de production de vinaigre (Fernández-Pérez *et al.* 2010). Pour cette étude, l'utilisation de l'enzyme de restriction *SpeI* puis de l'ECP a montré être une méthode fiable et discriminante pour le typage des souches. Cette technique a été pendant de nombreuses années définie comme étant le « gold-standard » des méthodes de génotypage.

(2) Le principe de la RFLP, repose également sur l'utilisation d'enzymes de restriction mais avec des enzymes qui clivent l'ADN plus fréquemment que celles utilisées pour l'ECP. Cela génère plus de fragments et rend difficile leur séparation par électrophorèse. Le gel obtenu, est ainsi analysé par Southern blot, l'utilisation de sonde permet la révélation des fragments d'intérêt par hybridation.

L'inconvénient de ces deux techniques est qu'il est nécessaire d'avoir une quantité initiale d'ADN importante et de bonne qualité.

### Méthodes basées sur l'amplification d'ADN par PCR

Les méthodes basées sur l'amplification d'ADN par PCR, sont basées sur l'amplification de séquences retrouvées en de nombreuses copies et conservées dans les génomes bactériens :

(1) L'analyse multi locus du nombre d'éléments répétés en tandem (MLVA) : dans le génome bactérien, un nombre variable de répétitions en tandem (VNTR) sont dispersés, dans des séquences non codantes ou codantes, cela est une source de polymorphisme génétique pour le typage de souches par leur rapide évolution (Belkum and Scherer, 1998). En utilisant des

amorces spécifiques à ces régions, l'ADN total du génome est amplifié et des profils de bandes sont obtenus pour chaque souche selon le nombre différent de ces tandems répétés dans leur génome. Malgré la rapidité, la facilité et la reproductibilité de cette technique, le fait que ces séquences VNTR aient une évolution rapide ne permettrait pas de déterminer les relations phylogénétiques entre les souches de manière fiable.

(2) Les méthodes REP-PCR sont basées sur l'amplification de régions d'ADN bactérien situées entre des séquences répétées présentes dans le génome (Versalovic *et al.* 1991). Ces séquences répétées sont retrouvées en de nombreuses copies dispersées dans tout le génome et ont l'avantage d'être très conservées. Il existe différentes séquences répétées retrouvées dans le génome bactérien : les séquences REP (Repetitive Extragenic Palindromic element), les séquences ERIC (Enterobacterial repeated Intergenic Consensus), les BOX éléments, et les séquences (GTG)<sub>5</sub>. Les avantages de cette technique sont qu'elle est peu coûteuse, rapide à réaliser et est reproductible du fait de l'utilisation d'amorces spécifiques. Concernant les BAA, l'utilisation des séquences ERIC et (GTG)<sub>5</sub> ont démontré être les plus efficaces pour le typage de souches (Rasschaert *et al.* 2005). Lors de l'analyse des isolats de BAA issues lors de fermentation de cacao, Papalexandratou *et al.* (2009), ont conclu que les séquences répétées (GTG)<sub>5</sub> étaient suffisamment pertinentes pour déterminer les relations taxonomiques et phylogéniques entre les espèces de BAA et pouvait être appliquée pour l'analyse intra spécifique des souches.

### Méthodes couplant amplification d'ADN et digestion enzymatique

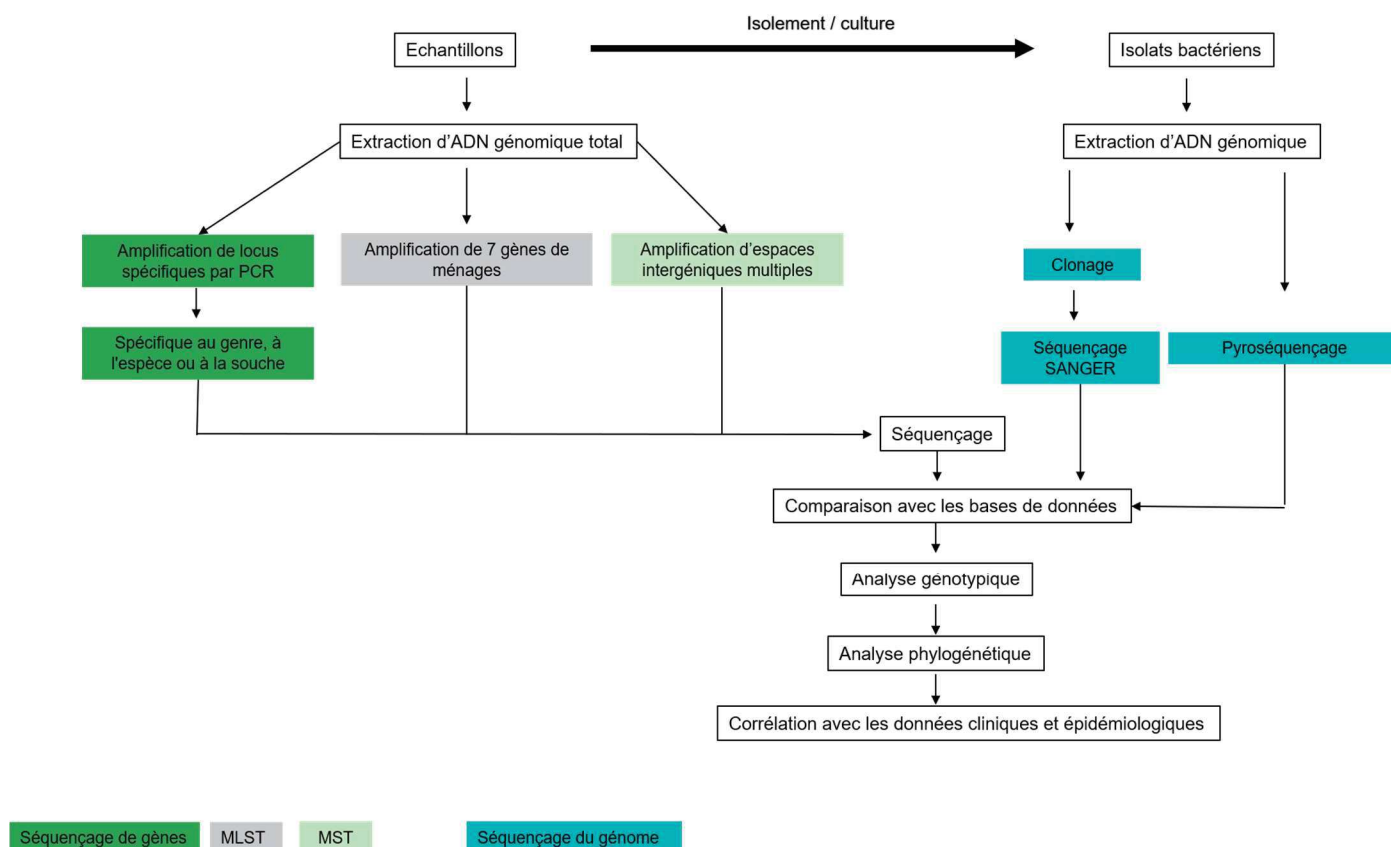
(1) La PCR-RFLP est comme son nom l'indique une combinaison entre la digestion enzymatique (1.5.2.1.1, (2)) d'un locus ciblé et amplifié par PCR, cette méthode ne nécessite pas l'étape de révélation par hybridation de sonde. Elle peut être utilisée sur les levures, les BAL et les BAA (Attchelouwa *et al.* 2018).

(2) L'AFLP (Polymorphisme de longueur des fragments amplifiés) est basée sur la digestion enzymatique de l'ADN, un adaptateur est ajouté aux fragments qui sera la cible de l'amplification par PCR. Ensuite, les profils des souches sont obtenus par migration sur gel d'électrophorèse en polyacrylamide.



## I.4.1.2.2 Méthodes basées sur le séquençage d'ADN

Dans cette catégorie de méthodes, l'analyse de la diversité intra spécifique se fait soit par le séquençage complet du génome des souches, soit par le séquençage de gènes d'intérêt (**Figure 20**).



**Figure 20:** Classification des méthodes de génotypage basées sur les séquences d'ADN pour les souches bactériennes (Li *et al.* 2009).

Cette méthode a déjà été utilisée pour des souches d'*Acetobacter* impliquées dans la fermentation du cacao : *Acetobacter aceti* (Sakurai *et al.* 2011), *Acetobacter pasteurianus* 386B (Illegheems *et al.* 2013) et *Acetobacter ghanensis* LMG 23848<sup>T</sup> et *Acetobacter senegalensis* LMG 108B (Illegheems *et al.* 2016).

Le séquençage de gènes d'intérêt est réalisé à partir du séquençage de plusieurs séquences choisies dans des gènes de ménage codant pour des protéines indispensables à la bactérie, cette technique est appelée : typage par séquençage multi locus, MLST, (Multi Locus Sequence Typing). Les séquences de ces gènes sont conservées entre les souches par le rôle essentiel de la protéine qu'elles encodent. Cependant, des variations d'un ou plusieurs nucléotides

retrouvés au sein de ces gènes apportent un polymorphisme génétique entre les souches, c'est ce que l'on étudie lors du séquençage. Ces variations sont neutres et n'affectent pas la fonction de la protéine. Généralement, 7 gènes sont séquencés et à chaque séquence différente que l'on peut obtenir entre les souches, un numéro leur est attribué. Donc chaque souche se voit donc attribuer un profil allélique à sept chiffres désignés comme type de séquence (ST) (Maiden *et al.* 1998).

Pour les BAL un schéma complet de 7 gènes élaboré pour *Lactobacillus plantarum*, avec comme gène : phosphoglucomutase (*pgm*), D-alanine-D-alanine ligase (*ddl*), la sous-unité  $\beta$  de l'ADN gyrase (*gyrB*), ATPase sous unité de phosphoribosylaminoimidazole carboxylase (*purK1*), glutamate dehydrogenase (*gdh*), Protéine de réparation de l'ADN (*mutS*) et transketolase (*tkk4*) un schéma complet avec 11 gènes (*clpX*, *dnaA*, *dnaK*, *groEL*, *murC*, *murE*, *pepX*, *pyrG*, *recA*, *rpoB*, and *uvrC*). Ces schémas sont des outils utiles pour une caractérisation précise et sans ambiguïté des isolats de *L. plantarum* et *L. fermentum* et semblent avoir un pouvoir discriminant important (De las Rivas 2006 ; Dan *et al.* 2015). Pour les BAA, un schéma incomplet existe, portant sur 3 gènes de ménage : *dnaK*, *groEL*, et *rpoB* codant pour une protéine de choc thermique, une protéine chaperonne et pour la sous unité  $\beta$  de l'ARN polymérase bactérienne (Cleenwerck *et al.* 2010; Li *et al.* 2014). Dans ces deux études, la MLST a été utilisée pour construire des arbres phylogénétiques en concaténant les séquences des 3 gènes et déterminer les relations de parenté entre les espèces de BAA.

Cependant, aucune étude n'a été réalisée sur la diversité intra spécifique des BAL et des BAA isolées lors de fermentation du cacao.

#### 1.4.2 Les bactéries lactiques (BAL)

Les BAL sont les microorganismes les plus utilisés dans la production d'aliments fermentés (produits à base de lait, viande, poisson, fruits...). Elles sont également retrouvées dans les microbiotes humains et animaux en particulier le microbiote intestinal et sont donc utilisées dans la production de probiotiques (Khandelwal *et al.* 2015).

Les BAL sont des bactéries à Gram positif, en forme de bâtonnet (bacille) ou cocci, non mobiles, à catalase négative, oxydase négative. Elles forment un groupe hétérogène car

certaines sont homofermentaires et d'autres hétérofermentaires avec production d'acide lactique, acétique et éthanol à la suite de la fermentation de glucides. Elles sont mésophiles avec un optimum de croissance de 30°C; certaines sont thermophiles avec un optimum de croissance de 37°C, mais elles sont capables de croître à des températures allant de 5°C à 45°C.

Toutes les BAL appartiennent au Phylum des *Firmicutes*, à l'ordre des *Lactobacillales* et forment un groupe très diversifié avec six familles, *Enterococcaceae*, *Leuconostocaceae*, *Lactobacillaceae*, *Streptococcaceae*, *Aerococcaceae* et *Carnobacteriaceae*.

Les quatre premières familles sont retrouvées dans les fermentations alimentaires alors que les deux dernières sont des familles non désirées dans les fermentations. Les bactéries appartenant aux familles *Aerococcaceae* et *Carnobacteriaceae* produisent des saveurs désagréables au produit fini mais également consomment les substrats nécessaires à la croissance des BAL désirées (Khandelwal *et al.* 2015).

Des milieux de cultures ont été élaborés afin de favoriser leur croissance : M17 (*Lactocoques*, *Streptocoques*), KAA mais surtout MRS (Man Rogosa Sharp). Ce sont des milieux riches car les BAL ont des besoins complexes en acide aminés, vitamine B, et bases azotées. Un pH faible est utilisé pour leur sélection (Trias Mansilla *et al.* 2008).

A partir du premier jour de fermentation du cacao, les conditions présentes au sein de la pulpe de cacao favorisent la croissance des BAL, avec un pH acide (3.5), la disponibilité des substrats et les conditions micro-aérophiles. Deux espèces sont majoritairement retrouvées dans les fermentations : *L. plantarum* et *L. fermentum* (Ardhana et Fleet 2003 ; Camu *et al.* 2007,2008 ; Papalexandratou *et al.* 2011a, 2011b, 2011c, 2013; Bartoloni *et al.* 2016, De Vuyst et Weckx 2016; Figueroa *et al.* 2019). Selon le mode de fermentation en caisse ou tas, la diversité des bactéries lactiques n'est pas affectée, comme le montre les études de Nielsen en 2007 et Visintin en 2016. Les espèces *L. fermentum* et *L. plantarum* dominent les fermentations, avec *L. fermentum* présente en début de fermentation et *L. plantarum* en fin de fermentation. D'autres espèces sont décelées dans les fermentations de cacao à travers les différents pays producteurs, *Leuconostoc pseudomesenteroides*, *Lactobacillus brevis*, *Enterococcus casseliflavus*, *Fructobacillus pseudoficulneus*, *Lactobacillus cacaonum*, *Lactobacillus fabifermentans*, *Lactobacillus ghanensis* et des espèces du genre *Pediococcus* (Nielsen *et al.* 2007; Camu *et al.* 2008; Kostinek *et al.* 2008; De Bruyne *et al.* 2009; Garcia-Armisen *et al.* 2010;

Papalexandratou *et al.* 2011a; Papalexandratou *et al.* 2011b; de Melo Pereira *et al.* 2013; Papalexandratou *et al.* 2013; Pereira *et al.* 2016) Un inventaire de la diversité des BAL a été réalisé et résumé dans le tableau 3.

**Tableau 3 :** Distribution de la fréquence des principales espèces de BAL présentes dans les fèves de cacao fermentées selon différents pays producteurs  
Ce tableau a été réalisé à partir d'une recherche dans la littérature scientifique, avec le nombre de fois que l'espèce a été identifiée/isolée dans une publication.

Lieu /pays	Australie	Bolivie	B r é s i l	Cameroun	Côte d'Ivoire	Equateur	G h a n a	Honduras	Indonésie	Malaisie	Nicaragua	Nigeria	République dominicaine	Total général
<i>Lactobacillus fermentum</i>	2	1	15	2	5	3	6	1	1	4	1	2		<b>43</b>
<i>Lactobacillus plantarum</i>	2	1	7	1	3	1	4	1	1	1	1	2	1	<b>26</b>
<i>Leuconostoc pseudomesenteroides</i>			3		2	1	2			1				<b>9</b>
<i>Lactobacillus brevis</i>			1				3	2			1		1	<b>7</b>
<i>Weissella paramesenteroides</i>				1	1		3			1				<b>6</b>
<i>Lactobacillus nagelii</i>			1	1	1	1	1							<b>5</b>
<i>Lactobacillus pentosus</i>	1		1							1	1		1	<b>5</b>
<i>Lactobacillus cacaonum</i>			1			1		2						<b>4</b>
<i>Lactobacillus ghanensis</i>				1	1		1							<b>3</b>
<i>Lactobacillus mali</i>			1				1			1				<b>3</b>
<i>Lactobacillus reuteri</i>			2				1							<b>3</b>
<i>Lactobacillus vaginalis</i>				1	1		1							<b>3</b>
<i>Fructobacillus pseudoficulneus</i>			2					1						<b>3</b>

<i>Fructobacillus sp.</i>	1				<b>1</b>
<i>Fructobacillus tropaeoli</i>		2		1	<b>3</b>
<i>Lactobacillus amylovorus</i>	1	1			<b>2</b>
<i>Lactobacillus casei</i>	1		1		<b>2</b>
<i>Lactobacillus durianis</i>	2				<b>2</b>
<i>Lactobacillus lactis</i>		1	1		<b>2</b>
<i>Lactobacillus sp</i>	2				<b>2</b>
<i>Leuconostoc mesenteroides</i>		1	1		<b>2</b>
<i>Pediococcus pentosaceus</i>	1		1		<b>2</b>
<i>Weissella confusa</i>	1		1		<b>2</b>
<i>Weissella fabaria</i>	1	1			<b>2</b>
<i>Fructobacillus ficulneus</i>		1			<b>1</b>
<i>Lactobacillus coryniformis</i>		1			<b>1</b>
<i>Lactobacillus curieae</i>		1			<b>1</b>
<i>Lactobacillus fabifermentans</i>		1			<b>1</b>
<i>Lactobacillus farraginis</i>		1			<b>1</b>
<i>Lactobacillus garvieae</i>		1			<b>1</b>
<i>Lactobacillus murinus</i>	1				<b>1</b>
<i>Lactobacillus paracasei susp paracasei</i>				1	<b>1</b>
<i>Lactobacillus pontis</i>	1				<b>1</b>

<i>Lactobacillus rhamnosus</i>			1												1
<i>Lactobacillus ruminis</i>									1						1
<i>Lactococcus lactis</i>			1												1
<i>Leuconostoc citreum</i>									1						1
<i>Leuconostoc gasicomitatum</i>									1						1
<i>Leuconostoc kimchii</i>									1						1
<i>Leuconostoc sp</i>			1												1
<i>Weissella beninensis</i>									1						1
<i>Weissella cibari</i>									1						1
<i>Weissella fabalis</i>													1		1
<i>Weissella sp</i>			1												1
<b>Total général</b>	<b>5</b>	<b>3</b>	<b>4</b>	<b>7</b>	<b>1</b>	<b>20</b>	<b>3</b>	<b>7</b>	<b>2</b>	<b>9</b>	<b>6</b>	<b>4</b>	<b>4</b>	<b>4</b>	<b>161</b>
(Ardhana et Fleet 2003; Camu <i>et al.</i> 2007,2008; Lagunez- Galvez <i>et al.</i> 2007; Kostinek <i>et al.</i> 2008; Garcia Armissen <i>et al.</i> 2010; Papalexandratou <i>et al.</i> 2011a, 2011b, 2011c, 2013, 2019; Illegheems <i>et al.</i> 2012; Pereira 2012, 2013a, 2013b; Hamdouche <i>et al.</i> 2013; Meersman <i>et al.</i> 2013; Moreira <i>et al.</i> 2013; Ho <i>et al.</i> 2014b, 2015; Bartoloni <i>et al.</i> 2016; Miescher Schwenninger <i>et al.</i> 2016; Ouattara <i>et al.</i> 2017, Bastos <i>et al.</i> 2018; Romanens <i>et al.</i> 2018; Agyirifo <i>et al.</i> 2019; Serra <i>et al.</i> 2019)															

### I.4.3 Les bactéries acétiques (BAA)

Les BAA sont des microorganismes retrouvés dans différents environnements, comme sur les fleurs (Jojima 2004; Tanasupawat *et al.* 2004; Yukphan *et al.* 2011; Lino *et al.* 2012), mais plus particulièrement dans les aliments et/ou les boissons fermentés (De Roos and De Vuyst, 2018). Leur production d'acide acétique les rend indésirables en œnologie ou zythologie car la qualité finale des produits est fortement détériorée. Cependant, leur présence est désirable, même nécessaire, pour la production de vinaigre et du cacao (Biehl *et al.* 1982).

#### I.4.3.1 Taxonomie

Elles font partie de la famille *Acetobacteraceae* dans le phylum des Alphaprotéobactéries. Dix-sept genres, dont 84 espèces et deux pseudo espèces sont retrouvés dans la famille des *Acetobacteraceae*, *Acetobacter*, *Acidomonas*, *Ameyamaea*, *Asaia*, *Endobacter*, *Frateuria*, *Gluconoacetobacter*, *Gluconobacter*, *Granulibacter*, *Komagataeibacter*, *Kozakia*, *Neosiasia*, *Neokomagataea*, *Nguyenibacter*, *Saccharibacter*, *Swaminathania*, *Swingsia* et *Tanticharoenia* (Yamada 2016).

#### I.4.3.2 Isolement des bactéries acétiques

Les bactéries acétiques sont reconnues pour être difficiles à isoler, cultiver et maintenir en culture pure (Teyssier et Hamdouche 2016). Elles ont besoin d'une grande quantité d'oxygène pour réaliser l'oxydation de l'éthanol en acide acétique. En fermentation du cacao, ce sont les brassages des fèves qui permettent cette entrée d'oxygène; En laboratoire, les cultures en aérobiose et l'agitation peuvent accélérer ce processus. De nombreux milieux de culture ont été développés pour l'isolement des bactéries et sont principalement composés des mêmes ingrédients en des concentrations différentes : le milieu AAB (pour Acetic Acid Bacteria), le milieu *Acetobacter*, le GY ou GYC (Glucose Yeast extract + carbonate de calcium), YMP (Yeast extract peptone, mannitol), DMS (deoxycholate-mannitol-sorbito) et le AE (Acetic Acid Ethanol) (Gomes *et al.*, 2018; Longin, 2016; Teyssier and Hamdouche 2016). L'ajout de carbonate de calcium (CaCO<sub>3</sub>) dans les milieux permet de détecter la production d'acide



acétique. En effet, la croissance des bactéries acétiques résulte de la dégradation du  $\text{CaCO}_3$  par l'acide acétique et forme donc un halo de dégradation autour des colonies ce qui permet de sélectionner les BAA. En œnologie, l'utilisation d'un milieu spécifique, le milieu Kneifel, qui est composé d'éthanol, d'un antifongique, d'un antibiotique et d'un indicateur de pH permet d'isoler et de détecter rapidement les bactéries acétiques. En effet, la production d'acide acétique provoque un changement de couleur du bleu au jaune/vert indiquant l'acidification du milieu (Organisme international de la vigne et du vin (OIV) 2010).

#### 1.4.3.3 Les bactéries acétiques impliquées dans la fermentation du cacao

Trois genres sont principalement retrouvés dans le processus de fermentation du cacao : *Acetobacter*, *Gluconobacter* et *Gluconacetobacter*. *Acetobacter* est le genre majoritairement isolé dominant les fermentations.

La diversité des BAA en fermentation des fèves de cacao est très limitée et est représentée en majorité par une seule espèce : *Acetobacter pasteurianus* qui domine les fermentations à travers tous les pays producteurs (De Vuyst and Weckx, 2016b; Figueroa-Hernández et al., 2019). Cependant d'autres espèces d'*Acetobacter* sont également détectées dans les différents pays producteurs mais ne dominent pas les fermentations. C'est le cas d'*A. senegalensis* et *A. ghanensis* trouvées au Ghana (Camu et al. 2008 ; Bortolini et al. 2016), Brésil (Papalexandratou et al. 2011a), Equateur (Papalexandratou et al. 2011b), Côte d'Ivoire (Bortolini et al. 2016), Cameroun (Bortolini et al. 2016) et Malaisie (Papalexandratou et al. 2013) qui seraient présentes en de fortes proportions en début du processus de fermentation. Cependant, au Brésil, l'espèce *A. tropicalis* a été isolée en tant qu'espèce dominante la fermentation; *A. syzygii*, *A. malorum*, *A. cerevisiae* et *A. ghanensis* ont été également isolées lors de ces fermentations (Pereira et al. 2012; Crafacck et al. 2013). Un inventaire bibliographique de la diversité des espèces de BAA retrouvées au cours de fermentations de fèves de cacao a été réalisé dans le cadre de la thèse et est résumé dans le **Tableau 4**.

**Tableau 4:** Distribution de la fréquence des principales espèces de BAA présentes dans les fèves de cacao fermentées selon différents pays producteurs  
Ce tableau a été réalisé à partir d'une recherche dans la littérature scientifique, avec le nombre de fois que l'espèce a été identifiée/isolée dans une publication

Espèces / Lieux	Brésil	Cameroun	Côte d'Ivoire	Equateur	Ghana	Indonésie	Malaisie	République dominicaine	Australie	Bolivie	Honduras	Total général	Références
<i>Acetobacter pasteurianus</i>	8	1	3	3	4	1	1		2	1	1	<b>23</b>	(Agyirifo <i>et al.</i> 2019; Ardhana 2003; Bortolini <i>et al.</i> 2016; Camu <i>et al.</i> 2007, 2008; Ho <i>et al.</i> 2014a, 2015; Magalhães da Veiga Moreira <i>et al.</i> 2016; Miescher Schwenninger <i>et al.</i> 2016; Mota-Gutierrez <i>et al.</i> 2018; Papalexandratou <i>et al.</i> 2011a, 2011b, 2013; Romanens <i>et al.</i> 2018; Serra <i>et al.</i> 2019)
<i>Acetobacter senegaliensis</i>	2	1	1	1	2		1			1	1	<b>10</b>	(Bortolini <i>et al.</i> 2016; Camu <i>et al.</i> 2008; Miescher Schwenninger <i>et al.</i> 2016; Papalexandratou <i>et al.</i> 2011a, 2011b, 2013; Pereira <i>et al.</i> 2013b)
<i>Acetobacter ghanensis</i>	2			1	1		1			1	1	<b>7</b>	(Cleenwerck <i>et al.</i> 2007; Camu <i>et al.</i> 2007, 2008; Garcia-Armisen <i>et al.</i> 2010; Papalexandratou <i>et al.</i> 2011a, 2013; Miescher Schwenninger <i>et al.</i> 2016; Romanens <i>et al.</i> 2018)
<i>Acetobacter fabarum</i>	2			1			1			1	1	<b>6</b>	(Cleenwerck <i>et al.</i> 2008; Garcia-Armisen <i>et al.</i> 2010; Papalexandratou <i>et al.</i> 2011a, 2011b, 2013; Romanens <i>et al.</i> 2018)

<i>Acetobacter syzygii</i>	1	1	3	2		<b>6</b>	(Bortolini <i>et al.</i> 2016; Hamdouche <i>et al.</i> 2015; Papalexandratou <i>et al.</i> 2011a)	
<i>Acetobacter tropicalis</i>	1	1	2	2		<b>5</b>	(Bortolini <i>et al.</i> 2016; Hamdouche <i>et al.</i> 2015; Pereira <i>et al.</i> 2013b)	
<i>Acetobacter lovaniensis</i>	1		2	1	1	<b>5</b>	(Lagunes Gálvez <i>et al.</i> 2007; Camu <i>et al.</i> 2007, 2008; Papalexandratou <i>et al.</i> 2011a; Hamdouche <i>et al.</i> 2015)	
<i>Acetobacter peroxydans</i>	1			1	1	<b>3</b>	(Papalexandratou <i>et al.</i> 2011a, 2011b, 2013)	
<i>Acetobacter sicerae</i>		1	1	1		<b>3</b>	(Bortolini <i>et al.</i> 2016)	
<i>Acetobacter malorum</i>	1			1		<b>2</b>	(Papalexandratou <i>et al.</i> 2011c; Pereira <i>et al.</i> 2013b)	
<i>Gluconobacter oxydans</i>	1			1	1	<b>2</b>	(Papalexandratou <i>et al.</i> 2011b, 2013; Agyirifo <i>et al.</i> 2019)	
<i>Gluconobacter frateuri</i>					2	<b>2</b>	(Ho <i>et al.</i> 2014, 2015)	
<i>Acetobacter cerevisiae</i>						1	<b>1</b>	(Romanens <i>et al.</i> 2018)
<i>Acetobacter cibirongensis</i>				1			<b>1</b>	(Papalexandratou <i>et al.</i> 2011a)
<i>Acetobacter indonesiensis</i>	1						<b>1</b>	(Papalexandratou <i>et al.</i> 2011b)
<i>Acetobacter nitrogenifigens</i>			1				<b>1</b>	(Hamdouche <i>et al.</i> 2015)

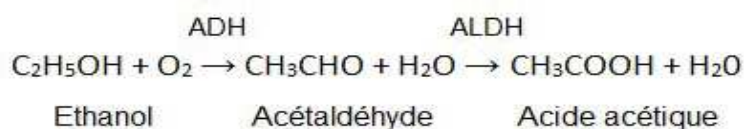
<i>Acetobacter orientalis</i>													1	(Papalexandratou, <i>et al.</i> 2011)
<i>Acetobacter orleanensis</i>	1												1	(Papalexandratou, <i>et al.</i> 2011b)
<i>Gluonoacetobacter saccharivorans</i>	1												1	(Papalexandratou, <i>et al.</i> 2011b)
<i>Gluconobacter sp</i>													1	(Papalexandratou <i>et al.</i> 2011a)
<i>Acetobacter calcoaceticus</i>												1	1	(Romanens <i>et al.</i> 2018)
<i>Acetobacter aceti</i>													1	(Ardhana 2003)
Total général	22	5	14	11	15	2	6	1	4	4	6	90		

#### 1.4.3.4 Métabolisme des bactéries acétiques

Les bactéries acétiques sont des bactéries à Gram négatif en forme de bâtonnet (bacille) ou ellipsoïdale, catalase positive, oxydase négative, parfois mobiles car elles possèdent des flagelles polaires ou péritriches. Elles ont un métabolisme strictement aérobie, avec l'oxygène comme accepteur d'électron terminal. Elles sont mésophiles avec un optimum de température de croissance compris entre 25 et 30°C (Sengun et Karabiyikli 2011; Teysier et Hamdouche 2016). Leur pH optimum pour la croissance est entre 5 et 6,5 mais elles peuvent croître à des pH inférieurs (3-4).

La particularité des BAA est qu'elles oxydent l'éthanol en acide acétique, cette réaction est catalysée par deux enzymes liées à la membrane : l'alcool déshydrogénase (ADH) et l'aldéhyde déshydrogénase (ALDH) en présence d'oxygène (Matsushita *et al.* 1994) (**Figure 21**). Les bactéries du genre *Asaia* n'oxydent pas d'éthanol et les genres *Saccharibacter* et *Granulibacter* oxydent de faibles quantités d'éthanol (Yamada *et al.* 2000). Selon la souche de BAA considérée, elles sont capables d'utiliser de nombreux substrats carbonés : acides organiques, hexoses, pentoses, alcools primaires et secondaires, aldéhydes (Longin 2016).

Les 3 genres impliqués (*Acetobacter*, *Gluconobacter* et *Gluconacetobacter*) dans les fermentations du cacao se différencient selon leurs capacités à oxyder les acides acétique et lactique. Par exemple *Gluconobacter* n'oxyde pas l'acide acétique alors que *Acetobacter* et *Gluconacetobacter* l'oxydent jusqu'au stade CO<sub>2</sub> + H<sub>2</sub>O. De plus, le genre *Acetobacter*, celui-ci dispose d'une ubiquinone de type Q-9 pour le transport d'électron dans la chaîne respiratoire alors que les autres une ubiquinone de type Q-10 (Cleenwerck and De Vos 2008).



**Figure 21:** Equation de la réaction d'oxydation de l'éthanol en acide acétique. Avec ADH : alcool déshydrogénase et ALDH : aldéhyde déshydrogénase

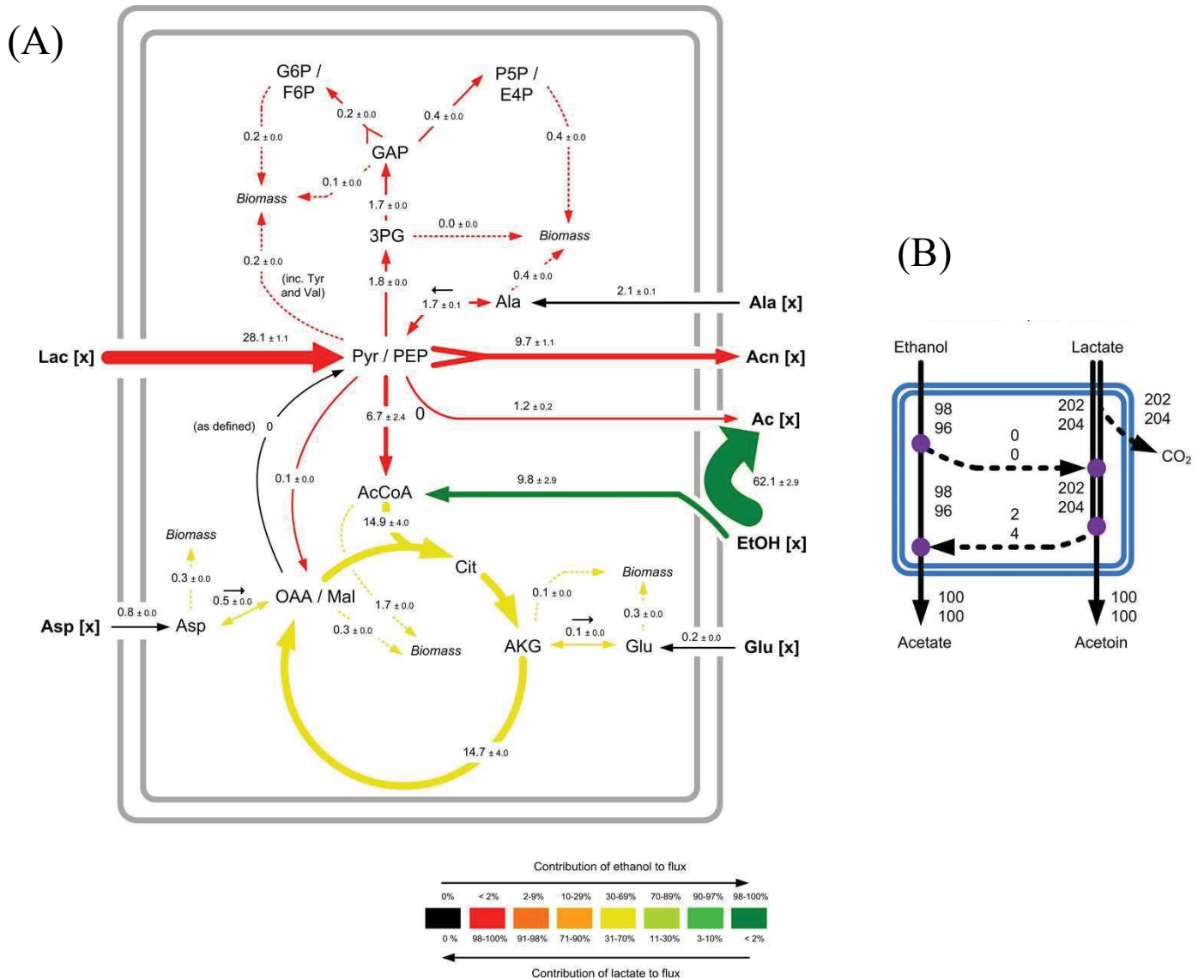
En outre, *Gluconobacter* possède un ensemble incomplet des enzymes du cycle des acides tricarboxyliques (TCA, cycle de Krebs) (Macauley *et al.* 2001; Toit et Pretorius 2002). La capacité des souches d'*Acetobacter* de suroxyder l'acide acétique est dû à une activité accrue des enzymes du cycle TCA et de l'acétyl-CoA synthétase pour un apport énergétique supplémentaire (Saeki *et al.* 1999; Sakurai *et al.* 2011; Adler *et al.* 2014; Yang *et al.* 2019).

De nombreuses études ont été réalisées afin d'élucider les flux métaboliques et leurs contributions dans la fermentation. Les génomes complets de 4 souches appartenant à des espèces différentes du genre *Acetobacter* (*Acetobacter aceti* (Sakurai *et al.* 2011), *Acetobacter pasteurianus* 386B (Illeghems *et al.* 2013), *Acetobacter ghanensis* LMG 23848<sup>T</sup> et *Acetobacter senegalensis* LMG 108B (Illeghems *et al.* 2016) ont été séquencés et ont fourni de nombreuses informations sur leurs métabolismes. En début de fermentation du cacao, les conditions d'anaérobiose ne sont pas favorables au développement des souches d'*Acetobacter*. Celles-ci doivent leur survie à leur capacité à dégrader la pectine, composant principal responsable de la viscosité de la pulpe dans la masse de cacao. En effet, lors du séquençage complet du génome d'*A. pasteurianus* 386B, la présence d'un gène codant pour l'enzyme endopolygalacturonase a été démontrée (Illeghems *et al.* 2013). Ces souches pourraient participer à la dépectinisation de la pulpe. Il a été démontré que les souches d'*Acetobacter* analysées partagent un unique métabolisme. Les bactéries du genres *Acetobacter* ont une voie Embden-Meyerhof-Parnas (glycolyse) incomplète dû à l'absence du gène codant la phosphofructokinase. Cependant, tous les gènes codants pour la voie de l'oxydation des pentoses phosphate sont retrouvés dans les génomes d'*A. pasteurianus* et *A. ghanensis* LMG 23848<sup>T</sup>, suggérant que le glucose est dégradé par cette voie. Pour *A. senegaliensis* 108B, l'absence du gène codant pour la gluconate kinase indique que cette souche a perdu la capacité à métaboliser le glucose. Des gènes codant pour des acteurs du cycle de Krebs sont retrouvés dans les 4 génomes suggérant l'implication de cette voie pour la dégradation des sucres pour la production d'énergie. L'utilisation du mannitol, du glycérol et de l'acide lactique comme source d'énergie en plus de l'éthanol a également été mise en évidence lors de ces études.

Afin d'élucider quels sont les apports de l'éthanol produit par les levures et de l'acide lactique produit par les bactéries lactiques, Adler *et al.* (2014) ont réalisé un marquage par isotope de ces deux substrats lors de suivis de fermentations. Ce marquage permet d'identifier les flux

intracellulaires lorsqu'ils sont introduits comme substrat premier dans le milieu simulant la pulpe de cacao lors de la fermentation (CPSM-AAB) (**Figure 22A**).

Cette étude montre que l'acide acétique est principalement synthétisé à partir de l'éthanol (96 à 98 %) et que seulement 2 à 4 % de la quantité finale d'acide acétique produite provient de l'acide lactique (**Figure 22B**). L'éthanol sert donc pour la production d'acide acétique mais, il est également la principale source d'énergie des bactéries. Le lactate quant à lui permet la production d'acétoïne via le pyruvate, de la biomasse et est impliqué dans la voie oxydative des pentoses phosphates. De toutes ces études, il a été mis en évidence que les *Acetobacter* présentent un métabolisme fonctionnel séparé en deux lors de la co-consommation de ces deux substrats à deux carbones (éthanol) et trois carbones (acide lactique). Cela est démontré lors de la production d'acides aminés, tous les acides aminés provenant d'intermédiaires en amont du pyruvate sont produits à l'aide de l'acide lactique, l'éthanol n'est pas impliqué dans leur production. L'éthanol et l'acide lactique sont deux substrats importants pour le métabolisme des *Acetobacter*, c'est pourquoi l'équilibre microbien dans la fermentation est important afin d'avoir un ratio équilibré entre les deux substrats pour une croissance optimale des *Acetobacter*.



**Figure 22:** (A) Flux métaboliques de la souche *A. pasteurianus* en milieu simulant les conditions de la pulpe de cacao lors de la fermentation. (B) Contribution en pourcentage du lactate et de l'éthanol dans la formation de l'acide acétique et de l'acétoïne (Adler *et al.* 2014).

#### 1.4.4.5 Organisation génomique des *Acetobacter pasteurianus*

*Acetobacter pasteurianus* est aussi retrouvée lors de la production de vinaigre, c'est pourquoi des études se sont focalisées sur l'explication de cette domination et quelles sont les particularités génomiques et génétiques de ces souches.



Le génome des souches d'*A. pasteurianus* est constitué d'un chromosome circulaire, d'environ 2,9 Mb et présente un nombre de plasmide variable selon la souche. Son chromosome est conservé entre les souches avec un fort degré d'homologie et est stable d'une génération à l'autre, mais peut être flexible sur une échelle de temps évolutive (Sokollek *et al.* 1998; Steiner et Sauer 2001). En effet, lorsque les souches sont maintenues dans des environnements dépourvus en acide acétique, les bactéries acétiques peuvent perdre rapidement leur capacité à résister dans des environnements hautement acides.

Le chromosome est relativement conservé alors que les plasmides peuvent être différents selon la souche, c'est ce qui participe à cette distinction génétique entre les souches (Wang *et al.* 2015).

*A. pasteurianus* est connue pour être une bactérie hyper mutable, son instabilité est aussi due aux nombreux réarrangements du génome (délétion, duplication de gènes) causés par des mutations spontanées ou par des éléments génétiques mobiles. La plupart de ces mutations sont silencieuses, elles n'ont aucun effet, mais d'autres ont conduit à des variations de phénotypes, à de la spéciation. La présence de nombreux tandems répétés hyper mutables et de transposons est une des principales sources de cette instabilité et des variations phénotypiques. D'après Azuma *et al.* (2009), 9 % du génome serait composé d'éléments transposables, avec la séquence d'insertion IS1380 comme étant la plus présente. Les principales conséquences de ces éléments transposables sont la résistance à l'acide acétique et la perte de la fonction d'oxydation de l'éthanol (Beppu 1994; Takemura *et al.* 1991). Grâce à cette flexibilité et instabilité génétique, les souches d'*A. pasteurianus* ont pu évoluer et s'adapter aux niches compétitives que l'on peut trouver dans les environnements fermentés. Cependant, certaines souches ont développé un mécanisme de protection contre ces insertions. Il y a une balance entre l'instabilité causée par les insertions et ce mécanisme de protection qui confère une certaine stabilité. Toutes les souches ne possèdent pas les mêmes capacités de protection, cette stabilité conférée par ce système de protection est variable d'une souche à l'autre (Wang *et al.* 2015).

#### 1.4.4 Starter de fermentations

Pour de nombreuses matrices alimentaires ou boissons, les fermentations ne se font plus de manière spontanée ; Mais contrôlée par l'industriel. Afin d'induire la fermentation une inoculation d'un cocktail d'espèces levuriennes et / bactériennes est réalisée au début des

fermentations. Ces starters sont souvent composés de souches qui interviennent naturellement dans les fermentations spontanées. C'est le cas pour la bière, le vin et le pain. De nombreuses études ont été menées afin de connaître les communautés microbiennes impliquées dans ces fermentations et de réaliser les inocula les plus adaptés. Les principaux avantages des fermentations inoculées sont une fermentation plus rapide, plus uniforme et un produit de qualité plus constant (Heard et Fleet 1993).

Dans le cas de la fermentation des fèves de cacao, la fermentation se déroule de manière artisanale et est spontanée, les fèves sont inoculées par l'environnement qui entoure l'ouverture des cabosses et les outils de fermentations (outils, mains, caisses de fermentations...) (Camu *et al.* 2007). Cependant, de premières études ont été menées afin d'établir un starter de fermentation, composé d'un cocktail de souches souvent retrouvées dans les fermentations à travers les pays producteurs. Le premier starter de culture a été élaboré par Schwan (1998), il était composé de *S. cerevisiae var. chevalieri*, *Lactobacillus lactis*, *L. plantarum*, *Acetobacter aceti* et *G. oxydans subsp. suboxydans*. Une analyse comparative entre fermentations inoculées et non inoculées (avec la flore naturelle) a été réalisée. Elle a démontré qu'il était possible de mimer les conditions de fermentations et la qualité sensorielle du cacao final à l'aide de cocktail de souches. Les études réalisées ces 20 dernières années sur la diversité microbienne ont permis de découvrir l'existence du consortium microbien et donc de réaliser des starters de cultures en adéquation avec ces découvertes.

Lefeber *et al.* (2012) ont proposé de contrôler les fermentations à l'aide de starter de cultures composé d'une souche de bactérie lactique de *L. fermentum* 222 et une bactérie acétique *A. pasteurianus* 386B dans un cas et avec ajout d'une levure *S. cerevisiae* H5S5K23 dans l'autre, ces souches ont été isolées lors de fermentations de fèves de cacao Ghanéennes. Après leur culture dans des milieux spécifiques, les souches ont été lyophilisées. Des fèves de cacao en provenance du Ghana, de la Malaisie et de la Côte d'Ivoire ont été inoculées avec ce starter avant la mise en fermentation. L'inoculation se fait avec les souches lyophilisées re suspendues dans du sérum physiologique stérile, à hauteur de 5 log de CFU (colony forming unit, unité formant une colonie). Dans les trois cas, le chocolat final produit est de meilleure qualité que le cacao produit lors de fermentation spontanée et est plus apprécié par les consommateurs.

Des résultats similaires ont été obtenus par Crafacck *et al.* (2013) qui ont étudié l'impact potentiel de *P. kluyveri* (souche avec un fort potentiel aromatique) et *Kluyveromyces marxianus* (souche avec un fort potentiel pectinolytique) sur l'arôme du cacao pendant la fermentation en association avec une souche de bactérie lactique *L. fermentum* L18 et avec une souche de bactérie acétique *A. pasteurianus* A149. Ils ont conclu que l'utilisation de ces souches comme starter de culture semblait avoir un effet positif sur le profil de saveur par rapport aux fermentations spontanées. Les cacaos réalisés avec ce cocktail de souches ont obtenu les scores les plus élevés pour les descripteurs sensoriels tels que la douceur, le fruité, le goût du cacao et l'appréciation générale du cacao final selon un panel de participants non formés. Dans ce cas, l'inoculation des fèves de cacao avec le starter s'est faite à hauteur de  $5 \cdot 10^5$  cellules par gramme pour les AAB et LAB et  $5 \cdot 10^6$  cellules par gramme pour les levures après culture de ces dits starters, l'inoculation des 3 groupes se fait de manière spontanée.

Au Brésil, Magalhães da Veiga Moreira *et al.* (2017) ont démontré l'impact positif de l'utilisation d'un cocktail microbien comme starter de culture sur la fermentation du cacao et la saveur du cacao final. Ils ont proposé une souche de levure, *Saccharomyces cerevisiae* UFLA CCMA 0200, une souche de bactérie lactique, *Lactobacillus plantarum* CCMA 0238 et une souche de bactérie acétique *Acetobacter pasteurianus* CCMA 0241, toutes trois isolées lors de fermentation du cacao au Brésil. L'inoculation se fait à partir de lyophilisats re suspendus dans une solution eau peptonée stérile, répartis en même temps sur les fèves de cacao atteignant une concentration d'environ  $10^5$  et  $10^7$  de cellules pour les AAB/LAB et les levures respectivement, par gramme de cacao. En comparaison avec une fermentation spontanée, le processus de fermentation est accéléré lors de l'inoculation. Il y a un changement dans les communautés microbiennes associées aux fermentations, dû aux souches introduites qui dominent la fermentation, mais dans les deux cas on retrouve les principales espèces microbiennes impliquées dans les fermentations du cacao. De plus, l'inoculation permet également une amélioration de la qualité du cacao final avec une plus forte amertume, un goût plus sucré et la saveur cacao plus dominante que sans inoculation.

## 1.5 Conclusion

De nombreuses études ont été réalisées pour mieux comprendre la diversité interspécifique des bactéries lactiques et acétiques impliquées dans la fermentation du cacao. Il a été

démontré que selon la zone géographique, un consortium microbien, composé de levures, de bactéries lactiques et acétiques est présent, mais que certaines espèces ne sont présentes que dans une origine géographique. Il est nécessaire d'étudier d'avantage la diversité des différentes souches isolées de ces principaux acteurs de la fermentation, notamment d'un point de vue intra spécifique, afin d'éviter de standardiser les fermentations à travers tous les pays producteurs. Pour trouver des souches de levures, de bactéries acétiques et lactiques propres à chaque pays afin de préserver la diversité des cacaos mondiaux, lors de l'élaboration de starter de culture.

## Chapitre II : Matériels et méthodes

## II.1 Echantillonnage des fèves de cacao

Deux catégories de fèves de cacao ont été utilisées au cours de cette étude : des échantillons de fèves fermentées puis séchées au soleil ainsi que des échantillons de fèves fraîches à différents temps de fermentation avec la présence de pulpe, en provenance de différents pays producteurs de cacao à travers le monde : Côte d'Ivoire, Mexique, Nicaragua, Pérou, Guyane (**Figure 23**). Lors de toutes les collectes, les prélèvements ont été réalisés à l'aide de gants.

**Chaque pays a des pratiques de récolte, de mise en fermentation et de déroulement de la fermentation différentes.**



**Figure 23** : Carte du monde indiquant les pays d'origine des échantillons des fèves fraîches (nom du pays en noir) et des fèves séchées (nom en rouge)

### II.1.1 Fèves de cacao séchées

Les échantillons de fèves de cacao sèches proviennent de trois pays : Nicaragua, Pérou et Côte d'Ivoire (**Tableau 5**). Les fèves ont été stockées au laboratoire dans des sachets stériles fermés hermétiquement à température ambiante.

**Tableau 5:** Liste détaillée des échantillons de fèves séchées.

Origine	Ville	Date de la collecte	Variété cacao	Récolte/écabossage	Type de fermentation	Conservation au laboratoire	Nombre d'échantillons	Fournisseur
Pérou	Tingo Maria	07/16	CCN-51	Machette	Caisse	Température ambiante	3	L. Berthiot*
Pérou	Tingo Maria	06/17	CCN-51	Machette	Caisse	Température ambiante	2	L. Berthiot*
Nicaragua	-	12/16	Trinitario (80%) Criollo (20%)	-	Caisse	Température ambiante	4	L. Berthiot*
Côte d'Ivoire	Abidjan	01/17	Forastero	Gourdins	Tas (feuilles de bananiers)	Température ambiante	36	T.Guehi**

\* Collaborateur du Cirad, Société Lallemand

\*\* Chercheur à l'université de Nangui Abrogoua, Abidjan – Côte d'Ivoire

#### - Echantillonnage au Pérou et Nicaragua

Les échantillons de fèves proviennent d'essais de fermentation menés par la société Lallemand avec une inoculation de souche de levure différente sélectionnée pour chaque échantillon (données confidentielles). Nous disposons également d'un témoin sans inoculum pour les deux pays (fermentation spontanée). Après une fermentation de 5 jours pour les fèves du Pérou et 8 jours pour celles du Nicaragua, les fèves ont été séchées au soleil.

En ce qui concerne le Pérou, 5 échantillons en provenance de deux campagnes de prélèvements en 2016 (P-1, P-2, PT) et 2017 (P-3, P-4) ont été utilisés. La variété de cacao utilisée est du CCN-51 (un hybride résistant à différents pathogènes né du croisement des variétés Iquitos, Criollo et Amelonado). Six cents kilogrammes de fèves de cacao ont été fermentées dans des caisses en bois, séparées en deux à l'aide de film plastique. L'écoulement des jus après l'écabossage a été réalisé pendant 15 h, puis une étape de pré-séchage solaire de 3 h a été réalisée avant l'inoculation et la mise en fermentation. Un brassage des fèves a été effectué au bout de 48 h puis toutes les 24 h de fermentation.

Pour le Nicaragua, 4 échantillons ont été collectés lors de la récolte de 2016 (N-1, N-2, N-3, N-T). La variété de cacao utilisée est un mélange de Trinitario à 80% et de Criollo à 20%. La fermentation qui a eu lieu dans des caisses en bois a duré 8 jours après inoculation. Les fèves ont été aussi brassées après 48 h de fermentation ensuite toutes les 24 h.

#### - Echantillonnage en Côte d'Ivoire

Les cabosses proviennent de quatre bassins producteurs de cacao (Daloa, Akoupé, Djékanou et Soubré-Gran Zatry). La collecte a été réalisée sur trois fermes par bassin avec 250 cabosses collectées. Cependant les fermentations n'ont eu lieu que sur un site de fermentation situé à Akoupé (**Figure 24**) :

- Daloa, située au Centre-Ouest à 400 km d'Abidjan
- Akoupé, au Sud-Est à 150 km d'Abidjan (site de fermentation)
- Djékanou, au Centre-Est à 250 km d'Abidjan
- Soubré-Gran Zatry, Sud/Centre-Ouest et à 250 km d'Abidjan.



**Figure 24:** Carte de la Côte d'Ivoire indiquant les bassins producteurs et le site de fermentation.

Vingt-cinq à vingt-six kg de fèves pour chaque ferme ont été mis en fermentation en tas sur feuilles de bananier pendant 5 jours au même site de fermentation (Akoupé), avec un brassage manuel des fèves à 48 h et 96 h de fermentation. Cinq cents grammes de fèves de cacao ont



été prélevés à trois temps de fermentations différents ; 2, 3 et 4 jours et séchés au soleil pendant 3 jours. Au total 36 échantillons de fèves ont donc été prélevés.

Pour l'ensemble des échantillons de fèves sèches, l'activité de l'eau ( $A_w$ ) a été mesurée à l'aide d'un  $A_w$  mètre préalablement calibré à 0,75 à 25°C avec une solution de NaCl (358,5 g/L). Elle a été mesurée sur dix grammes d'échantillon de fèves broyées. Pour la croissance des bactéries une  $A_w$  supérieure à 0,9 est nécessaire.

### II.1.2 Fèves de cacao fraîches avec la pulpe

Nous disposons d'échantillons de fèves de cacao fermentées en provenance de 3 pays : la Guyane Française, la Côte d'Ivoire et le Mexique (**Tableau 6**).

**Tableau 6:** Liste détaillée des échantillons de fèves fraîches avec pulpe.

Origine	Ville	Date de la collecte	Variété cacao	Récolte /écabossage	Fermentation	Conservation Au laboratoire	Nombre d'échantillons	Fournisseur
Guyane TX_G	Combi	12/17	Mélange de variétés	Machette	Caisse (polystyrene)	Eau+ glycérol (50/50, v/v) T= -20°C	7	G. Fourny*
Côte d'Ivoire JX	Abidjan	05/18	Mercedes	Gourdin	Tas	Eau+ glycérol (50/50, v/v) T= -20°C	7	T. Guehi***
Mexique TXCY a/b	Comalcalco	02/18	Criollo	Machette	Caisse (bois)	Bouillon (MRS) + glycérol (50/50 v/v) T= -80°C	12	M. Suarez Quiroz**
Mexique CXTY	Comalcalco	10/18	Criollo	Machette	Caisse (bois)	Eau+ glycérol (50/50, v/v) T= -20°C	8	L. Farrera (Séjour de 3 semaines)

\* Chercheur dans l'UMR QualiSud

\*\* Chercheur à l'Institut technologique de Veracruz, Mexico

\*\*\* Chercheur à l'Université Nangui Abrogoua, Abidjan – Côte d'Ivoire

X : nombre de jours de fermentation

Y : caisse de fermentation

a/b : répliques pour chaque caisse de fermentation

TX-G, JX, TXCY a/b et CXTY est le code pour les extractions d'ADN en provenance de 4 campagnes d'échantillonnage.

#### - Echantillonnage en Guyane Française

Nous disposons de 7 échantillons de fèves fraîches provenant d'un suivi de fermentation réalisé sur 7 jours avec un prélèvement chaque 24 h. La récolte et la fermentation ont eu lieu en décembre 2017 à la plateforme expérimentale du Cirad située à Combi (Guyane). La station

expérimentale du Cirad dispose d'une collection importante de cacaoyers, avec une plantation de 1,5 hectares comptant 1600 arbres (Guyanais primitifs, sauvages et une collection de matériel international). Les fermentations ont été réalisées avec un mélange de fèves de variétés différentes en provenance de cette collection. Le délai entre la récolte et l'écabossage a été de 1 jour. Quinze à vingt kilogrammes de fèves ont été mis à fermenter dans des caisses en polystyrène qui sont recyclées d'une fermentation à l'autre. La fermentation a duré 7 jours et des brassages à la main ont été réalisés tous les deux jours jusqu'à la fin de la fermentation. Quinze à vingt grammes de fèves ont été prélevés toutes les 24 h après la mise en fermentation. Elles ont été transférées dans un flacon stérile de 50 mL avec ajout de 10-15 mL d'un mélange eau/glycérol (v/v : 50/50) puis congelés à  $-20^{\circ}\text{C}$ .

Le code utilisé pour les échantillons est le suivant TX-G, avec X représentant le jour de fermentation.

#### - **Echantillonnage en Côte d'Ivoire**

Nous disposons de 7 échantillons de fèves fraîches provenant d'un suivi de fermentation réalisé sur 7 jours avec un prélèvement chaque 24 h.

La récolte a eu lieu le 28 mai 2018 dans la ville d'Akoupé au Sud-Est de la Côte d'Ivoire à 150 km d'Abidjan, la variété de cacao est du Mercedes. Le délai entre la récolte et l'écabossage a été de 1 jour. Vingt kilogrammes de fèves fraîches ont été mis en fermentation en tas sur des feuilles de bananiers. La fermentation a duré 6 jours et des brassages manuels des fèves ont été réalisés à 48h et 95h de fermentation. Quinze à vingt grammes de fèves ont été prélevés toutes les 24h après la mise en fermentation, elles ont été transférés dans un flacon stérile de 50 mL avec ajout de 10-15 mL d'un mélange eau/glycérol, (v/v : 50/50) puis congelés à  $-20^{\circ}\text{C}$ .

Le code utilisé pour les échantillons est le suivant JX, avec X représentant le jour de fermentation.

#### - **Echantillonnage au Mexique**

Deux campagnes de prélèvements ont été réalisées, une en février 2018 (période de grande récolte) et une en octobre 2018 (petite récolte) provenant d'un « beneficio » mexicain, une coopérative de plantations de cacao où les fèves sont fermentées, séchées puis stockées

située dans la région de Tabasco, à Comalcalco, à 475 km de Veracruz. La variété de cacao utilisée pour les fermentations est du Criollo. Les cabosses ont été récoltées à l'aide de machettes au champs et écabossées également à l'aide de machettes. Les fermentations se sont déroulées dans des caisses en bois. Un premier brassage à l'aide d'une pale en bois a été réalisé au bout de 48 h puis toutes les 24 h. Les fermentations ne durent pas plus de 5 jours généralement dû à la faible quantité de pulpe dans les cabosses de cacao.

Pour la première campagne, des échantillons de fèves avec la pulpe ont été prélevés à 3 temps de fermentations différents (2, 4 et 5 jours) dans deux caisses de fermentations (réalisées en duplicata). Les fèves ont été mises en culture 48 h dans du bouillon MRS (Man-Rogosa Sharp, milieu sélectif des bactéries lactiques) et dans un bouillon MRS enrichi en alcool (pour les bactéries acétiques). Du glycérol a été ajouté au bouillon (50/50, v/v) pour la conservation à -80 °C. Ces cultures ont donc été réalisées sur place au Mexique à l'institut technologique de Veracruz, grâce à l'aide du Dr Mirna Suarez-Quiroz. Les extractions d'ADN sur les fèves de cacao et pulpe de ces échantillons ont aussi été réalisés au Mexique. Les cultures et les extractions d'ADN ont été transportées de manière réfrigérée (dans une glacière avec des blocs de glace) jusqu'à Montpellier.

Le code utilisé pour les échantillons est le suivant TXCY a/b, avec X représentant le jour de fermentation, Y la caisse de fermentation et a/b représentant les répliques pour chaque caisse de fermentation.

Pour la seconde campagne, j'ai effectué une mission de 3 semaines (financée par la bourse DUFRENOY en partenariat avec le Crédit Agricole d'Ile-de-France) à Comalcalco afin de prélever des fèves de cacao, de réaliser les fermentations et les extractions d'ADN des prélèvements d'échantillons aux différents temps de fermentation. La récolte a eu lieu le 23 octobre 2018, le délai entre l'écabossage et la mise en fermentation a été de 1 jour. Trois caisses de fermentations ont été ainsi réalisées : deux à 25 kg (caisses 1 et 2) et une à 20 kg (caisse 3). Seules les caisses 1 et 2 ont été échantillonnées pour l'isolement bactérien et les caisses 1, 2 et 3 pour les extractions d'ADN. Un suivi de fermentation a été réalisé du temps 0 au temps 4 avec prélèvement de 150 g de fèves. Pour le transport jusqu'en France, 15 à 20 g de fèves ont été conservés à -20°C dans du glycérol + sérum physiologique à 9 g/L (50/50, v/v, stérile). Les isollements bactériens ont été faits sur place au Mexique. Deux millilitres de

chaque bouillon de culture réalisés au Mexique pour l'isolement ont été aussi conservés à -80°C dans du glycérol (50/50, v/v).

Le code utilisé pour les échantillons est le suivant CXTY, avec X la caisse de fermentation et Y le jour de fermentation.

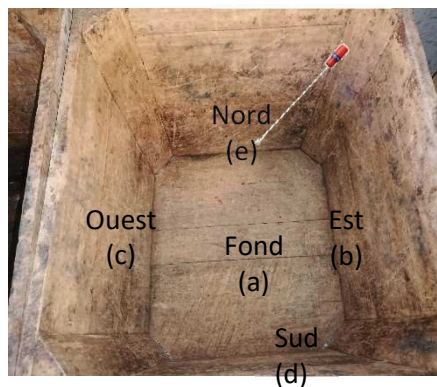
## II.2 Echantillonnage environnemental

Lors de la seconde campagne d'échantillonnage de fèves de cacao au Mexique (Octobre 2018), 57 échantillons environnementaux ont également été prélevés par moi-même. Ces prélèvements ont été réalisés dans l'environnement au niveau du champ lors de la récolte et au niveau du site de fermentation afin d'analyser les communautés bactériennes associées. Les prélèvements ont été effectués à l'aide d'écouvillons stériles, trempés dans une solution de Tween 20 à 1 % (avec NaCl à 0,9 %). La surface des objets ou mains a été échantillonnée par frottement jusqu'à l'assèchement de l'écouvillon puis la partie en coton a été découpée, déposée dans un micro tube et conservée à -20°C jusqu'à l'extraction d'ADN (**Figure 25**).



**Figure 25:** Méthode d'échantillonnage avec des écouvillons stériles.

Nous avons échantillonné au niveau du champ de cacaoyers lors de la récolte : l'écorce de 3 arbres, 3 fleurs, 3 feuilles, 3 cabosses, les mains gauches et droites des deux récolteurs (droitiers) et la machette avec laquelle les cabosses sont coupées de l'arbre mais également les sacs dans lesquels les cabosses ont été transportées du champ au lieu de fermentation. Au niveau de l'environnement de la fermentation, les prélèvements ont été réalisés sur : la machette à écabosser, le sol où ont été posées les cabosses et les fèves après l'écabossage, les mains gauches et droites des 3 hommes qui ont écabossé (droitiers), les 3 caisses avant la mise en fermentation (côté Sud, Nord, Est, Ouest et le fond, **Figure 26**), le seau pour peser les fèves, la pêle qui permet de brasser les fèves et les sacs couvrant les caisses. Les prélèvements sont récapitulés dans le **tableau 7**.



**Figure 26:** Prélèvement des caisses avant la mise en fermentation.

**Tableau 7:** Liste et code des prélèvements environnementaux réalisés au Mexique en Octobre 2018.

<b>M1a=</b> Main droite récolteur 1	<b>C1d=</b> Caisse de fermentation Sud
<b>M1b=</b> Main gauche récolteur 1	<b>C1e=</b> Caisse de fermentation Nord
<b>M2a=</b> Main droite récolteur 2	<b>C2a=</b> Caisse de fermentation fond
<b>M2b=</b> Main gauche récolteur 2	<b>C2b=</b> Caisse de fermentation Est
<b>M3a=</b> Main droite écabosseur 1	<b>C2c=</b> Caisse de fermentation Ouest
<b>M3b=</b> Main gauche écabosseur 1	<b>C2d=</b> Caisse de fermentation Sud
<b>M4a=</b> Main droite écabosseur 2	<b>C2e=</b> Caisse de fermentation Nord
<b>M4b=</b> Main gauche écabosseur 2	<b>C3a=</b> Caisse de fermentation Fond
<b>M5a=</b> Main droite écabosseur 3	<b>C3b=</b> Caisse de fermentation Est
<b>M5b=</b> Main gauche écabosseur 3	<b>C3c=</b> Caisse de fermentation Ouest
<b>Ca1=</b> Cabosse 1	<b>C3d=</b> Caisse de fermentation Sud
<b>Ca2=</b> Cabosse 2	<b>C3e=</b> Caisse de fermentation Nord
<b>Ca3=</b> Cabosse 3	<b>C3f=</b> Caisse -> toile d'araignée
<b>Ma1a=</b> Machette 1 (récolte)	<b>Z1=</b> Sac de transport des cabosses 1
<b>Ma1b=</b> Machette 1 (récolte)	<b>Z2=</b> Sac de transport des cabosses 2
<b>Ma2a=</b> Machette 2 (récolte)	<b>Z3=</b> Sac de transport des cabosses 2
<b>Ma2b=</b> Machette 2 (récolte)	<b>Pa=</b> Pale pour faire les brassages (côté 1)
<b>Fe1=</b> Feuille 1 (cacaoyer)	<b>Pb=</b> Pale pour faire les brassages (côté 2)
<b>Fe2=</b> Feuille 2 (cacaoyer)	<b>G1=</b> Sac polypropylène recouvrant caisse 1
<b>Fe3=</b> Feuille 3 (cacaoyer)	<b>G2=</b> Sac polypropylène recouvrant caisse 2
<b>F1=</b> Fleur 1	<b>G3=</b> Sac de jute recouvrant caisse 3
<b>F2=</b> Fleur 2	<b>Mq1=</b> Machette pour ouvrir les cabosses (côté 1)
<b>F3=</b> Fleur 3	<b>Mq2=</b> Machette pour ouvrir les cabosses (côté 2)
<b>T1=</b> Tronc 1	<b>Tma=</b> Fond de la boîte pour peser les fèves
<b>T2=</b> Tronc 2	<b>Tmb=</b> Intérieur de la boîte pour peser les fèves
<b>T3=</b> Tronc 3	<b>S1=</b> Sol (autour de l'écabossage)
<b>C1a=</b> Caisse de fermentation fond	<b>S2=</b> Sol (autour de l'écabossage)
<b>C1b=</b> Caisse de fermentation Est	<b>S3=</b> Sol (autour de l'écabossage)
<b>C1c=</b> Caisse de fermentation Ouest	

## II.3 Analyse des communautés microbiennes associées aux fèves en fermentation et à l'environnement de la fermentation

### II.3.1 Extraction, mesure de la qualité et de la quantité d'ADN total

#### II.3.1.1 Extraction des ADN par la méthode de « phénol/chloroforme »

L'extraction des ADN des microorganismes présents à la surface des fèves a été réalisée selon la méthode d'extraction au phénol/chloroforme/alcool iso amylique développée au laboratoire (Ampe *et al.* 1999 ; Leesing *et al.* 2005 ; Le Nguyen *et al.* 2008 ; El Sheikha *et al.* 2009).

Pour que les microorganismes présents à la surface des fèves de cacao se détachent, 10 grammes d'échantillon ont été mis sous agitation pendant 2 h dans du sérum physiologique

stérile supplémenté de Tween 80 (0,1 %). Un millilitre de chaque solution de lavage a été prélevé et centrifugé pendant 15 min à 12 000 g. Le culot obtenu a été alors suspendu dans 100 µL de TE (10 mM Tris-HCl, 1 mM EDTA, pH 8,0 (Promega) et l'extraction a commencé avec les étapes de lyse cellulaire. Cent microlitres de solution de lysozyme (25 mg/mL, Eurobio) ont été ajoutés puis incubés à température ambiante pendant 5 min puis 100 µL de protéinase K (20 mg/mL, Biosolve) ont été ajoutés. Après agitation manuelle, les échantillons ont été mis à 42 °C pendant 20 min. Cinquante microlitres de SDS 20% (Dodécylsulfate de sodium) ont été ajoutés puis les échantillons ont été mis à 42 °C pendant 10 min. Quatre cents microlitres de CTAB (Hexadecyltrimethylammonium bromure) ont été versés dans chaque tube puis mis à 65 °C pendant 10 min. Les lysats ainsi obtenus ont subi deux extractions au phénol/chloroforme. Sept cents microlitres de phénol/chloroforme/alcool iso amylique (25/24/1, v/v/v) ont été ajoutés au lysat, agités manuellement et ensuite centrifugés à 12 000 g pendant 15 min. La phase aqueuse qui correspond à la phase supérieure contenant les acides nucléiques a été alors recueillie dans un nouveau micro tube après chaque étape de centrifugation, en évitant de prélever l'interphase qui contient les protéines précipitées et en évitant de prélever la phase organique (inférieure). Le phénol résiduel a été éliminé par extraction avec 600 µL de solution de chloroforme/alcool iso amylique puis centrifugé 12000 g pendant 10 min. Enfin, l'extraction s'est terminée par les étapes de précipitation et de purification de l'ADN. Trente microlitres d'acétate de sodium (3 M, pH 5) ont été ajoutés à la phase aqueuse recueillie, après l'élimination du phénol résiduel, un volume d'isopropanol glacé a été ensuite ajouté. Le mélange a été agité manuellement puis mis pendant 12 h à -20°C. Les micro-tubes ont été centrifugés à 12 000 g pendant 30 min, le surnageant a été éliminé et le culot lavé avec 500 µL d'éthanol à 70 % afin d'éliminer les molécules organiques. Les micro-tubes ont été alors centrifugés à 12 000 g pendant 10 min. L'éthanol a été ensuite éliminé et les culots séchés à l'air libre sous la hotte jusqu'à évaporation totale de l'éthanol. Les culots ont été remis en suspension dans 50 µL de TE puis stockés une nuit à 4°C et enfin mis à -20°C jusqu'à l'analyse.

### *II.3.1.2 Détermination de la quantité et de la qualité des ADN*

La pureté et la quantification des ADN extraits ont été déterminées à l'aide d'un spectrophotomètre Nanodrop (BiospecNano). La concentration en ADN est mesurée à la longueur d'onde de 260 nm, et la pureté de l'ADN est estimée à l'aide de deux ratios. Le ratio 260/280 permet de rendre compte de la teneur en acides nucléiques et protéines. Il doit être compris entre 1,8 et 2; s'il est supérieur à 2 cela veut dire qu'il y a présence d'ARN dans l'échantillon et s'il est inférieur à 1,8, cela indique une teneur en protéine élevée. Le ratio 260/230 correspond au ratio : acides nucléiques/inhibiteur de PCR, celui-ci doit être compris entre 1,6 et 1,8.

L'ADN a été aussi quantifié à l'aide d'un fluoromètre Qubit 3.0<sup>®</sup> et du kit Qubit<sup>®</sup> dsDNA HS Assay (Invitrogen, Thermo Fisher Scientific), qui permet une mesure plus précise de la quantité d'ADN. Le protocole fourni avec le kit est le suivant : la solution de travail est préparée avec 199 µL de tampon dsDNA HS Buffer et 1 µL de dsDNA HS Reagent contenant le fluorochrome par échantillon. Deux solutions d'ADN standards (une à 0 ng/µL et une à 10 ng/µL) ont été préparées en parallèle de nos échantillons afin d'établir la courbe d'étalonnage : 190 µL de la solution de travail + 10 µL de standard dans un tube de 500 µL (Qubit<sup>®</sup> Assay tube). Puis pour chaque échantillon : 198 µL de la solution de travail et 2 µL d'extrait d'ADN ont été mélangés. Les tubes ont été vortexés 2 à 3 sec puis incubés à température ambiante pendant 5 min. Puis l'analyse a été effectuée avec le Qubit<sup>®</sup>.

Après quantification par spectrophotométrie et fluorométrie, la qualité de l'ADN a été vérifiée sur gel d'agarose à 0,8 % avec du tampon TAE 1X (40 mM Tris-HCl, 20 mM d'acétate de sodium et 1 mM de Na<sub>2</sub>-EDTA, Euromedex). Cinq microlitres d'extrait d'ADN additionnés de 2 µL de tampon de charge 6X (additionné de Gel red, Biotium) ont été soumis à une migration de 100 V pendant 30 min. Un marqueur de taille de 1 kb (Promega) a été ajouté pour vérifier la migration. Le gel a été visualisé puis photographié sur un transilluminateur UV 312 nm avec un appareil photo numérique contrôlé par le logiciel Gel Smart 7.3 (Clara Vision, France).



### II.3.2 Analyse des communautés bactériennes associées aux fèves de cacao par PCR-DGGE

Afin d'analyser la structure des communautés bactériennes présentes à la surface des fèves des échantillons, la méthode de PCR-DGGE a été employée. Elle comprend l'amplification partielle du gène de l'ARNr 16S par PCR puis l'électrophorèse en gel d'acrylamide avec gradient dénaturant.

#### *II.3.2.1 Amplification partielle du gène de l'ARNr 16S par PCR*

Pour l'analyse des communautés bactériennes, la région variable V3 du gène de l'ARNr 16S a été amplifiée en utilisant les amorces bactériennes universelles 338GC F (5' CGC CCG CCG CGC GCG GCG GGC GGGGCG GGG ACT CCT ACG GGA GGC AGCAG 3') et 518 R (5' ATT ACC GCG GCT GCT GG 3') avec une taille d'amplicon d'environ 236 pb (Le Nguyen *et al.* 2008). Il faut noter que l'amorce 338GC F comprend un GC clamp (enchaînement de GC) nécessaire pour la DGGE ultérieure. L'amplification par PCR a été réalisée dans 50 µL de mélange réactionnel (**Tableau 8**) avec le programme suivant : une étape de dénaturation initiale de 3 min à 94°C suivie de 35 cycles de PCR de 30 sec à 94°C pour la dénaturation, 30 sec à 65°C pour l'hybridation des amorces et 30 sec à 72°C pour l'élongation et enfin une étape d'élongation terminale de 10 min à 72°C. Cinq microlitres des produits PCR additionnés de 2 µL de tampon de charge 6X (additionné de Gel red) ont été déposés sur un gel d'agarose à 2% avec du tampon TAE 1X. Le marqueur de taille de 100 pb (Promega) a été ajouté afin d'estimer la taille des amplicons. Après migration à 100 V pendant 30 min, le gel a été visualisé puis photographié de la même manière que précédemment (II.3.1.2).

**Tableau 8:** Composition du mélange réactionnel pour l'amplification partielle du gène de l'ARNr 16S par PCR.

Réactif	Volume	Concentration finale
GoTaq® Reaction Buffer (Promega)	10 µL	1X
dNTP (Promega)	1 µL	200 µM pour chaque dNTP
Amorce 338 GC F (Sigma)*	10 µl	0,2 µM
Amorce 518 R (Sigma)*	10 µL	0.2 µM
GoTaq® G2 DNA Polymerase	0.25 µL	1.25 unités/réaction
Extrait d'ADN	5 µL	10-50 ng
Eau ultra pure qsp	50 µL	

\*Concentration initiale des amorces à 10 µM

### II.3.2.2 Electrophorèse en gel d'acrylamide avec gradient dénaturant (DGGE)

Les produits PCR ont été ensuite analysés dans un gel d'acrylamide/bisacrylamide à 8 % (37,5/1 Biosolve) contenant un gradient de dénaturant chimique : composé d'urée (Eurobio) et de formamide (Thermo scientific). Le gel présente un gradient de 30 à 60 % qui permet la dénaturation de l'ADN double brin (**Tableau 9**). Afin que celle-ci ne soit pas complète un GC-clamp a été ajouté à l'amorce forward lors de l'amplification. Pour la préparation du gradient, deux solutions ont été réalisées séparément, leur composition est donnée dans le **Tableau 9**.

**Tableau 9:** Composition des solutions d'agents dénaturants.

Composants	Fournisseur	30%	60%
Urée	Eurobio	12,6 g	25,2 g
Acrylamide/Bis-acrylamide	Biosolve	20 mL	20 mL
Formamide	PanReac AppliChem	12 mL	24 mL
TAE 50X	Euromedex	2 mL	Qsp 100 mL
Eau		Qsp 100 mL	Qsp 100 mL

Seize microlitres des deux solutions (30 et 60%) ont été mélangés avec 50  $\mu$ L de persulfate d'ammonium (10 %, p/v, Promega). Elles ont été ensuite transvasées dans deux cuves séparées par une vanne du préparateur de gradient (BioRad Model 485, USA). Le préparateur de gradient est relié à une pompe péristaltique (Percom-I, Watson-Marlon, USA) réglée à 30 tours/min. Dans chacune des cuves un agitateur magnétique a été ajouté ainsi que 50  $\mu$ L de Trétraméthyléthylenediamine (TEMED, Promega). Ensuite, à l'aide d'une seringue le gel a été coulé entre deux plaques de verre séparées à l'aide de deux joints situés à leurs extrémités avec une épaisseur de 0,75 mm. Un peigne permettant la formation de puits a été intégré entre les deux plaques. La polymérisation du gel s'est faite en deux heures. Puis, le gel a été inséré dans une cuve remplie de TAE 1X préalablement chauffée à 60°C. Quarante-deux microlitres de produits PCR additionnés de 8  $\mu$ L de tampon de charge (colorant bleu/orange 6X, Promega) ont été déposés dans les puits.

Le gel a été alors placé dans la cuve d'électrophorèse et l'électrophorèse (Biorad Dcode, USA) a été exécutée selon le programme de migration suivant : une première étape de 10 min à 20 V, suivie d'une seconde étape de 12 h à 80 V. Après migration, le gel a été enlevé délicatement des plaques de verre, puis mis en immersion dans une solution de GelRed (Biotium) (250 mL de TAE 1X + 25  $\mu$ L de GelRed) pendant 30 min. Enfin, les profils de chaque échantillon ont été observés sous lumière UV grâce à l'appareil Benchtop UV Transilluminator, BioDo-It™ Imaging System et le logiciel TS software puis enregistrés en format .TIFF en négatif.

### *II.3.2.3 Traitement des images*

L'image des profils DGGE enregistrée au format .TIFF a été analysée à l'aide du logiciel ImageQuant TL version 2007 (Amersham Biosciences, USA). Le logiciel détecte automatiquement les bandes d'ADN des profils (une correction manuelle peut être nécessaire) et génère les distances de migration ( $R_f$ ). Chaque bande correspond à une séquence microbienne donc à une Unité Taxonomique Opérationnelle (OTU) et l'ensemble des bandes représente la communauté bactérienne. L'intensité des bandes correspond à l'abondance de cette séquence, donc de l'OTU détectée. Pour chaque profil, le nombre, la

position et l'intensité des bandes ont été calculés et comparés deux à deux à l'aide d'outils statistiques précisés dans la partie II.3.4.

#### *II.3.2.4 Identification des bandes*

Les différentes bandes obtenues ont été récupérées, à l'aide d'un scalpel stérile. Les bandes ont été éluées dans 100 µL de TE une nuit à 4°C puis purifiées selon le protocole suivant. Cent microlitres de l'ADN élués à partir de chaque bande ont été transférés dans un nouveau micro-tube. Dix microlitres d'acétate de sodium (3 M, pH 5) ont été ajoutés et mélangés par pipetage puis 1 µL de glycogène a été également ajouté et mélangé par pipetage aux 100 µL d'ADN élués. Puis 300 µL d'éthanol absolu (mis à -20°C préalablement) ont été ajoutés. Les tubes ont été mis à -20°C pendant 1 h puis centrifugés durant 30 min à vitesse maximale (18 000 g). Le surnageant a été éliminé et 500 µL d'éthanol à 70 % ajoutés. Les tubes ont été centrifugés à vitesse maximale pendant 5 min. L'éthanol a été éliminé et les tubes ont été mis à sécher sous flux d'air (sorbonne) afin que tout l'éthanol s'évapore. Ensuite 20 µL de TE ont été ajoutés après homogénéisation, l'ADN ainsi purifié a été conservé à -20°C jusqu'à utilisation.

L'ADN contenu dans les bandes a été amplifié à l'aide des amorces 338 F (sans GC-clamp) et 518 R de la même manière que dans le paragraphe II.3.2.1.1 puis l'amplicon d'environ 190 paires de bases a été envoyé pour un séquençage Sanger à la plateforme Genseq de l'ISEM de l'université de Montpellier.

#### **II.3.3 Analyse des communautés bactériennes par metabarcoding (Miseq, Illumina) sur le gène de l'ARNr 16S**

Cette technique a été uniquement utilisée pour l'étude des communautés bactériennes associées aux fèves de cacao en fermentation ainsi qu'aux échantillons environnementaux collectés au Mexique (Octobre 2018).

La technique de metabarcoding présente trois grandes étapes : l'amplification partielle du gène de l'ARNr 16S, la préparation de la banque d'amplicons (purification, quantification et indexation) précédant le séquençage haut débit des amplicons par la technologie Miseq (Illumina). L'amplification partielle du gène de l'ARNr 16S a été réalisée à l'UMR QualiSud et

les étapes de purification, de quantification de l'ADN, d'indexation et de séquençage ont été sous traités à la plateforme Genseq à l'ISEM (Université de Montpellier).

### II.3.3.1 Amplification partielle par PCR du gène de l'ARNr 16S (PCR1)

La région V3/V4 du gène de l'ARNr 16S a été amplifiée à l'aide d'un mix d'amorces 341F1N,2N,3N et 785R 1N,2N,3N (**Tableau 10**). La taille attendue des amplicons est d'environ 600 pb (Bukin *et al.* 2019; Klindworth *et al.* 2013; Thijs *et al.* 2017).

**Tableau 10:** Liste des amorces utilisées pour l'amplification partielle du gène de l'ARNr 16S.

Amorces	Sens/Anti-sens	Séquences (5'-3')
<b>341 F 1N</b>	Sens	TCGTCGGCAGCGTCAGATGTGGTATAAGAGACAG <del>NCCT</del> ACGGGNGGCWGCAG
<b>341 F 2N</b>	Sens	TCGTCGGCAGCGTCAGATGTGGTATAAGAGACAG <del>NNCCT</del> ACGGGNGGCWGCAG
<b>341 F 3N</b>	Sens	TCGTCGGCAGCGTCAGATGTGGTATAAGAGACAG <del>NNNC</del> CTACGGGNGGCWGCAG
<b>785 R 1N</b>	Anti-sens	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG <del>NGAC</del> TACHVGGGGTATCTAATCC
<b>785 R 2N</b>	Anti-sens	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG <del>NNGA</del> CTACHVGGGGTATCTAATCC
<b>785 R 3N</b>	Anti-sens	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG <del>NNN</del> GACTACHVGGGGTATCTAATCC

Les ajouts de N nucléotide dans les séquences des amorces permettent d'apporter de la variabilité. En 5' des amorces, une extension de nucléotides (en gras dans le **Tableau 10**) a été ajoutée afin de réaliser les prochaines étapes (ancrage des adaptateurs Illumina). Pour éviter que les amorces ne s'hybrident avec l'ADN mitochondrial ou chloroplastique potentiellement extrait lors de l'extraction d'ADN sur la pulpe et les fèves de cacao, des bloqueurs de PCR appelés PNA ont été ajoutés au mix PCR (Mpna et Ppna, PNA Bio Inc). Ce sont des séquences nucléotidiques capables de se fixer aux séquences d'ADN mitochondriaux et chloroplastiques. L'utilisation de ces bloqueurs de PCR permet de réduire le biais introduit par ces séquences et donner un résultat de séquençage plus spécifique. L'amplification par PCR (PCR1) a été réalisée dans 25 µL de mélange réactionnel décrit dans le **Tableau 11** et avec le programme suivant : une première étape de dénaturation initiale de 2 min à 95°C suivie de 25 cycles de PCR de 10 sec à 95°C pour la dénaturation, 30 sec à 78°C pour l'hybridation des PNA, 40 sec à 56,5°C pour l'hybridation des amorces et 30 sec à 78°C pour l'élongation et enfin une étape d'élongation terminale de 10 min à 72°C.

**Tableau 11:** Composition du mélange réactionnel pour l'amplification partielle du gène de l'ARNr 16S par PCR.

Réactif	Volume	Concentration finale
<b>Phusion mastermix buffer (Promega)</b>	12.5 µL	1X
<b>Amorces 341 F (Sigma)*</b>	0.625 µl	0,2 µM
<b>Amorces 785 R (Sigma)*</b>	0.625 µL	0.2 µM
<b>Ppna (pna)</b>	1	0.5 µM
<b>Mpna</b>	1	0.5 µM
<b>ADN</b>	6 µL	5 ng
<b>Eau ultra pure qsp</b>	25 µL	qsp

\*Concentration initiale des amorces à 10 µM

Les PCR ont été réalisées en duplicata puis la vérification de l'amplification des produits PCR a été réalisée par électrophorèse en gel d'agarose à 2 % dans du TAE 1X (30 min à 100 V) avec un marqueur de taille 100 pb (Promega). Puis 12,5 µL des produits PCR de chaque duplicata ont été regroupés dans des plaques 96 puits semi-jupées (Eppendorf).

### *II.3.3.2 Préparation de la banque d'amplicons : purification, PCR d'indexation (PCR2) et séquençage*

Les étapes suivantes ont été réalisées par le personnel technique de la plateforme GenSeq à Montpellier. Les produits PCR ont été purifiés avec le Kit CleanPCR (Proteigene) utilisant des billes magnétiques. Dix microlitres des produits d'amplification ont été additionnés avec 8 µL de billes magnétiques pour la purification. Le dosage au Nanodrop des produits d'amplification purifiés a permis de valider la purification et a donné une estimation de la concentration obtenue.

Une PCR d'indexation (PCR2) a ensuite été réalisée afin d'ajouter une séquence unique pour chaque échantillon (index) permettant l'identification de l'origine des séquences produites. Trois cent quatre-vingt-quatre combinaisons d'index de 9 nucléotides sont disponibles (I5A001 à I5D384 et I7A001 à I7D384). Chaque échantillon reçoit une combinaison unique

index I5 (pour la séquence forward) + index I7 (pour la séquence reverse). Les index I5 et I7 sont identiques 2 à 2 mais associés à chacun des deux adaptateurs Illumina du système Nextera (mêmes amorces de séquençage) par exemple la séquence de l'index I5A001 et I7A001 = GCTCACTTG.

Exemple de la construction des séquences avec index, adaptateur et séquence locus spécifique :

#### Index 2 Read

<sup>5'</sup>AATGATACGGCGACCACCGA GATCTACAC[I5XXX] TCGTCGGCAGCGTC –

Pour l'amorce forward : <sup>5'</sup> TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGYR- [séquence locus spécifique]<sup>3'</sup>

#### Index 1 Read

<sup>5'</sup>CAAGCAGAAGACGGCATACGAGAT[I7XXX]GTCTCGTGGGCTCGG –

Pour l'amorce reverse: <sup>5'</sup>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGYR-[séquence locus spécifique]<sup>3'</sup>

**Surligné** : séquence de fixation sur la flow-cell, adaptateur Illumina

**Rouge** : séquence ajoutée lors de la PCR 1 pour l'ancrage de l'index et de l'adaptateur illumina

**Bleu** : index individuel

Les plaques d'amorces indexées s'utilisent par paires :

Plaquette 1 (i5001 à i5096) avec la Plaquette 5 (i7001 à i7096)

Plaquette 2 (i5097 à i5192) avec la Plaquette 6 (i7097 à i7192)

Plaquette 3 (i5193 à i5288) avec la Plaquette 7 (i7193 à i7288)

Plaquette 4 (i5289 à i5384) avec la Plaquette 8 (i7289 à i7384)

La PCR2 a été réalisée dans un volume final de 18 µL contenant 5 µL d'ADN, 2 µL de chaque index et 9 µL du Phusion® High-Fidelity PCR masterMix (Biolabs) avec le programme suivant : une première étape à 95°C pendant 3 min suivie de 8 à 10 cycles à 95°C pendant 30 sec, 55°C pendant 30 sec et 72°C pendant 30 sec et enfin l'étape finale d'élongation à 72°C pendant 5 min (le nombre de cycles dépendant de la quantité d'amplicons obtenu lors de la PCR1). Ensuite les produits d'amplification de la PCR2 ont été regroupés ; 5 µL de chacun des produits d'amplification ont été transférés dans un tube Low-Binding 1,5 mL (~480 µL par plaquette). Soixante-dix µL de chaque pool ont été transférés dans un nouveau tube Low-Binding et purifiés à l'aide de billes magnétiques.

L'analyse sur le Fragment Analyzer (Agilent) avec le kit qualitatif DNF-910 a été réalisée afin de vérifier la taille de la banque ainsi que l'efficacité de la purification (absence de dimère d'amorces), l'augmentation de la taille des fragments entre la PCR1 et la PCR2 (75 pb attendue), le contrôle de l'absence de fragments pour les témoins négatifs. Les pools d'amplicons (ou banque) purifiés ont également été dosés par qPCR.

Le séquençage a été ensuite réalisé sur le MiSeq de la plateforme GenSeq avec un kit de séquençage en paired-end 2X250 nt (15 M de séquences attendues) pour une profondeur maximum de 20 000X.

### II.3.3.3 Analyse des données brutes

Le traitement des données issues du séquençage a été réalisé à l'aide des outils VSEARCH<sup>1</sup>, CUTADAPT<sup>2</sup> et SWARM<sup>3</sup> (Mahé *et al.* 2014, 2015; Rognes *et al.* 2016) . Les séquences nucléotidiques obtenues pour chaque échantillon sont des reads; L'ensemble des reads de tous les échantillons ont été regroupés dans un fichier fastq.

VSEARCH permet de vérifier l'encodage des fichiers fastq et de fusionner les paires de read pour rassembler les séquences reverse et forward pour les séquences similaires. CUTADAPT permet de sélectionner les reads avec les différents index, de sélectionner les reads contenant l'amorce forward et celles contenant l'amorce reverse.

<sup>1</sup><https://github.com/torognes/vsearch>

<sup>2</sup><https://github.com/marcelm/cutadapt/>

<sup>3</sup><https://github.com/torognes/swarm>

Pour chaque échantillon, les informations de la qualité ont été extraites et un fichier Fasta dérépliqué a été créé (les séquences identiques ont été fusionnées et le complémentaire inversé des séquences reverses a été réalisé). L'ensemble de données dérépliquées a été regroupé par SWARM en OTU, une OTU correspond *a priori* à une espèce bactérienne et les représentants des OTU sont transmis à VSEARCH pour détecter les chimères et les séquences de mauvaise qualité. Ensuite, le tableau des OTU a été créé et les séquences des OTU ont été comparées avec la base de données SILVA 123 (<http://www.arb-silva.de>) à l'aide du programme SILVAngs afin d'obtenir l'affiliation taxonomique.



### II.3.4 Analyses statistiques

Les analyses statistiques et les calculs d'indices de la diversité  $\alpha$  et  $\beta$  ont été réalisées à l'aide du logiciel RStudio version 1.1.453 – (© 2009-2018 RStudio, Inc, USA). La diversité microbienne repose sur deux concepts : la richesse spécifique (le nombre de taxons) et l'abondance relative des différents taxons (domination et / ou équitabilité). Cette diversité peut être calculée pour un échantillon (diversité  $\alpha$ ) ou en comparant entre plusieurs échantillons (diversité  $\beta$ ).

#### La diversité $\alpha$

Elle correspond à la diversité au sein d'un échantillon, a été mesurée à l'aide de trois indices :

- Indice de Shannon ( $\alpha$ )

Cet indice permet de mesurer la diversité spécifique des échantillons en tenant compte de la richesse spécifique (nombre de taxons) et de la distribution des individus au sein du taxon (équitabilité spécifique). C'est un indice compris entre 0 et 5, plus cet indice se rapproche de 0 plus la population est représentée par un faible nombre d'individu pour un taxon, et il est maximal lorsque les individus sont répartis de manière homogène au sein du taxon.

- Indice d'Inverse Simpson ( $\alpha$ ) (uniquement pour la partie DGGE)

Cet indice permet de mesurer la probabilité pour que des individus choisis aléatoirement dans la population appartiennent au même taxon. Plus cet indice se rapproche de 0 plus la probabilité d'obtenir des individus du même taxon est élevée.

- Indice d'équitabilité de Pielou ( $\alpha$ ) (uniquement pour la partie NGS)

Cet indice permet de mesurer la répartition des individus au sein des taxons (abondance), indépendamment de la richesse spécifique ; sa valeur varie de 0 (dominance d'un taxon) à 1 (équi-répartition des individus dans les taxons).

La diversité  $\beta$  (diversité entre les échantillons) a été évaluée en calculant la similarité entre les échantillons à l'aide de l'indice de Jaccard qui prend en compte l'absence ou la présence d'un taxon. De la même manière, la dissimilarité entre les échantillons a été calculée à l'aide de l'indice de Bray-Curtis qui tient compte aussi de l'abondance relative des taxons et permet une analyse descriptive des communautés bactériennes des échantillons.

Ces deux outils statistiques permettent de générer une matrice de distance entre les différents échantillons. A partir de cette matrice un dendrogramme est construit par la méthode UPGMA (Unweighted Pair-Wise Grouping with Mathematical Averages) pour établir le regroupement des échantillons. Un positionnement multidimensionnel non-métrique (NMDS) a également été réalisé sur ces échantillons. Une analyse multivariée de la variance par permutations (PERMANOVA / Adonis) a été effectuée afin de déterminer quelles variables qualitatives (jour de fermentation, pays d'origine) ont un impact sur le regroupement des échantillons.

## II.4 Isolement de souches de BAL et BAA impliquées dans la fermentation du cacao

### II.4.1 Protocole d'isolement

Pour l'isolement des BAL et BAA à partir des échantillons de fèves sèches de cacao en provenance de Côte d'Ivoire, le milieu de culture liquide sélectif MRS (Man Rogosa-Sharp) a été employé pour la croissance des BAL et un bouillon nutritif PCA (Plate Count Agar), ainsi que deux milieux gélosés sélectif MRS (BAL) et *Acetobacter* (**Tableau 13**). Puis, nous avons optimisé le protocole d'isolement en utilisant deux milieux de cultures liquides simulant les conditions de la pulpe de cacao durant les étapes de fermentations : le CPSM-LAB (Cocoa Pulp Simulation Media) (Lefeber *et al.* 2010) et le CPSM-AAB (Moens *et al.* 2014) (**Tableau 12**). Le CPSM-LAB correspond à la composition de la pulpe au premier jour de la fermentation et le CPSM-AAB à sa composition à partir du second jour de fermentation, ce qui permet d'isoler les BAL et les BAA respectivement.

**Tableau 12:** Composition des milieux de culture liquides utilisés pour la croissance des BAL et des BAA.

<b>CPSM-LAB</b>	<b>CPSM-AAB</b>
<b><u>Composition pour 1 L :</u></b>	<b><u>Composition pour 1 L :</u></b>
<b>pH à 5,5</b>	<b>pH à 4,5</b>
<b>Fructose (Fluka) : 25 g</b>	D-mannitol (Sigma-Aldrich) : 10 g
<b>Glucose (Sigma) : 25 g</b>	Acide lactique (Sigma) : 10 g
<b>Acide citrique : 10 g</b>	Extrait de levure : 5 g
<b>Extrait de levure (Difco BD) : 5 g</b>	Peptone de soja (Merck KGaA) : 5 g
<b>Peptone de soja (Merck KGaA) : 5 g</b>	Tween 80 (Panreac) : 1 mL
<b>MgSO<sub>4</sub>·5H<sub>2</sub>O (Fluka) : 0,5 g</b>	Eau distillée : qsp 1 L
<b>MnO<sub>4</sub>S - H<sub>2</sub>O (Fluka) : 0,2 g</b>	Ethanol absolu : 20 mL*
<b>Tween 80 (Panreac) : 1 mL</b>	
<b>Eau distillée : qsp 1 L</b>	

Autoclavage à 121 °C pendant 15 min

\*à rajouter après autoclavage

Dix grammes de chaque échantillon de fèves de cacao ont été mis pendant 48 h dans 90 mL de bouillon CPSM-LAB (ou MRS pour les fèves sèches) et dans 90 mL de bouillon CPSM-AAB (nutritif, pour les fèves sèches), dans un bain marie sous agitation à 1400 tours/min à 30°C. Ensuite, des dilutions en cascade au dixième ont été réalisées dans de l'eau physiologique stérile (NaCl 9 g/L). Le nombre de dilutions dépend de la turbidité de la culture obtenue. Cent microlitres des deux dernières dilutions de chacune des cultures ont été ensemencés sur milieu gélosé sélectif des BAL, MRS (aérobiose et anaérobiose, 48h à 30°C) et sur milieu gélosé Kneifel, ou milieu gélosé *Acetobacter* sélectif des BAA (aérobiose, 48h à 30°C).

Le milieu MRS (Biokar) comprend notamment un antifongique, le cycloheximide (Sigma-Aldrich) (100 mg/L) afin d'inhiber la croissance de la flore fongique ainsi que du carbonate de calcium CaCO<sub>3</sub> (10 g/L). L'acidification du milieu MRS + CaCO<sub>3</sub> et *Acetobacter* provoque la dégradation du CaCO<sub>3</sub> et forme donc un halo de dégradation autour des colonies ce qui traduit

la production d'acide. Les bactéries lactiques sur le milieu MRS et les bactéries acétiques sur le milieu *Acetobacter* présentent des colonies entourées d'un halo de dégradation du  $\text{CaCO}_3$ .

Le milieu Kneifel est composé d'un indicateur de pH le vert de Bromocrésol (Sigma-Aldrich), qui lorsque le milieu est acidifié, passe de la couleur bleue à vert/jaune. Un halo vert ou jaune autour de la colonie traduit la production d'acide. Il contient aussi un antifongique pour inhiber la croissance de la flore fongique (cycloheximide, Sigma-Aldrich) et un antibiotique (Pénicilline-G, Sigma-Aldrich) pour inhiber la croissance des bactéries lactiques (**Tableau 13**). Les bactéries acétiques sur ce milieu présentent des colonies entourées d'un halo jaune / vert dû à la consommation d'éthanol transformé en acide acétique. Les colonies présentant les aspects macroscopiques correspondants aux BAL et BAA sont repiquées sur ces mêmes milieux (sans ajout d'indicateur de pH ou de  $\text{CaCO}_3$ ) pour l'obtention de cultures pures.

**Tableau 13:** Composition du milieu de culture gélosé pour l'isolement des BAA

<u>Kneifel</u>	<u>Acetobacter</u>
<b>Composition pour 1 L :</b>	<b>Composition pour 1 L :</b>
<b>Agar/ gélose : 20 g (Biokar)</b>	Agar/ gélose : 12 g (Biokar)
<b>Extrait de levure : 30 g (Biokar)</b>	Extrait de levure : 8 g (Biokar)
<b>Solution de vert de Bromocrésol dans de l'éthanol (2.2 %) : 1 mL</b>	Carbonate de calcium $\text{CaCO}_3$ : 7.6 g
<b>Qsp : 1000 mL</b>	BD Peptone Bacto™: 5 g
<b>Ethanol : 20 mL*</b>	Glucose (Sigma-Aldrich): 20 g
<b>Cycloheximide : 100 mg*</b>	Ethanol: 10 mL*
<b>Pénicilline-G : 12,5 mg*</b>	

Autoclavage à 121 °C pendant 15 min

\*à rajouter après autoclavage

#### II.4.2 Caractérisations phénotypiques des isolats

Une coloration de Gram a été effectuée sur chacune des colonies présentant l'aspect macroscopique attendu afin de déterminer leur morphologie. Le protocole utilisé pour la coloration de Gram est le suivant : un frottis cellulaire est réalisé puis fixé à la chaleur. Celui-ci est coloré au violet de gentiane pendant 1 min, après un rinçage à l'eau distillée quelques gouttes de Lugol sont ajoutées. Un rinçage à l'acétone/alcool est réalisé afin de décolorer les

cellules, après le rinçage à l'eau les cellules sont colorées avec une solution de safranine pendant 1 min puis rincées à l'eau. Les lames sont observées au microscope optique au grossissement 1000 après ajout d'une goutte d'huile à immersion.

A l'issue de la coloration de Gram, un test de recherche de la catalase a été réalisé sur les bacilles à Gram positif pour l'identification des souches potentielles de BAL. Les souches correspondant aux bacilles à Gram positif et catalase négative ont été sélectionnées et mises en cultures dans le milieu CPSM-LAB (30°C, 48 h). Les bacilles, ovoïdes à Gram négatif ont été ensuite sélectionnés comme souches potentielles de BAA et mises en culture dans le milieu CPSM-AAB (30°C, 48 h) pour conservation. Cinq cents microlitres de glycérol stérile ont été ajoutés à 1,5 mL de ces bouillons de culture dans un cryotube. Les souches ont ainsi été conservées à -80°C.

#### II.4.3 Identification moléculaire des isolats au niveau de l'espèce

Les extractions d'ADN ont été réalisées sur des cultures pures obtenues après 48 h d'incubation à 30°C.

##### *II.4.3.1 Extraction de l'ADN par choc thermique*

Une colonie a été déposée dans un micro-tube contenant 50 µL d'eau moléculaire stérile. Ce tube a été porté à ébullition pendant 10 min puis immédiatement mis pendant 10 min à -20°C.

##### *II. 4.3.2 Pré-identification des BAL par PCR*

Une pré-identification des BAL a été réalisée à l'aide d'amorces : LAC1 (5'- AGC AGT AGG GAA TCT TCCA – 3') et LAC2 (5'- ATT YCA CCG CTA CAC ATG – 3') spécifiques aux genres *Lactobacillus*, *Pediococcus*, *Weissella*, et *Leuconostoc*. Ces amorces permettent d'amplifier une séquence de 340 pb au sein du gène de l'ARNr 16S (Walter *et al.* 2001). L'amplification a été réalisée dans un volume réactionnel de 50 µL (**Tableau 14**) selon le programme suivant : une dénaturation initiale à 94°C pendant 2 min, puis 35 cycles incluant une dénaturation à 94°C pendant 30 sec, l'hybridation des amorces à 55 °C pendant 30 sec et une élongation à 72°C pendant 1 min et enfin une élongation finale de 5 min à 72°C.

**Tableau 14** : Composition du volume réactionnel de la PCR pour la pré-identification des BAL.

Réactif	Volume*	Concentration finale
GoTaq® Reaction Buffer (Promega)	10 µL	1X
dNTP (Promega)	1 µL	200 µM pour chaque dNTP
Amorce LAC1 (Sigma)*	10 µl	0,2 µM
Amorce LAC2 (Sigma)*	10 µL	0.2 µM
GoTaq® G2 DNA Polymerase (Promega)	0.25 µL	1.25 unités/réaction
Extrait d'ADN par choc thermique	5 µL	
Eau ultra pure qsp	50 µL	

\*Concentration initiale des amorces à 10 µM

Après amplification, 5 µL d'ADN ont été mélangés avec 2 µL de colorant bleu/orange 6X (Promega) comprenant l'intercalant d'ADN GelRed (Biotium) puis analysés par électrophorèse sur gel d'agarose à 2 % avec du tampon TAE 1X afin de vérifier l'amplification. La taille des fragments d'ADN a été estimée à l'aide du marqueur de poids moléculaire 100 pb (Promega).

Après migration à 100 V pendant 30 min, le gel a été visualisé puis photographié de la même manière que précédemment (II.3.1.2).

#### II. 4.3.3 Identification des BAL et BAA par amplification et séquençage du gène de l'ARNr 16S

La quasi-totalité du gène de l'ARNr 16S (1500 pb) a été amplifié par PCR à l'aide des amorces 27 F (5'- GTGCTGCAGAGAGTTTGATCCTGGCTCAG- 3') et 1492 R (5'- CACGGATCCTACGGGTACCTGTTACGACTT - 3') (Marchesi *et al.* 1998) avec un mélange réactionnel de 50 µL (**Tableau 14** modifié pour les amorces). Une première étape de dénaturation initiale de 3 min à 94°C a été réalisée suivie de 35 cycles incluant une dénaturation de 30 sec à 94°C, 30 sec à 55°C pour l'hybridation des amorces et 30 sec à 72°C pour l'élongation et enfin une étape d'élongation terminale de 5 min à 72°C.

Ensuite, après vérification de l'amplification par électrophorèse sur gel d'agarose à 1 % avec du tampon TAE 1X pendant 30 min à 100 V, les produits PCR ont été envoyés pour séquençage Sanger (Genseq, Montpellier). Les séquences d'ADN sens et anti-sens ont été analysées et corrigées à l'aide du logiciel Sequence Scanner v1.0 Software (Applied Biosystems) puis le

logiciel Geneious Prime<sup>®</sup> 2019.0.4 Build 2018, restricted version (Invitrogen Corporation) nous a permis d'obtenir la séquence complète. Les séquences ont été alors comparées avec la base de données Genbank à l'aide du programme Blastn (Altschul, 1997) sur le site web du NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

## II.5 Analyse de la diversité intra spécifique des BAA isolées lors des fermentations de cacao

### II.5.1 Analyse du polymorphisme génomique : PCR sur séquences répétées

Afin d'analyser le polymorphisme génomique, la méthode basée sur l'amplification de régions situées entre des séquences répétées (rep-PCR) a été utilisée. Ces séquences répétées sont présentes en de nombreuses copies et hautement conservées dans les génomes, ce qui permet d'obtenir des profils génomiques pour chaque souche (Versalovic *et al.* 1991). Il existe différentes séquences répétées retrouvées dans le génome bactérien : les séquences REP (Repetitive Extragenic Palindromic element), les séquences ERIC (Enterobacterial repeated Intergenic Consensus) les BOX éléments, et la séquence (GTG)<sub>5</sub> (**Tableau 15**).

**Tableau 15** : Listes des amorces utilisées pour l'analyse par PCR sur séquences répétées (Versalovic *et al.* 1994).

Amorces	Séquences (5'-3')	Température d'hybridation
<b>REP1R-I</b>	CGICGICATCIGGC	40 °C
<b>REP2-I</b>	CGNCGNCATCNGGC	
<b>ERIC1R</b>	ATGTAGATCCTGGGGATTAC	52 °C
<b>ERIC2</b>	AAGTAAGTGACTGGGGTGAGCG	
<b>Box1AR</b>	CTACGGCAAGGCGACGCTGACG	52 °C
<b>(GTG)<sub>5</sub></b>	GTGGTGGTGGTGGTG	40°C

#### II.5.1.1 PCR sur séquences répétées : ERIC-PCR, REP-PCR, Box-PCR, (GTG)<sub>5</sub>-PCR

Dans un volume réactionnel de 50 µL (**Tableau 16**), des amplifications par PCR à l'aide d'amorces qui ciblent un type de séquences répétées ont été réalisées selon le programme suivant : une première étape de dénaturation initiale de 5 min à 95°C suivie de 30 cycles de 30 sec à 90°C pour la dénaturation, 30 sec à 40°C ou 52°C (selon le type d'amorces, **Tableau**

**15)** pour l'hybridation des amorces, 4 min à 65°C pour l'élongation et enfin une étape d'élongation terminale de 8 min à 65°C (Gonzalez *et al.* 2004). Après migration à 100 V pendant 1h30, le gel a été immergé 1 h dans une solution de Gelred dilué dans du TAE 1X. Les profils ont été visualisés et photographiés de la même manière que précédemment (II.3.1.2) puis enregistrés en format .TIFF en négatif.

**Tableau 16 :** Composition du mélange réactionnel pour l'analyse par PCR sur séquences répétées

Réactif	Volume	Concentration finale
Master Mix (Promega)	25 µL	1X
Amorces rep*	10 µl	2 µM
Extrait d'ADN par choc thermique	1 µL	10 ng
DMSO	5 µL	10 % (v/v)
Eau ultra pure qsp	50 µL	

\* Concentration initiale des amorces 100 µM

#### II.5.1.2 PCR Multiplexe sur séquences répétées (Box1AR, ERIC, REP)

Un test comparatif a été réalisé en utilisant la méthode de multiplexe, avec 3 jeux d'amorces de séquences répétées différentes : Box1AR, ERIC et REP (**Tableau 15**). Dans un volume réactionnel de 50 µL (**Tableau 16**) avec une concentration en amorces de 0,2 µM et de DMSO à 3,5 % (v/v). Les séquences ont été amplifiées par PCR selon le programme suivant : une première étape de dénaturation initiale de 3 min à 95°C suivie de 30 cycles de 30 sec à 90°C pour la dénaturation, 1 min à 40°C pour l'hybridation des amorces et 4 min à 72°C pour l'élongation et enfin une étape d'élongation terminale de 16 min à 72°C (Dupont *et al.* 2015). Ensuite, les produits PCR ont été mis à migrer sur gel d'agarose à 1,5 % dans du TAE 1X pendant 1h30 puis dans un bain de gel red + TAE 1X pendant 1 h. Les profils ont été observés comme précédemment (II.3.1.2).

#### II.5.1.3 Analyse des profils génomiques

Les photos des gels ont été analysées à l'aide du logiciel CLIQS version 1.2.0.044 (Totalab, UK). Le logiciel détecte automatiquement les bandes présentes pour chaque échantillon, une



correction manuelle peut être nécessaire et génère les fronts de migration (Rf). Chaque bande a sa valeur de Rf, position donnée sur le gel. Le logiciel permet de réaliser une normalisation des Rf entre différents gels d'électrophorèses pour pouvoir les comparer entre eux.

Les PCR ont été réalisées en dupliquât pour chaque souche, afin de s'assurer de la reproductibilité de la méthode. Seules les bandes qui étaient partagées entre les deux amplifications ont été conservées pour obtenir le profil numérique de chaque souche et ainsi réaliser l'analyse statistique.

L'absence et la présence de bandes à différentes positions du gel (Rf) a été notée 1 (présence) et 0 (absence) pour chaque souche afin de réaliser une matrice en deux dimensions en code binaire. Ensuite, à l'aide du logiciel Rstudio, une matrice de distance a été réalisée en calculant l'indice de Sørensen-Dice, qui est un indice qui mesure la similarité entre deux échantillons en tenant compte de la présence ou de l'absence des bandes. A partir de cette matrice un dendrogramme a été construit par la méthode UPGMA (Unweighted Pair-Wise Grouping with Mathematical Averages) pour établir le regroupement des souches. Une analyse multivariée de la variance par permutations (PERMANOVA/ Adonis) a été effectuée afin de déterminer quelles variables qualitatives (jour de fermentation, pays d'origine) ont un impact sur le regroupement des souches.

D'autres tests statistiques ont été utilisés, Shapiro-Wilk (test de normalité) et ANOVA et test de Student (comparaisons de moyennes).

## II.5.2 Caractérisation biochimique des souches

### *II.5.2.1 Suivi de cinétique de croissance à l'aide d'un spectrofluorimètre*

#### 1) Préculture en milieu liquide

Les souches conservées à -80°C ont été mises en culture sur gélose Kneifel pendant 2 jours à 30°C. Puis une à deux colonies ont été mises en suspension dans 5 mL de milieu liquide CPSM-AAB durant une nuit à 30°C sous agitation à 140 rpm. La densité optique (DO) des cultures a été mesurée à l'aide du spectrophotomètre UV-1800 (Shimadzu) à la longueur d'onde  $\lambda$  à 600 nm (DO600). Puis les cultures ont été transférées en triplicata sur plaque 96 puits à fond plat (Nunclon™Surface) à une DO de 0,01.

## 2) Protocole

La cinétique de croissance bactérienne a été réalisée à l'aide du spectrofluorimètre Spark 10M (TECAN®) en mesurant l'absorbance des cellules ( $\lambda$  à 600 nm). Le logiciel Spark Control v2.1-2016 (GmbH – TECAN Austria) nous a permis d'établir un protocole adapté à notre étude. Ainsi, la plaque a été agitée à 150 rpm pendant 600 sec puis une première mesure d'absorbance a été réalisée. Ensuite, la plaque a été agitée toutes les 10 min et 30 sec et une nouvelle mesure d'absorbance a été réalisée. Cent-vingt cycles ont été réalisés pour une durée totale de 20 h.

## 3) Calcul du taux de croissance spécifique

Le taux de croissance spécifique a été calculé en extrayant la pente de  $\ln(\text{DO600}) = f(\text{temps})$  et en prenant la valeur maximale.

### *II.5.2.2 Analyse des substrats consommés et du produit formé*

Les substrats consommés lors de cette étude par les BAA sont le mannitol, l'éthanol et l'acide lactique et le produit formé est l'acide acétique.

L'analyse des métabolites a été réalisée en chromatographie en phase liquide à haute performance (HPLC). L'appareil HPLC utilisé au cours de cette étude est le Ultimate 3000, Fluorescence detector (ThermoScientific), le logiciel associé à l'appareil est Chromeleon 7.2.7 Thermo Scientific Inc® 2009-2017. La détection se fait à l'aide d'un indice de réfraction et des ultraviolets (UV). La séparation des composés est réalisée sur une colonne Aminex HPX-87H (Biorad, Hemel Hempstead, UK) de dimension 300 \* 7,8 mm; L'élution des échantillons a été effectuée à 30°C avec de l'acide sulfurique (5 mM) à un débit de 0,6 mL/min. La détection est effectuée en utilisant le détecteur UV (210 nm) pour les acides et le réfracteur pour les alcools et sucres. Une courbe d'étalonnage à 4 points de concentration (2,5 g/L, 5 g/L, 7,5 g/L et 10 g/L) définissant la gamme d'étalonnage de la méthode a été réalisée. Pour cela, une solution mère avec un mélange d'acide lactique, de mannitol, d'éthanol et d'acide acétique a été traitée par l'HPLC avant toute analyse.

Différents paramètres ont été évalués pour l'analyse comparative des souches. Le rendement en biomasse a été calculé comme suit :  $Y_{X/S}$  où X est la biomasse (DO600) et S la consommation en substrats (éthanol et acide lactique) en mM (valeur par HPLC). Le rendement de production

de l'éthanol en acide acétique ( $Y_{AA/ETH}$ ) a également été calculé, c'est le rendement expérimental sur le rendement théorique. Une conversion de 1 mole d'acétique par mole d'éthanol a été envisagée, le rendement théorique devrait être de 1. Le rendement en biomasse produit par mole d'acide acétique produite a été calculé comme suit :  $Y_{X/AA}$ , où X est la biomasse (OD) et AA la production d'acide acétique en mM. Enfin, le rapport acide lactique sur éthanol ( $R_{LA/ETH}$ ) a également été calculé afin de déterminer l'affinité relative pour l'éthanol et pour l'acide lactique comme source de carbone pour leur croissance et/ou pour produire de l'acide acétique.

### *II.5.2.3 Analyse statistique*

Toutes les analyses statistiques ont été réalisées avec les logiciels XLSTAT version 2019.1.3 (Addinsoft, France) et Rstudio version 1.1.453 – (© 2009-2018 RStudio, Inc, USA). Le prétraitement des données HPLC a consisté en une normalisation sur le produit de la biomasse (DO) pour corriger les différences entre les souches dans la consommation et dans les expériences de fermentation à l'échelle du laboratoire. Ensuite, une conversion en Z-scores pour corriger le bruit, a été calculée comme suit :  $Z\text{-score} = (X - \mu) / \sigma$ , où X est la mesure de la valeur transformée logarithmique,  $\mu$  est la valeur moyenne et  $\sigma$  est l'écart type des valeurs de toutes les souches. L'analyse de regroupement a été effectuée en utilisant la distance euclidienne pour construire la matrice de distance entre les souches et l'algorithme UPGMA (Unweighted pair group method with arithmetic mean) a été utilisé pour construire le dendrogramme. Les données ont été analysées statistiquement à l'aide d'une analyse de variance à sens unique (Kruskal-Wallis) afin de déterminer les différences significatives ( $p < 0,05$ ). La méthode de comparaison multiple de Dunn a été utilisée pour examiner plus en détail toute différence significative entre les souches après avoir vérifié la normalité des données (Test de Shapiro-Wilk,  $p\text{-value} < 0,05$ ).

# Chapitre III : Résultats et discussion

### III.1 Influence de l'origine géographique et de la période de récolte sur la structure des communautés bactériennes impliquées dans la fermentation de fèves de cacao

Lors de la fermentation des fèves de cacao en plus des levures une succession de groupes bactériens évoluent au cours du temps. Ces communautés successives sont les Entérobactéries, les BAL, les BAA et les *Bacillus* formant le consortium bactérien. Ce chapitre présente la composition de ce consortium incluant les espèces dominant les débuts et fins de fermentation. La composition des communautés bactériennes associées aux fèves de cacao en fermentation a été étudiée dans 3 pays différents : le Mexique, la Guyane et la Côte d'Ivoire à différents jours de fermentation. Cette étude consiste à répondre aux questions de recherches suivantes :

- Quel est l'impact de l'origine géographique sur les communautés bactériennes ?
- Existe-il des taxons spécifiques à un pays ?

De plus, lors de ce chapitre nous nous sommes également intéressés à l'effet de la grande récolte et de la petite récolte sur les communautés bactériennes. Sachant que le nombre de fèves par cabosse est plus faible lors de la petite récolte et que les fèves sont moins chargées en pulpe, afin de répondre à la question de recherche suivante :

- Quel est l'impact de la période de récolte des cabosses sur les communautés bactériennes associées à la fermentation des fèves ?

Pour répondre à ces questions, deux méthodes ont été employées : la PCR-DGGE et le séquençage haut débit par la méthode de metabarcoding à l'aide du Miseq (Illumina) après extractions par phénol/chloroforme des ADN totaux des bactéries présentes à la surface des fèves et de la pulpe de chaque échantillon. Leur diversité spécifique ( $\alpha$ ), la diversité entre les échantillons ( $\beta$ ) et leur composition en phyla et genres bactériens ont été évalués pour décrire les communautés bactériennes. Les différents pays échantillonnés pour cette étude, ont chacun des pratiques de récolte et de mise en fermentation différente. Chacune de ces différences peuvent influencer les communautés bactériennes impliquées dans les fermentations : comme le type de fermentation

en caisse ou en tas, le matériau de la caisse en bois ou en polystyrène, les conditions météorologiques (plus d'humidité) et les manipulateurs eux-mêmes.

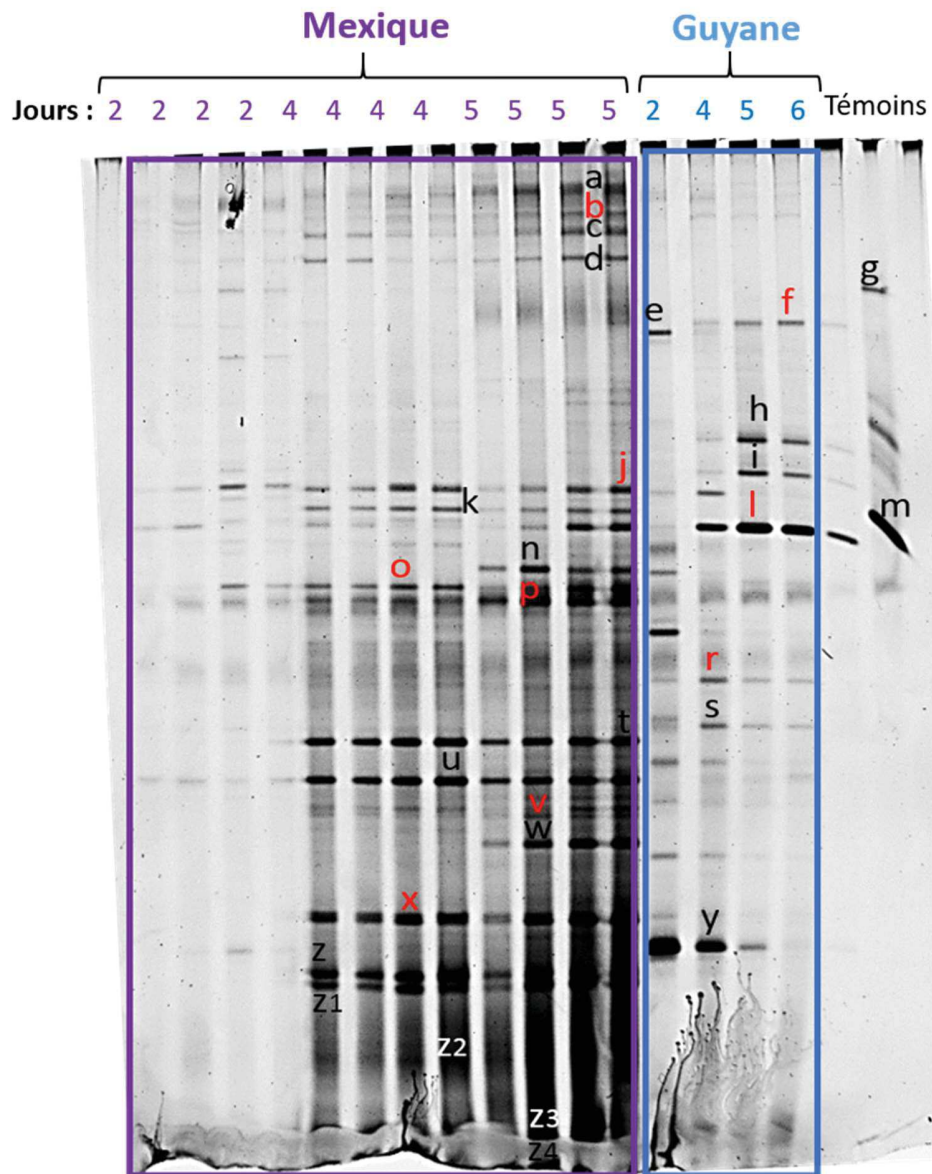
### III.1.1 Etude comparative des communautés bactériennes des fèves de cacao en fermentation en fonction de leurs origines géographiques

#### *III.1.1.1 Comparaison des empreintes globales bactériennes en fonction de 2 pays*

Les communautés bactériennes présentes à la surface des fèves de cacao fraîches et de la pulpe des échantillons provenant de Guyane et du Mexique (Février 2018) ont été analysées par la méthode moléculaire d'analyse globale, la PCR-DGGE. Au moment de l'analyse nous ne disposions pas des échantillons de fèves fraîches en provenance de Côte d'Ivoire.

Deux caisses de fermentations en bois ont été échantillonnées lors de la campagne réalisée au Mexique (février 2018). On a obtenu des extraits d'ADN dont les concentrations varient de 5,95 à 48,5 ng/ $\mu$ L. La région variable V3 du gène de l'ARNr 16S a été amplifiée par PCR avec des amorces universelles, puis les amplicons obtenus ont été analysés par DGGE suivant le protocole énoncé dans le chapitre, Matériels et méthodes (II.3.2.2).

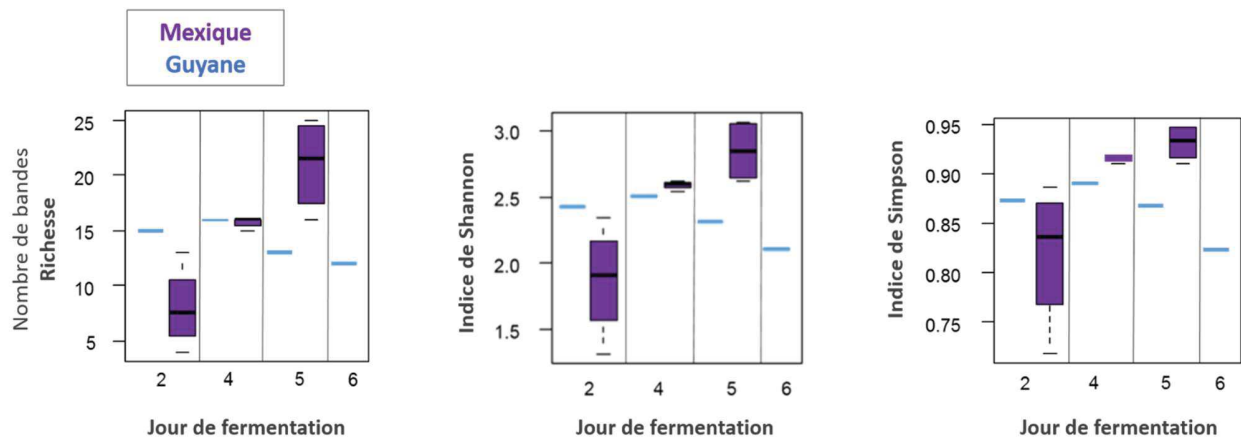
L'observation des profils DGGE, qui correspondent à la communauté bactérienne présente à la surface des fèves et pulpe de cacao, a montré une variation au cours du temps de fermentation en termes d'abondance (intensité des bandes) et de richesse (nombre de bandes) quel que soit le pays d'origine et ont également montré des variations entre les profils des deux pays d'origine (**Figure 27**).



**Figure 27** : Profils DGGE des ADNr 16S présents à la surface des fèves de cacao en provenance du **Mexique** et de **Guyane** échantillonnées à différents temps de fermentation (2, 4, 5 et 6 jours (uniquement Guyane))

Visuellement, les profils des échantillons semblent plus diversifiés au Mexique qu'en Guyane, par leur nombre plus important de bandes présentes. Seules 10 bandes sur 30 bandes sont communes entre les différents profils (b, f, j, l, o, p, r, u, v et x) des deux pays.

### III.1.1.1.1 Comparaison de la diversité $\alpha$ des échantillons de fèves de cacao en fermentation en provenance de 2 pays



**Figure 28** : Indices de diversité  $\alpha$  exprimés sous forme de boîtes à moustaches : (A) la richesse (nombre de bandes), (B) Shannon et (C) Inverse Simpson sur les profils obtenus après analyse par PCR-DGGE des communautés bactériennes associées aux fèves de cacao fermentées en provenance de [Guyane](#) et du [Mexique](#).

Nous avons pu constater que la communauté bactérienne au 2<sup>ème</sup> jour de fermentation des fèves issues de Guyane est plus riche comparée à celles issues du Mexique. En effet, le nombre de bandes d'ADN (richesse) retrouvées dans le profil de l'échantillon de Guyane est plus important (15 bandes) contre 8 bandes (valeur médiane) pour les profils du Mexique. Cependant, pour les échantillons au 4<sup>ème</sup> jour de fermentation, le nombre de bandes présent dans les communautés bactériennes pour les deux pays est identique avec 16 bandes (valeur médiane). Pour le 5<sup>ème</sup> jour de fermentation, les communautés bactériennes des échantillons du Mexique sont plus diverses que celle de Guyane : avec 21 bandes pour le Mexique et 13 bandes pour la Guyane (**Figure 28A**). Le nombre de bandes est identique pour les échantillons au 5<sup>ème</sup> et 6<sup>ème</sup> jour de fermentation en provenance de la Guyane. Pour la plupart des bandes, l'intensité augmente au cours du temps, excepté pour la bande  $\gamma$ , présente uniquement en Guyane ; son intensité diminue au cours du temps. Le nombre de bandes obtenu pour les échantillons au Mexique augmente au cours du temps de fermentation. Une différence notable est observée entre les



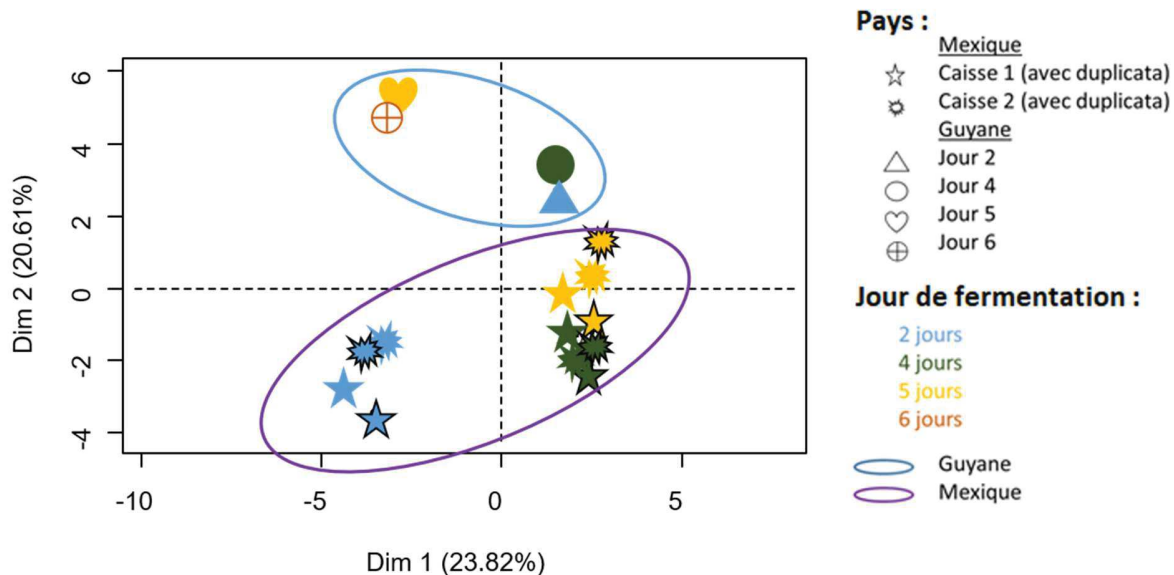
deux pays. En effet, on observe une augmentation du nombre de bandes aux premiers jours de fermentation pour la Guyane, puis à partir du 4<sup>ème</sup> jour de fermentation on observe une diminution de celui-ci contrairement aux échantillons du Mexique.

Les indices de diversité Shannon et Inverse Simpson (**Figure 28A et 28B**), ont été utilisés pour décrire la diversité  $\alpha$ , c'est-à-dire la diversité au sein de chaque échantillon. Les résultats ont confirmé ceux observés en analysant la richesse des échantillons : il y a un effet du jour de la fermentation pour le Mexique, avec une diversité plus riche qui augmente au cours du temps de fermentation. Également, il y a un effet du pays d'origine, avec une diversité plus importante en Guyane (Shannon = 2,4 et Inv. Simpson = 0,80) qu'au Mexique (Shannon = 1,8 et Inv. Simpson = 0,83 – valeurs médianes) au 2<sup>ème</sup> jour de fermentation, et une diversité plus importante au Mexique qu'en Guyane aux autres jours de fermentation. Tous les échantillons ont une diversité riche avec des indices de Shannon supérieurs à 1,8 et des indices d'Inverse Simpson proche de 1, supérieurs à 0,83. Cette différence de diversité peut être expliquée par la sélection d'un genre bactérien en Guyane, lié aux pratiques de mise en fermentation qui diffèrent entre les deux pays. En effet, en Guyane, les fermentations sont réalisées dans des caisses en polystyrène qui sont recyclées de fermentation en fermentation ce qui permet le réensemencement des fèves. Cette pratique a sûrement sélectionné un genre bactérien, qui aurait pris le dessus sur les autres genres impliqués lors des fermentations de fèves de cacao.

#### *III.1.1.1.2 Analyse de la diversité $\beta$ des échantillons de fèves de cacao en fermentation en provenance de 2 pays*

Les profils DGGE ont été comparés par analyses statistiques, tout d'abord une analyse en composantes principales (ACP) (**Figure 29**) nous a permis d'obtenir une représentation graphique de l'ensemble des échantillons. L'ACP, a permis d'illustrer un regroupement des échantillons selon le pays mais également selon le jour de fermentation, la somme des pourcentages des deux premières composantes a expliqué à 44,4 % les variations observées dans l'échantillonnage. La première composante a permis de séparer les échantillons en fonction du pays d'origine et la

seconde en fonction du jour de fermentation, les pourcentages obtenus pour ces regroupements ont été de 20,61 % et 23,83 % respectivement.



**Figure 29 :** Analyse en composantes principales (ACP) des profils des communautés bactériennes obtenus après analyse par PCR-DGGE des communautés bactériennes associées aux fèves de cacao fermentées en provenance de [Guyane](#) et du [Mexique](#).

Ensuite, l'indice de similarité de Bray-Curtis a été calculé, il prend en compte la présence ou l'absence des bandes ainsi que de leur abondance, une matrice de distance est ainsi générée.

A partir de cette matrice un positionnement multidimensionnel non-métrique (NDMS) (**Figure 30A**) et un dendrogramme (**Figure 30B**) ont été réalisés pour le regroupement des échantillons.



**Figure 30 :** (A) Positionnement multidimensionnel non-métrique (NMDS) et (B) dendrogramme pour le regroupement des profils obtenus après analyse par PCR-DGGE des communautés bactériennes associées aux fèves de cacao fermentées en provenance de **Guyane** et du **Mexique**.

L'indice de Bray-Curtis a été utilisé pour analyser la similarité, un positionnement multidimensionnel non-métrique (NMDS) et un algorithme de regroupement UGPMA a été appliqué pour regrouper les échantillons.

Ces graphiques ont permis de confirmer le regroupement des échantillons selon leur pays d'origine et selon le jour de fermentation.

Le dendrogramme (**Figure 30B**) nous a permis d'évaluer le degré de similarité entre les profils. Le dendrogramme est divisé en 3 groupes, le premier est composé des 4 échantillons au 2<sup>ème</sup> jour de fermentation du Mexique, le second, des échantillons aux 4<sup>ème</sup> et 5<sup>ème</sup> jour de fermentation du Mexique et le troisième groupe de l'ensemble des échantillons en provenance de Guyane. Le premier groupe présente moins de 40 % de similarité avec les deux autres groupes, leurs profils ont démontré une faible diversité spécifique, avec un nombre de bandes faibles (7 bandes valeurs médiane **Figure 28A**) et de plus faibles indices de diversité Shannon et Inverse Simpson que les autres échantillons (1,8 et 0,83 respectivement, **Figure 28B et 28C**). Les deux autres groupes sont regroupés avec 50 % de similarité. On peut également constater que les deux répliques biologiques (2 caisses de fermentations différentes) du Mexique ont été regroupés ensemble avec une forte similarité, de 70 à 95 %.

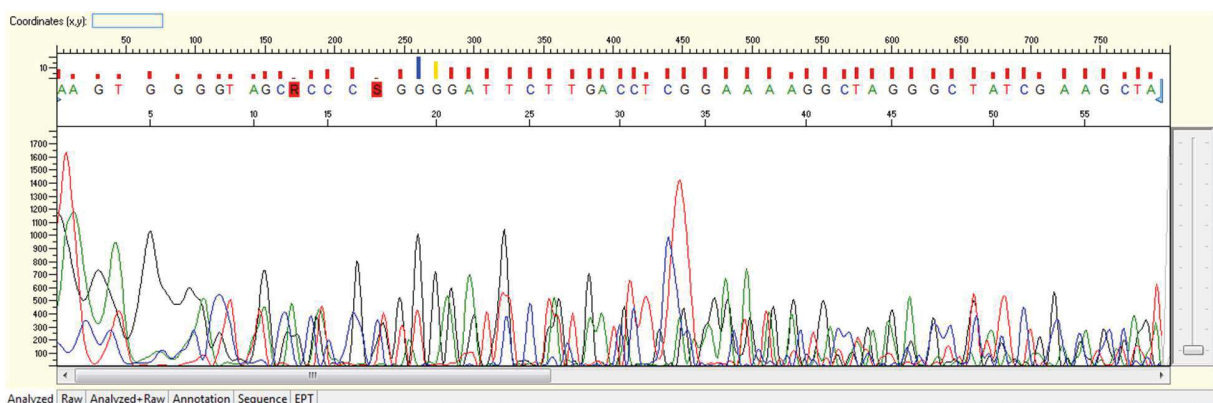
Le test statistique réalisé en analysant les matrices de distance créées avec l'indice Bray-Curtis a montré que le pays d'origine et le jour de fermentation ont eu un effet important sur le regroupement des échantillons avec une p-value=0,003\*\*\* et 0,0013\*\*\* respectivement.

### III.1.1.1.3 Identification des OTUs bactériennes

L'identification des bandes ADN d'intérêt récupérées à partir des empreintes bactériennes (DGGE) a été réalisée par séquençage Sanger afin d'identifier les taxons majoritaires et ainsi avoir un aperçu de la composition des communautés bactériennes présentes à la surface des fèves de cacao.

Les 29 bandes majoritaires au sein des profils DGGE ont été excisées, purifiées, amplifiées puis séquencées par la méthode Sanger. Les séquences obtenues ont été comparées dans les bases de données et présentent un pourcentage d'identité entre 92 et 99 % avec le taxon le plus proche. La plupart des identifications n'ont pas pu être obtenues au niveau de l'espèce, mais pour la majorité l'identification obtenue était au niveau du genre ou de la famille. L'impossibilité d'obtenir l'identification au niveau de l'espèce est dû au fait que la taille des séquences est trop courte, inférieure à 200 pb.

De plus, certaines bandes avaient une forte intensité et étaient rapprochées au niveau du gel. Pour un peu moins de la moitié (13/29), le résultat du séquençage a montré un mélange de plusieurs séquences, empêchant l'identification (**Figure 31**).



**Figure 31** : Exemple de chromatogramme issue de l'analyse d'une séquence de bande ADN issue du gel DGGE. Visualisation avec le logiciel Sequence Scanner v1.0 (Applied Biosystems).

Nous avons pu obtenir l'identification par séquençage pour 16 bandes d'ADN (**Tableau 17**).

**Tableau 17** : Identification par séquençage de la région V3 du gène de l'ARNr 16S des bandes d'ADN bactériens des profils DGGE communes aux deux pays (Mexique et Guyane).

Bande	Famille, Genre, espèce	% d'identité	% Query cover	E-value
<b>d</b>	<i>Bacillus sp</i>	92	100	2 <sup>e-15</sup>
<b>f</b>	<i>Lactobacillus fermentum</i>	98.2	99	5 <sup>e-61</sup>
<b>g</b>	<i>Lactobacillus sp</i>	99.3	90	2 <sup>e-71</sup>
<b>h</b>	<i>Lactobacillus fermentum</i>	98.45	100	5 <sup>e-61</sup>
<b>i</b>	<i>Lactobacillus fermentum</i>	96.7	100	2 <sup>e-60</sup>
<b>k</b>	<i>Lactobacillus sp</i>	98.3	100	3 <sup>e-53</sup>
<b>l</b>	<i>Lactobacillus fermentum</i>	98.4	100	3 <sup>e-58</sup>
<b>m</b>	<i>Lactobacillus sp</i>	99.3	100	2 <sup>e-70</sup>
<b>n</b>	<i>Lactobacillus sp</i>	96.7	95	3 <sup>e-74</sup>
<b>o</b>	<i>Lactobacillus sp</i>	95	96	1 <sup>e-35</sup>
<b>r</b>	<i>Acetobacteraceae</i>	100	93,2	2 <sup>e-39</sup>
<b>s</b>	<i>Pseudomonas sp</i>	95	97	7 <sup>e-14</sup>
<b>z</b>	<i>Acetobacteraceae</i>	99	100	1 <sup>e-51</sup>
<b>z2</b>	<i>Acetobacteraceae</i>	100	96	4 <sup>e-51</sup>
<b>z3</b>	<i>Acetobacteraceae</i>	96.6	99	3 <sup>e-52</sup>
<b>z4</b>	<i>Acetobacteraceae</i>	99.07	97	2 <sup>e-49</sup>

Neuf bandes sur les 16 ont été identifiées comme appartenant au genre *Lactobacillus*, l'identification au niveau de l'espèce a pu être obtenue pour 4 bandes (f, h, i, et l) comme étant l'espèce *L. fermentum*. Les 5 bandes, r, z, z2, z3 et z4 ont pu être identifiées comme appartenant à la famille des *Acetobacteraceae*. La bande d, qui était présente uniquement dans les échantillons du Mexique, a été identifiée comme un *Bacillus sp*. Sa présence a été détectée du début à la fin de la fermentation. Enfin la bande s, présente uniquement dans les profils des échantillons de Guyane a été identifiée comme étant une espèce du genre *Pseudomonas*.

Sur les 10 bandes communes entre les profils des deux pays, l'identification a pu être possible seulement sur 4 bandes (f, l, o, et r) et appartiennent aux deux familles principalement retrouvées en fermentation du cacao : *Lactobacillaceae* et *Acetobacteraceae*.

#### III.1.1.1.4 Conclusion partielle

L'analyse par PCR DGGE nous a permis d'observer une grande diversité des communautés bactériennes au sein des échantillons du Mexique et de la Guyane. Les communautés bactériennes semblent être différentes selon le pays d'origine, et selon le temps de fermentation. En effet, au cours de la fermentation, les conditions changent et favorisent différents groupes de bactéries. Il a déjà été démontré dans la littérature qu'en début de fermentation les Entérobactéries et les BAL dominant, puis à partir du second jour les BAA et les BAL sont dominantes et en fin de fermentation ce sont les *Bacillus* (De Vuyst et Weckx 2016). L'augmentation du nombre de bandes observée est expliquée par la mise en place de ces groupes bactériens, qui deviennent majoritaire au cours de la fermentation.

Concernant les différences de profils obtenus entre les échantillons des deux pays, les calculs d'indices de diversité  $\alpha$  et  $\beta$  ont montré que ces différences sont observées au niveau de la diversité bactérienne, les échantillons du Mexique présentent une plus grande diversité que ceux de Guyane. Cette diversité plus importante au Mexique peut être dû à la présence d'espèces ou de genres différents qu'on ne retrouve pas en Guyane. L'emploi de caisse en matériaux différents (bois vs en polystyrène), les différents manipulateurs peuvent expliquer cette différence.

La communauté bactérienne en Guyane doit être représenté par un seul genre ou une seule espèce, expliquant cette plus faible diversité. Cependant, l'identification partielle de l'ensemble des bandes de DGGE sélectionnées ne nous permet pas de vérifier cette hypothèse et ni de conclure quels sont les taxons qui expliquent la répartition des échantillons en fonction du pays.

La DGGE nous a permis d'avoir une vision globale de la diversité des communautés bactériennes présentes à la surface des fèves de cacao en provenance du Mexique et de la Guyane, en tenant compte de la majorité des taxons présentes dans ces communautés. Cependant, cette méthode n'est pas quantitative, car l'abondance relative des taxons ne peut pas être correctement évaluée dû à la saturation du signal lié au mode de coloration et surtout au nombre de copies du gène codant pour l'ARNr 16S qui peut être variable selon les espèces. Nous avons par la suite, appliqué une méthode de séquençage à haut débit (NGS, Next generation sequencing) afin d'avoir une

vision plus complète, en termes d'abondance et d'affiliation taxonomique des communautés bactériennes associées aux fèves de cacao lors de la fermentation.

*III.1.1.2 Analyse par metabarcoding des communautés bactériennes associées aux fèves de cacao en fermentation de 3 pays*

Les communautés bactériennes présentes à la surface des fèves de cacao et de la pulpe des échantillons provenant de trois pays : Guyane, Côte d'Ivoire et Mexique (deux campagnes d'échantillonnage) ont été analysées par la méthode de metabarcoding. Il est à noter que pour une des caisses de fermentation réalisée au Mexique en Octobre 2018 ( la caisse 3), un défaut de fermentation a été observé, elle s'est déroulée plus rapidement que les deux autres, avec un développement de moisissures en fin de fermentation, cette caisse aura été traitée comme les autres.

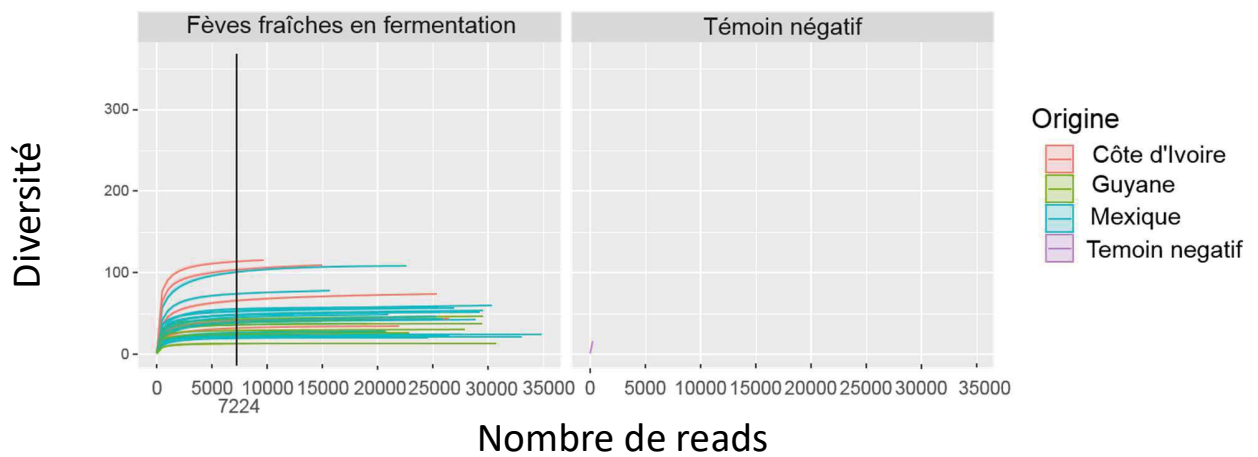
La technique de séquençage Miseq (Illumina) a été utilisée pour l'analyse des communautés bactériennes. Pour cela, les régions V3-V4 du gène de l'ARNr 16S ont été amplifiées suivant le protocole énoncé dans le chapitre, Matériels et méthodes (II.3.3.1) et les amplicons ont été envoyés à la plateforme Genseq pour la PCR d'indexation, la purification et le séquençage. L'amplification du gène de l'ARNr 16S n'a pas pu être obtenu pour les échantillons suivants : la caisse 3 au 2<sup>ème</sup> jour de fermentation (Mexique, Février 2018) et de la caisse 1 au 4<sup>ème</sup> jour de fermentation (Mexique, Octobre 2018).

Après le nettoyage des séquences (trimming) afin d'éliminer les séquences de mauvaise qualité (trimming) et les chimères, un total de 907 673 séquences a été obtenu, avec une moyenne de 23 221 séquences par échantillon.

L'analyse de la raréfaction nous indique que la richesse des échantillons a été entièrement séquencée ; Pour la plupart des échantillons leurs courbes ont atteint une phase de saturation (asymptote) indiquant que la profondeur de séquençage a été suffisante pour décrire les communautés bactériennes associées aux échantillons (**Figure 32**).

Le nombre de séquences entre les échantillons a été normalisé à 7224 OTUs, pour cette valeur, tous les échantillons ont atteint l'asymptote. Si on prend une valeur supérieure à 7224, aucune

information supplémentaire ne sera apportée. Le tableau d'abondance des OTUs a été obtenu à la suite de cette normalisation.

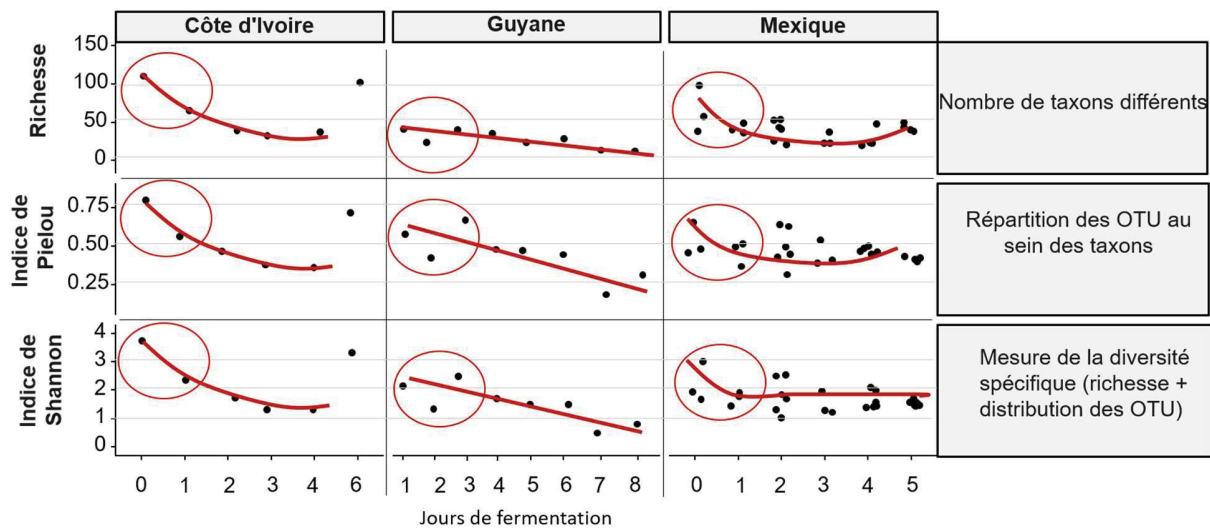


**Figure 32** : Courbes de raréfaction pour les échantillons de fèves de cacao de 3 pays, Mexique, Côte d'Ivoire et Guyane à différents de fermentation. Les courbes représentent le nombre d'unités taxonomiques opérationnelles (OTUs) détectées par échantillon en fonction du nombre de reads par échantillon.

#### *III.1.1.2.1 Comparaison de la diversité $\alpha$ des échantillons de fèves de cacao en fermentation en provenance de 3 pays*

Dans un premier temps la diversité  $\alpha$  des échantillons a été analysée, en tenant compte de la richesse observée (OTUs observées) et en utilisant les indices de Shannon, Inverse Simpson et Pielou pour chaque échantillon (**Figure 33**).





**Figure 33** : Indices de diversité  $\alpha$  des échantillons de fèves de cacao à différents temps de fermentations de de 3 pays (Mexique, Côte d’Ivoire et Guyane).

La richesse spécifique (nombre d’OTUs présentes, observées), l’indice de Pielou (mesurer la répartition des séquences au sein des OTUs), l’indice de Shannon (mesure la diversité spécifique des échantillons en tenant compte du nombre d’OTUs et de la distribution des séquences au sein de l’OTU), et l’indice d’inverse Simpson (mesure la probabilité que les séquences choisis aléatoirement dans la population soient de la même OTU).

Concernant la richesse spécifique, elle varie entre 13 OTUs (T8, Guyane) et 111 OTUs (J0, Côte d’Ivoire). De manière globale, pour l’ensemble des échantillons, le nombre d’OTUs total diminue en fonction du temps de fermentation sauf pour la Côte d’Ivoire, où l’échantillon au 6<sup>ème</sup> jour de fermentation (fin de la fermentation) présente un nombre d’OTUs (103) supérieur aux nombres d’OTUs obtenus aux autres temps de fermentation. Les échantillons du Mexique présentent un nombre d’OTUs plus important que ceux de Guyane et Côte d’Ivoire et stable au cours de la fermentation.

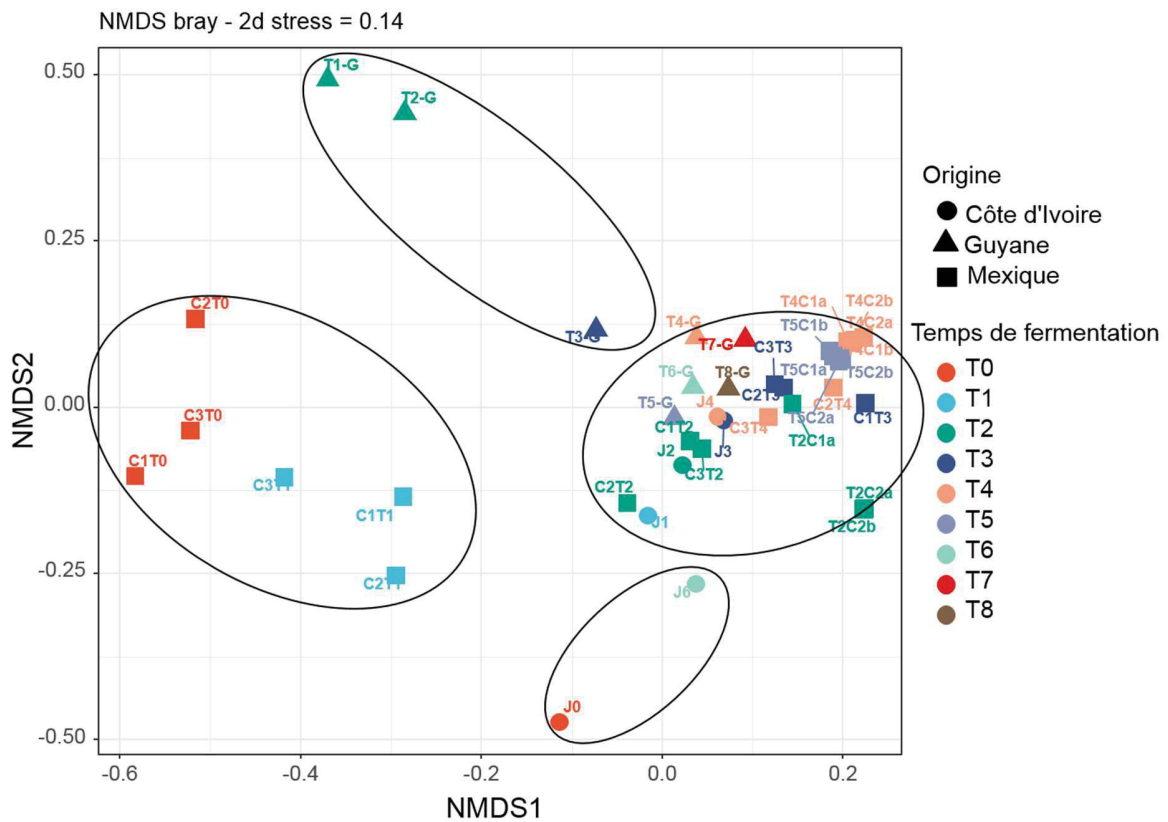
L’indice de diversité Pielou, permet de mesurer la répartition des séquences au sein des OTUs, l’abondance relative. Pour la Côte d’Ivoire, l’indice se rapproche de 0 au cours du temps, reflétant la dominance de certaines OTUs. Pour la Guyane, la diversité au T<sub>0</sub> est très faible, indiquant la dominance d’OTUs représentés majoritairement. Cet indice augmente jusqu’au 3<sup>ème</sup> jour de

fermentation indiquant une équi-répartition des séquences dans les OTUs puis celui-ci diminue jusqu'à la fin de la fermentation indiquant la dominance de certaines OTUs. Pour le Mexique, cet effet a été moins important, les indices sont autour de 0,5, avec un léger rapprochement à 0 vers la fin de fermentation, indiquant que les communautés bactériennes ne sont pas homogènes avec des OTUs sur représentées.

Les mêmes observations ont été réalisées avec les indices de Shannon et Inverse Simpson. Ces résultats indiquent que les échantillons ont une importante diversité bactérienne, qui au cours du temps de fermentation diminue. Ceci est dû à la mise en place des communautés bactériennes impliquées dans la fermentation du cacao et dominant celle-ci. L'échantillon obtenu au 6<sup>ème</sup> jour de fermentation en Côte d'Ivoire présente une plus grande diversité que ceux obtenus au Mexique et en Guyane. Cette grande diversité peut s'expliquer par la mise en place d'une communauté bactérienne plus riche, qui pourra être identifiée par la suite avec l'analyse en taxons.

#### *III.1.1.2.2 Analyse de la diversité $\beta$ des échantillons de fèves de cacao en fermentation en provenance de 3 pays*

Dans un second temps, l'analyse de la diversité  $\beta$  des échantillons a été réalisée. Cette analyse permet de comparer les échantillons entre eux. Pour cela les indices de diversité Bray-Curtis, Jaccard ont été appliqués. L'indice de diversité  $\beta$ , Bray-Curtis permet la mesure de la dissimilarité entre deux échantillons, en tenant compte des abondances relatives des OTU. Une matrice de dissimilarité (appelée aussi de distance) est obtenue puis à partir de cette matrice un positionnement multidimensionnel non-métrique est réalisé afin d'obtenir le regroupement des échantillons (**Figure 34**).



**Figure 34** : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de dissimilarité de Bray-Curtis obtenus sur les échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane) à différents temps de fermentation.

Cette figure nous montre que les échantillons sont regroupés en quatre groupes. Un groupe majoritaire qui est composé de la majorité des échantillons à partir du 3<sup>ème</sup> jour de fermentation jusqu'à la fin de la fermentation. Ce regroupement indique que les communautés bactériennes associées à ces échantillons présentent une importante similarité.

Ensuite, les échantillons au jour 0 à 2 sont regroupés dans 3 groupes selon le pays d'origine. Le premier groupe est composé des échantillons du Mexique (seconde campagne) au T<sub>0</sub> (sans fermentation) et au 1<sup>er</sup> jour de fermentation. Le deuxième groupe est composé des échantillons de Guyane du 1<sup>er</sup> au 3<sup>ème</sup> jour de fermentation. Le troisième groupe est quant à lui composé des de Côte d'Ivoire au début (J<sub>0</sub>) et à la fin (J<sub>6</sub>) de la fermentation. Ces trois regroupements indiquent

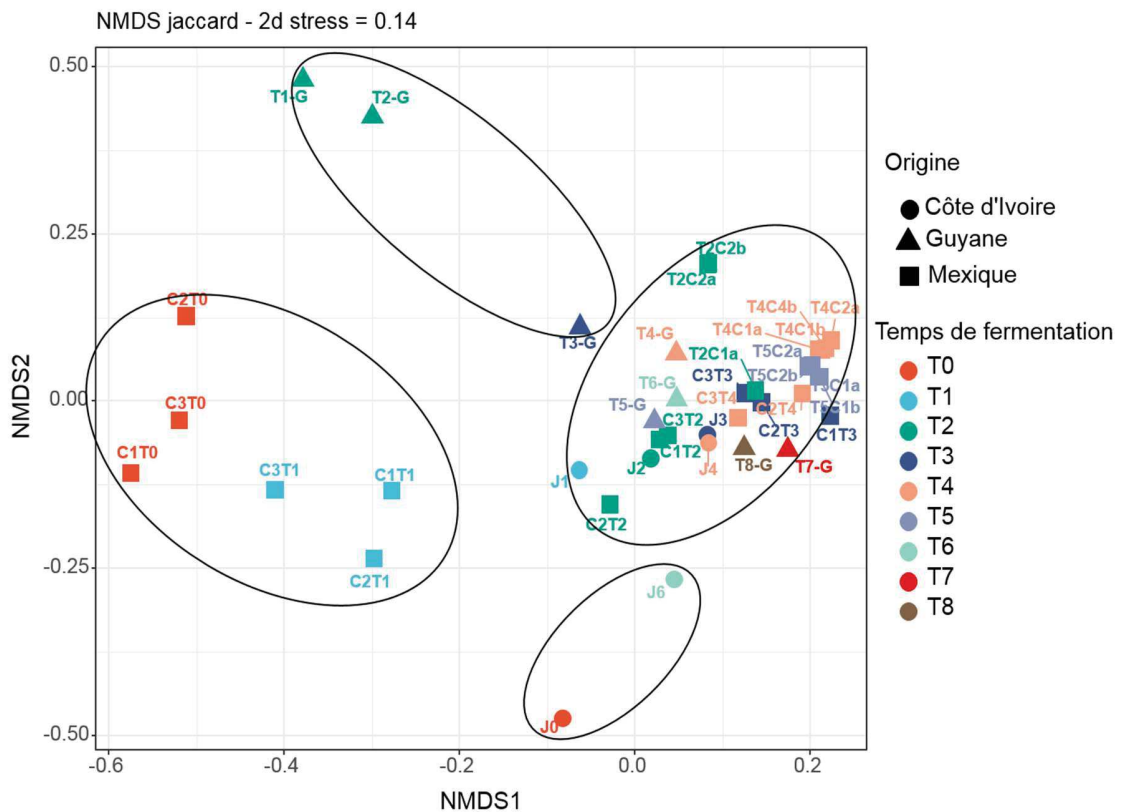
que les communautés bactériennes sont différentes aux premiers jours de la fermentation selon le pays.

Les tests réalisés sur la matrice de distance, ont montré que le pays d'origine des échantillons et le jour de fermentation ont un effet sur le regroupement des échantillons avec une p-value de 0.001\*\*\* et 0.001\*\*\* respectivement (**Tableau 18**). Le  $R^2$  du test statistique représente le pourcentage de variation de la réponse expliquée par le modèle ; plus la valeur  $R^2$  est élevée plus le modèle est ajusté aux données. Le  $R^2$  a montré que la variable « jour de fermentation » explique à elle seule 46% de la répartition des échantillons alors que la variable « pays d'origine » seulement 16 % (**Tableau 18**).

**Tableau 18** : Résultats du test statistique PERMANOVA sur la matrice de distance réalisée avec les indices de dissimilarité de Bray-Curtis obtenus pour les échantillons de fèves de cacao de 3 pays, Mexique, Côte d'Ivoire et Guyane à différents temps de fermentation

Variables	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr..F.
Pays	2.000	1.494	0.747	7.855	0.159	0.001
Jour de fermentation	8.000	4.320	0.540	5.678	0.459	0.001
Interaction des deux	8.000	1.704	0.213	2.240	0.181	0.003
	20.000	1.902	0.095		0.202	
	38.000	9.420			1.000	

Afin de tester, si cette différence au sein des communautés bactériennes est dû à l'abondance relative des séquences au sein des échantillons ou seulement à un nombre d'OTU différent, l'indice de diversité  $\beta$ , Jaccard a été calculé. Cet indice prend en compte uniquement l'absence ou la présence d'une OTU. De la même manière qu'avec l'indice de Bray-Curtis, une matrice de dissimilarité est obtenue puis à partir de cette matrice un positionnement multidimensionnel non-métrique a été réalisé afin d'obtenir le regroupement des échantillons (**Figure 35**).



**Figure 35** : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l’aide des indices de similarité Jaccard issus des échantillons de fèves de cacao de 3 pays (Mexique, Côte d’Ivoire et Guyane) à différents temps de fermentation.

Avec cet indice, les échantillons sont regroupés de la même manière qu’avec l’indice de Bray-Curtis, indiquant donc que les communautés bactériennes diffèrent selon leur composition en OTUs et non par l’abondance relative des OTUs présents dans les communautés bactériennes.

Les tests statistiques réalisés sur la matrice de distance, ont également montré un effet du pays d’origine et du jour de fermentation avec une p-value=0.0001\*\*\* (**Tableau 19**). La variable « jour de fermentation » explique le plus la répartition des échantillons avec un pourcentage de R<sup>2</sup> de 38 % alors que la variable « pays » d’origine seulement 14 %.

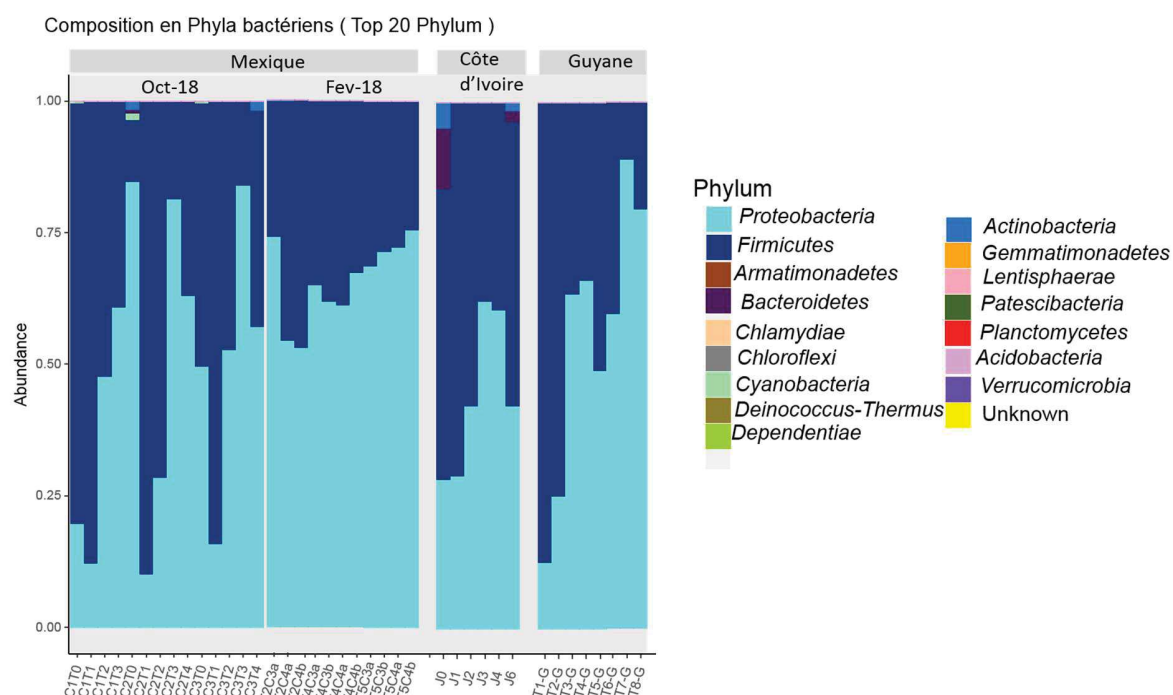
**Tableau 19** : Résultats du test statistique PERMANOVA sur la matrice de distance réalisée avec les indices de similarité de Jaccard obtenus pour les échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane) à différents temps de fermentation.

Variables	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr..F.
Pays	2.000	1.691	0.846	5.374	0.144	0.001
Jour de fermentation	8.000	4.542	0.568	3.608	0.386	0.001
Interaction des deux	8.000	2.381	0.298	1.892	0.202	0.001
	20.000	3.147	0.157		0.268	
	38.000	11.762			1.000	

*III.1.1.2.3 Comparaison de la composition des communautés bactériennes associées aux échantillons de fèves de cacao en fermentation en provenance de 3 pays*

Les séquences des OTUs dans le tableau des OTUs ont été comparées dans les bases de données (SILVA 123) enfin d'affilier la taxonomie.

Composition en phyla bactériens



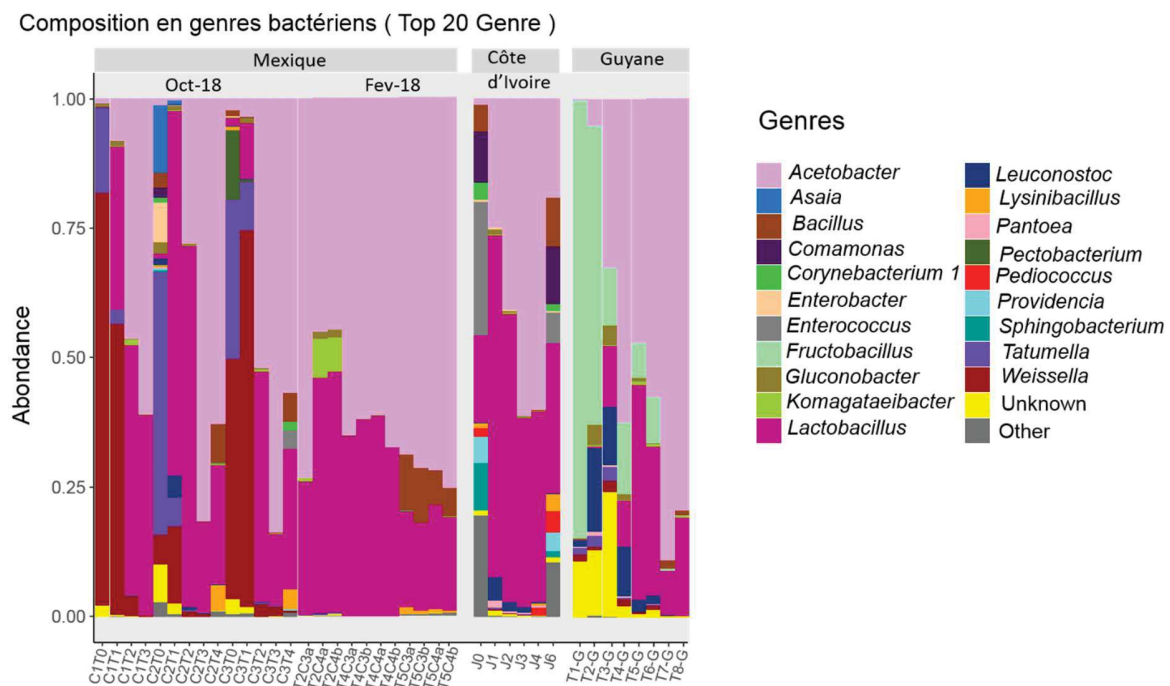
**Figure 36** : Composition des échantillons en phylum bactériens pour les échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane) à différents temps de fermentation.

Les échantillons en provenance des 3 pays sont composés d'OTUs affiliées en majorité aux phyla des *Firmicutes* et *Proteobacteria* (**Figure 36**). Pour les échantillons en provenance du Mexique (Octobre 2018), Côte d'Ivoire et de Guyane, la proportion des séquences appartenant aux *Firmicutes* est plus importante en début de fermentation et diminue au cours du temps, alors que la proportion en *Proteobacteria* augmente au cours du temps. L'échantillon J6 en Côte d'Ivoire diffère des autres échantillons, avec une proportion de *Firmicutes* plus élevée, expliquant donc la plus grande diversité observée lors de l'analyse avec les indices de diversité. Les échantillons J<sub>0</sub> et J<sub>6</sub> de Côte d'Ivoire sont également composés de séquences appartenant aux phyla *Actinobacteria* et *Bacteroidetes*, ce qui explique qu'ils ne soient pas regroupés avec l'ensemble des échantillons lors de l'analyse de la diversité  $\beta$ . Les échantillons en provenance du Mexique (Février) sont composés d'une proportion en *Proteobacteria* plus importante tout au long de la fermentation contrairement à l'ensemble des échantillons des autres campagnes. On peut également observer la présence du phylum *Actinobacteria*, présent en fin des fermentations réalisées en Octobre 2018, notamment la caisse 3, la même caisse où la fermentation s'est déroulée plus rapidement recouverte par la toile de jute. Cette proportion plus importante en *Actinobacteria* peut être expliquée par une météorologie plus humide lors de ces fermentations. La conséquence d'un environnement plus humide peut conduire au développement de moisissures, nous n'avons pas étudié les communautés fongiques liées aux fermentations de fèves de cacao, nous ne pouvons pas vérifier cette hypothèse. Le développement de moisissures est peut-être lié au développement de bactéries de la famille *Actinobacteria*.

#### Composition en genres bactériens

L'affiliation au niveau de l'espèce n'a pas été possible pour toutes les OTUs étant donné les séquences courtes fournies par le séquençage Miseq (Illumina, 2\*250 pb). Cependant, l'affiliation au niveau du genre a pu être possible pour la majorité des OTUs (**Figure 37**). La composition en genres bactériens explique la répartition des échantillons obtenu lors de l'analyse de la diversité  $\beta$  à l'aide des indices Bray-Curtis et Jaccard. En effet, à partir du 3<sup>ème</sup> jour de fermentation, les communautés bactériennes de l'ensemble des échantillons sont composées en majorité des genres *Acetobacter* et *Lactobacillus* avec en fin de fermentation le genre *Bacillus*. Le genre *Acetobacter* est présent dès le début de la fermentation dans les 3 pays et le genre *Lactobacillus*

à partir du second jour de la fermentation. L'analyse de la diversité  $\alpha$  de l'ensemble des communautés bactériennes associées aux premiers jours de fermentation a démontré une plus importante diversité spécifique. Celle-ci est expliquée par leur composition en OTUs appartenant à des genres différents et uniques selon le pays. Pour les échantillons en provenance de la première campagne au Mexique, cela n'a pas pu être démontré car nous ne disposons pas des échantillons au début et au premier jour de fermentation.



**Figure 37** : Composition des échantillons en genre bactériens pour les échantillons de fèves de cacao de 3 pays (Mexique, Côte d'Ivoire et Guyane) à différents temps de fermentation.

Concernant les échantillons en provenance du Mexique, des OTU ont pu être identifiés appartenant aux genres : *Weissella*, *Tatumella* dans les trois caisses de fermentations. Les bactéries du genre *Tatumella* sont connues pour être impliquées dans la dépectinisation de la pulpe en début de fermentation et les bactéries du genre *Weissella* sont des BAL fréquemment isolées lors de fermentations de cacao (Camu *et al.* 2008; Papalexandratou *et al.* 2011c, 2013, 2019; Bortoloni *et al.* 2016; Agyirigo *et al.* 2019; Serra *et al.* 2019). Cependant, selon la caisse de



fermentation, la proportion de ces genres est différente. En effet, le genre *Weissella* est majoritaire dans les caisses 1 et 3 (entre 50 et 75 %) alors que le genre *Tatumella* est majoritaire (50 %) dans la caisse 2. De plus, d'autres genres sont retrouvés seulement dans la caisse 2 qui présente une richesse spécifique plus importante. Dans cette, caisse 2, deux genres appartenant à la famille des BAA sont détectées : *Asaia* et *Gluconobacter* ; Ces deux genres sont connus pour être des pathogènes opportunistes de l'Homme (Alauzet *et al.* 2010). Certaines bactéries du genre *Gluconobacter* sont impliquées et fréquemment isolées lors de la fermentation des fèves de cacao (Papalexandratou *et al.* 2011a, 2011b, 2013 ; Ho *et al.* 2014a, 2015; Agyirifo *et al.* 2019). Une proportion d'OTUs est affiliée à un genre appartenant à la famille des BAL : *Leuconostoc*, qui est souvent décelé dans la fermentation du cacao (Camu *et al.* 2007; Papalexandratou 2011a,2011b,2011c,2013; Schwenninger *et al.* 2016; Ouattara *et al.* 2017; Agyirifo *et al.* 2019; Serra *et al.* 2019). Les genres *Enterobacter*, *Comamonas*, *Corynebacterium*, *Acinetobacter* et *Bacillus* sont également présents dans la communauté bactérienne. Les bactéries du genre *Enterobacter* sont retrouvées principalement dans l'intestin humain, c'est un indicateur de la mauvaise hygiène, indiquant une contamination des fèves. *Comamonas sp.* sont des bactéries environnementales retrouvées dans le sol et pouvant être pathogènes pour l'homme (Steinberg et Burd 2015). Les bactéries du genre *Corynebacterium*, sont présents de manière ubiquitaire dans tous les habitats (eau, sol, air) mais également chez les animaux (poissons, volailles) et plus particulièrement dans le microbiote cutané de l'homme et sont de rares pathogènes opportunistes (Bernard 2012). Les bactéries du genre *Acinetobacter*, sont des bactéries non fermentaires, pathogènes opportuniste de l'homme (Touchon *et al.* 2014). Ce genre a été rarement identifié lors des fermentations des cacaos et aucune étude n'a été réalisée pour déterminer le rôle qu'il pourrait avoir dans la fermentation des fèves de cacao. Les bactéries du genre *Bacillus* sont fréquemment décelées lors des fermentations du cacao, pouvant être impliquées dans la dépectinisation de la pulpe (Ouattara *et al.* 2008). La caisse 3, est composé d'un genre bactérien qui n'est pas retrouvé dans les autres caisses de fermentation : *Pectobacterium*, appartenant à la famille des *Enterobacteriaceae*, ce sont des bactéries phytopathogènes, produisant des enzymes pectolytiques permettant la dégradation des parois cellulaire de leurs hôtes (Mansfield *et al.* 2015).

La composition de la communauté bactérienne du premier jour de fermentation pour la Côte d'Ivoire ( $J_0$ ) est identique à celle du dernier jour de fermentation ( $J_6$ ). La composition des communautés bactériennes de ces échantillons est en majorité composée d'OTU appartenant aux genres, *Comamonas*, *Sphingobacterium*, *Providencia*, *Corynebacterium*, *Bacillus*, *Lysinibacillus*, et *Lactobacillus* (*Acetobacter* uniquement pour  $J_6$ ). Le nombre important de genres différents explique l'importante diversité observée pour  $J_0$  et  $J_6$  avec l'analyse à l'aide des indices de diversité  $\alpha$ . Les bactéries du genre *Bacillus* et *Lysinibacillus* ont déjà été identifiées dans les fermentations du cacao (Ardhana 2003; Moreira *et al.* 2013; Pereira *et al.* 2013a ; Illegheims *et al.* 2015; Magalhães da Veiga *et al.* 2016) et ont une potentielle activité antagoniste contre certains pathogènes du cacaoyer (*Moniliophthora roreri*, *Moniliophthora perniciosa* et *Phytophthora capsici*) (Melnick *et al.* 2011). Les bactéries du genre *Sphingobacterium* sont retrouvées dans l'environnement plus particulièrement le sol et peuvent être, pour certaines espèces des pathogènes opportunistes de l'Homme. Les bactéries du genre *Providencia*, peuvent également être des pathogènes opportunistes et sont principalement retrouvés dans l'hémolymphe d'insectes comme les drosophiles, indiquant une colonisation bactérienne des fèves de cacao par les drosophiles. Une proportion importante d'OTUs correspondante à d'autres genres minoritaires est identifiée (other, pour les genres où le nombre d'OTU représente moins de 1% de la totalité).

Enfin, les échantillons Guyanais au début de la fermentation jusqu'au 3<sup>ème</sup> jour de fermentation montrent une composition de la communauté bactérienne riche en OTUs appartenant au genre *Fructobacillus*, avec plus de 75 % des OTUs en début de fermentation puis de 10 à 20 % jusqu'au 6<sup>ème</sup> jour. Ce qui expliquerait la faible diversité observée lors de l'analyse de la diversité  $\alpha$  pour  $T_0$ . Les bactéries du genre *Fructobacillus* sont des BAL impliquées dans la fermentation du cacao mais sont généralement dominés par les bactéries du genre *Lactobacillus*. En effet, on peut voir au cours de la fermentation la proportion en *Fructobacillus* a diminué tandis que la proportion en *Lactobacillus* a augmenté. Deux autres genres de BAL sont identifiés au sein de ces communautés bactériennes: *Leuconostoc* et *Weissella*, qui sont retrouvés dans les fermentations du cacao à travers le monde, Brésil, Equateur, Côte d'Ivoire, Ghana (Camu *et al.* 2007; Papalexandratou *et al.* 2011b, 2011c; Schwenninger *et al.* 2016; Agyirigo *et al.* 2019; Serra *et al.* 2019). Le genre *Tatumella* est également identifié, ces 3 derniers genres cités sont minoritaires dans les

communautés bactériennes associées aux fèves de cacao en Guyane, puis qu'ils représentent à eux 3 moins de 3% des OTUs totales. Cependant la proportion en genre *Leuconostoc* est présent tout au long de la fermentation, avec une plus forte abondance au 2<sup>ème</sup> jour (20 %).

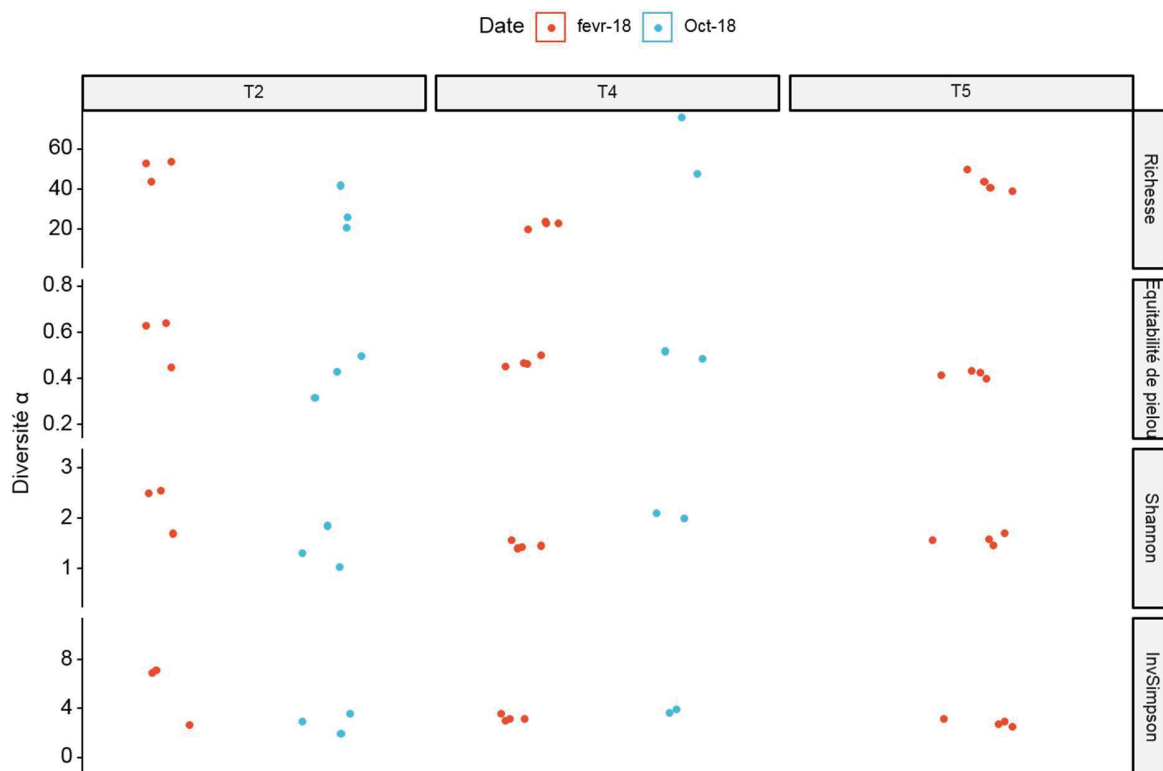
### III.1.2 Etude comparative des communautés bactériennes des fèves de cacao en fermentation en fonction de la grande et de la petite récolte au Mexique

Lors de l'analyse des communautés bactériennes selon les trois pays, on a pu constater que les échantillons du Mexique au cours d'une campagne à l'autre semblent différents, c'est pourquoi nous nous sommes intéressés dans cette partie à l'effet que pouvait avoir la période de récolte des cabosses sur la diversité des communautés bactériennes associées aux fèves de cacao en fermentation. Les échantillons de fèves de cacao en provenance du Mexique lors de la grande récolte (Février 2018) et de la petite récolte (Octobre 2018) ont été analysés. Nous avons gardé uniquement les échantillons présents aux jours de fermentation : 2,4 et 5 jours (pour Mexique-Février 2018) pour les deux campagnes. En Octobre 2018, la fermentation a duré 4 jours au lieu de 5, dû aux fèves qui étaient moins chargées en pulpe. Les analyses proviennent du séquençage réalisé dans la partie précédente par la méthode de metabarcoding. Au total, 424 801 séquences ont été obtenues, avec une moyenne de 24 988 séquences par échantillon.

#### *III.1.2.1 Comparaison de la diversité $\alpha$ des échantillons des fèves de cacao en fermentation en fonction de la grande et de la petite récolte au Mexique*

La diversité spécifique des échantillons a été analysée avec les mêmes indices que dans la partie précédente (III.1.1.2.1) (**Figure 38**). Au 2<sup>ème</sup> jour de fermentation, les communautés bactériennes des fèves de cacao en provenance de la grande récolte (Fév-18) semblent avoir une diversité plus importante que celles associées aux fèves de cacao de la petite récolte (Oct-18). En effet, le nombre d'OTU observé (Richesse spécifique) a été de 42, 26 et 21 pour les échantillons d'Oct-18 (caisse 1, 2 et 3 respectivement) et de 44, 53 et 54 pour ceux de Fév-2018 (caisse, les indices de diversité Shannon, Pielou et Inverse Simpson ont montré également une meilleure diversité pour

la grande récolte que la petite récolte. Les échantillons présentent comme valeurs pour l'indice de Pielou : 0,49, 0,32 et 0,43 (Oct-18) et 0,45, 0,64 et 0,63 (Fev-18), pour l'indice de Shannon : 1,85, 1,03 et 1,31 (Oct-18) et 1,7, 2,55 et 2,51 (Fev-18) et pour l'indice Inverse Simpson : 3,57, 1,95 et 2,97 (Oct-18) et 2,65, 7,17 et 6,93 (Fev-18) indiquant une meilleure répartition des séquences au sein des OTUs et une diversité spécifique plus importante dans les échantillons de Fev-18 que ceux d'Oct-18. Cependant au 4<sup>ème</sup> jour, on a observé que la diversité spécifique des échantillons d'Oct-18 a été plus importante que celle dans les échantillons de Fev-18, avec une richesse spécifique plus importante, des valeurs d'indice de Pielou, Shannon et Inverse Simpson plus élevées.



**Figure 38** : Indices de diversité  $\alpha$  des échantillons de fèves de cacao à différents temps de fermentation en provenance du Mexique.

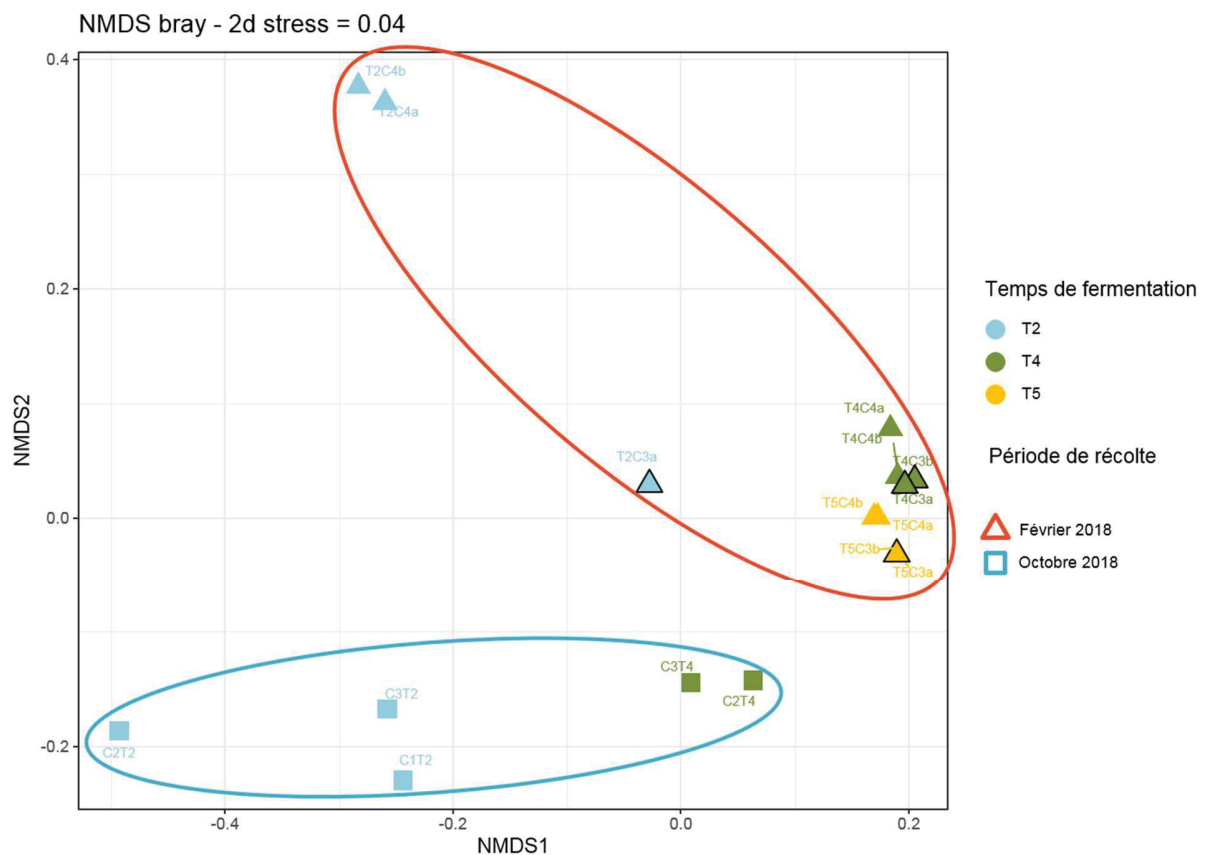
Grande récolte en févr-18 et petite en Oct-18

La richesse spécifique (nombre d'OTUs présentes, observées), l'indice de Pielou (mesurer la répartition des séquences au sein des OTUs), l'indice de Shannon (mesure la diversité spécifique des échantillons en tenant compte du nombre d'OTUs et de la distribution des séquences au sein de l'OTU), et l'indice d'inverse Simpson (mesure la probabilité que les séquences choisies aléatoirement dans la population soient de la même OTU).

### III.1.2.2 Analyse de la diversité $\beta$ des échantillons des fèves de cacao en fermentation en fonction de la grande et de la petite récolte au Mexique

L'analyse de la diversité  $\beta$  des échantillons a été réalisée. Cette analyse permet de comparer les échantillons entre eux. Pour cela les indices de diversité Bray-Curtis, Jaccard ont été appliqués, comme précédemment (III.1.1.2.2).

L'indice de diversité  $\beta$ , Bray-Curtis permet la mesure de la dissimilarité entre deux échantillons, en tenant compte des abondances relatives des OTU. Une matrice de dissimilarité (appelée aussi de distance) est obtenue puis à partir de cette matrice un positionnement multidimensionnel non-métrique est réalisé afin d'obtenir le regroupement des échantillons (**Figure 39**).



**Figure 39** : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de dissimilarité de Bray-Curtis pour les échantillons de fèves de cacao à différents temps de fermentation en provenance du Mexique, grande et petite récolte.

A partir de cette figure, nous avons pu constater que les échantillons sont regroupés par rapport à la période de récolte (Février 2018 ou Octobre 2018), on a pu voir également qu'au sein de cette répartition par pays, les échantillons se regroupaient en fonction du jour de fermentation. Avec pour la campagne en Février 2018 (grande récolte) les échantillons au 4<sup>ème</sup> et 5<sup>ème</sup> jour regroupés ce qui indique une forte similarité entre les communautés bactériennes, ces résultats ont été obtenus également lors de l'analyse des communautés par la méthode de fingerprinting, DGGE.

Les échantillons d'une caisse de fermentation au deuxième jour (T2C4 a/b) de la grande récolte sont éloignés des autres, indiquant une proportion différente en genre. L'hypothèse émise plus haut (partie I.1.1) selon laquelle la masse de fèves ait été plus aérée lors de ces fermentations pourrait expliquer ce regroupement.

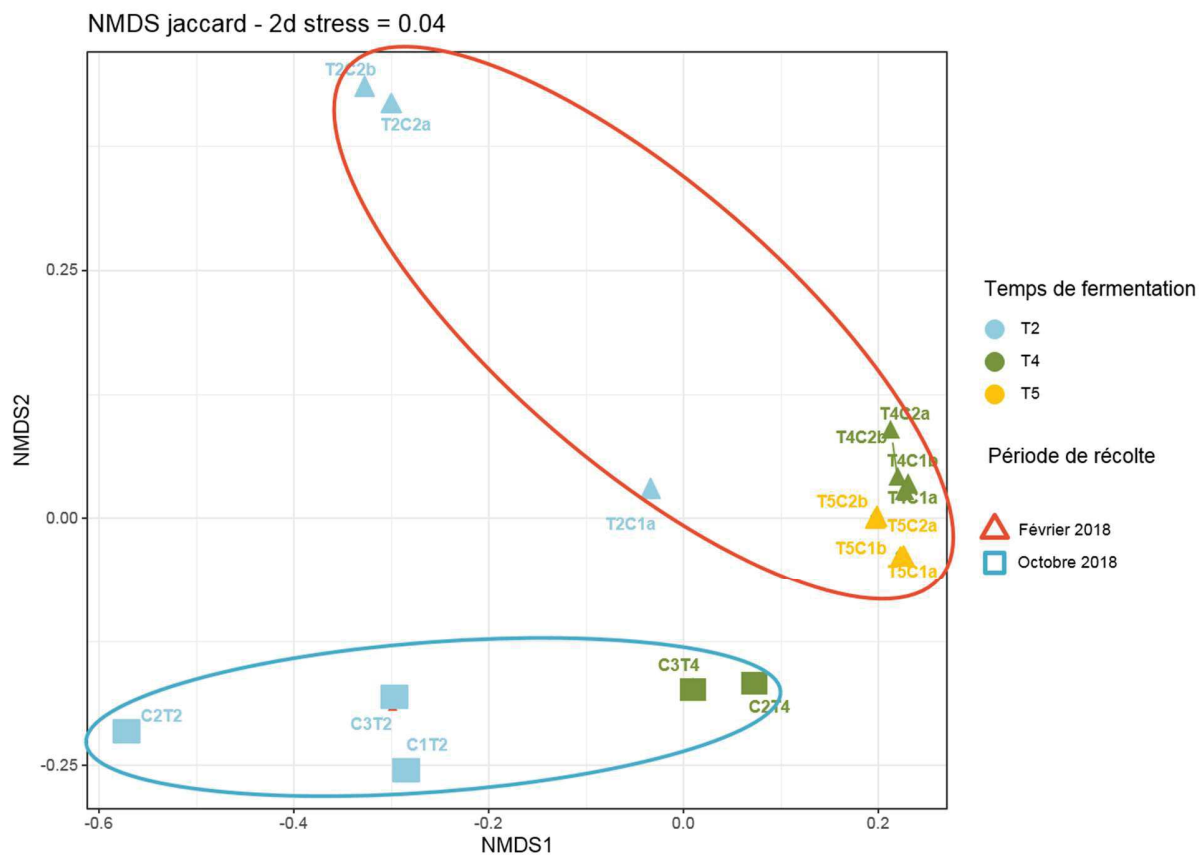
Les tests statistiques réalisés sur la matrice de distance, ont confirmé que le pays d'origine des échantillons et le jour de fermentation avaient un effet sur le regroupement des échantillons avec une p-value de 0.001\*\*\* et 0.001\*\*\* respectivement (**Tableau 20**). Le R<sup>2</sup> (représente le pourcentage de variation de la réponse expliqué par le modèle, plus la valeur R<sup>2</sup> est élevée, plus le modèle est ajusté aux données), du test statistique a montré que la variable « jour de fermentation » a expliqué 38,8 % de la répartition des échantillons et que la variable « période de récolte » 28 %.

**Tableau 20** : Résultats du test statistique PERMANOVA sur la matrice de distance réalisée avec les indices de dissimilarité de Bray-Curtis des échantillons de fèves de cacao à différents temps de fermentation en provenance du Mexique, grande et petite récolte.

Variable	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr..F.
Période de récolte	1.000	0.470	0.470	16.099	0.280	0.001
Jour de fermentation	2.000	0.651	0.326	11.148	0.388	0.001
Interaction des deux	1.000	0.237	0.237	8.117	0.141	0.001
	11.000	0.321	0.029		0.191	
	15.000	1.680			1.000	

Afin de tester, si cette différence au sein des communautés bactériennes était dû à l'abondance relative des séquences au sein des échantillons ou seulement au nombre d'OTUs différent. L'indice de diversité  $\beta$ , Jaccard a été employé. Cet indice prend uniquement en compte l'absence ou la présence d'une OTU. De la même manière qu'avec l'indice de Bray-Curtis, une matrice de

dissimilarité est obtenue puis à partir de cette matrice un positionnement multidimensionnel non-métrique est réalisé afin d'obtenir le regroupement des échantillons (**Figure 40**). Le même regroupement des échantillons a été obtenu avec cet indice qu'avec l'indice de Bray-Curtis, indiquant que les communautés bactériennes sont différentes selon leur composition en OTUs et non de la répartition des séquences au sein des OTUs.



**Figure 40** : Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de dissimilarité de Jaccard pour les échantillons en provenance du Mexique, grande et petite récolte.

Les tests statistiques réalisés sur la matrice de distance de Jaccard, confirment que le pays d'origine des échantillons et le jour de fermentation avaient un effet sur le regroupement des échantillons avec une p-value de 0.001\*\*\* et 0.001\*\*\* respectivement (**Tableau 21**).

**Tableau 21** : Résultats du test statistique PERMANOVA sur la matrice de distance réalisée avec les indices de dissimilarité de Jaccard des échantillons de fèves de cacao à différents temps de fermentation en provenance du Mexique, grande et petite récolte.

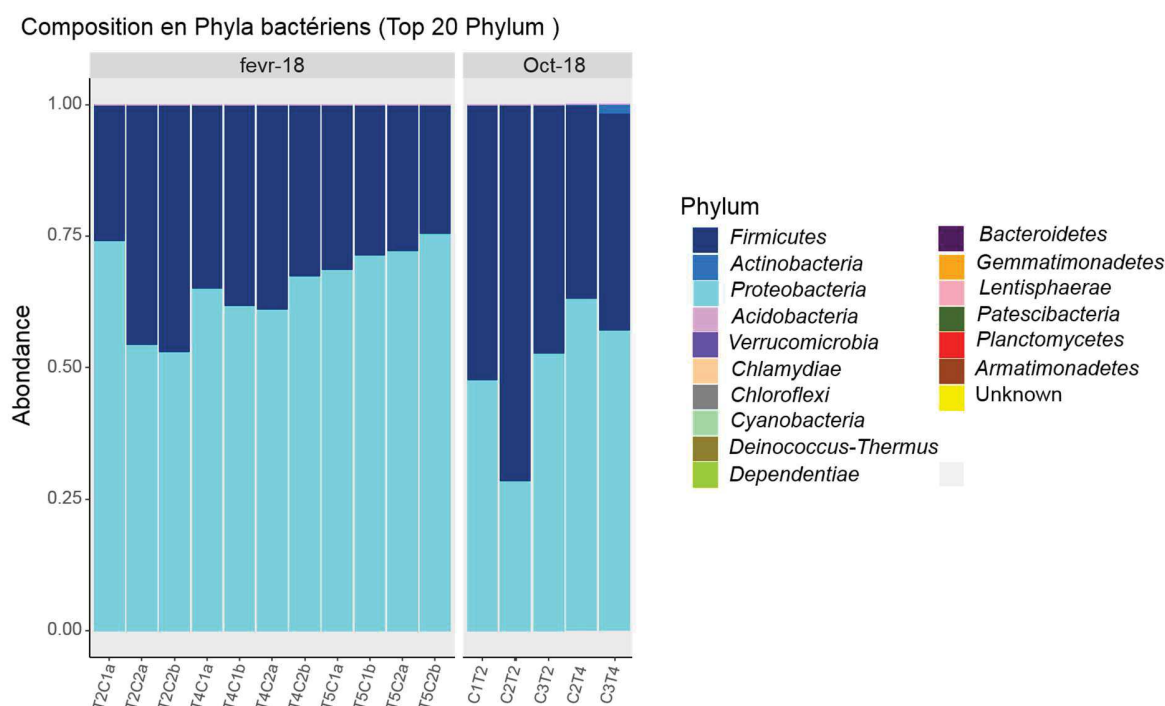
Variable	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr..F.
Période de récolte	1.000	0.690	0.690	10.927	0.248	0.001
Jour de fermentation	2.000	0.996	0.498	7.888	0.358	0.001
Interaction des deux	1.000	0.399	0.399	6.323	0.144	0.001
	11.000	0.695	0.063		0.250	
	15.000	2.780			1.000	

### III.1.2.3 Comparaison de la composition des communautés bactériennes des échantillons des fèves de cacao en fermentation en fonction de la grande et de la petite récolte au Mexique

#### Composition en phyla bactériens

Les échantillons en provenance des deux récoltes du Mexique sont composés d'OTUs affiliées en majorité aux phyla des *Firmicutes* et *Proteobacteria* (**Figure 41**). Cependant, on a pu observer que les échantillons en provenance de février 2018 (grande récolte) sont composés en majorité d'OTU affiliées au phyla *Proteobacteries* (de 55 à 75 %) contrairement aux échantillons en provenance de la petite récolte, où la proportion de *Firmicutes* est plus importante. Au 4<sup>ème</sup> jour de la fermentation, pour les échantillons de la petite récolte, le phylum *Actinobactérie* est également présent mais à un faible pourcentage (moins de 5 %).





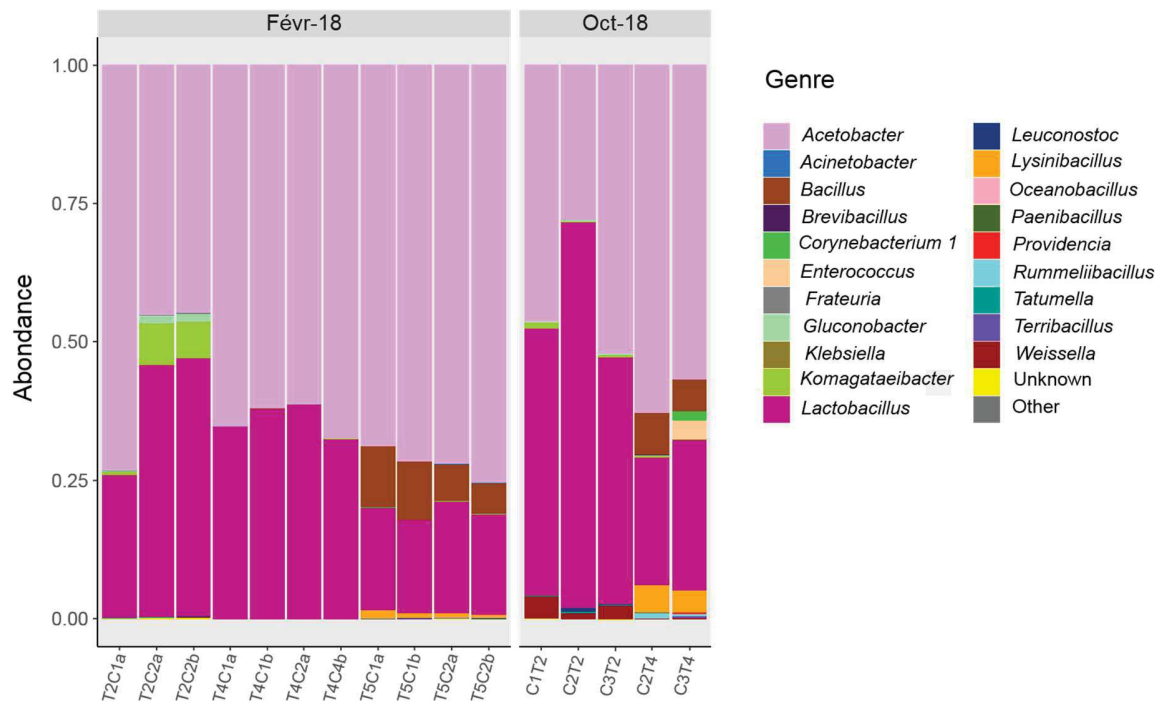
**Figure 41** : Composition en phylum des échantillons en provenance des deux récoltes réalisées au Mexique, grande (fevr-18) et petite récolte Oct-18).

### Composition en genres bactériens

L'analyse de la composition des communautés bactériennes en genre (**Figure 42**) indique que les échantillons de fèves fermentées sont composés majoritairement d'OTU affiliées aux genres *Acetobacter* et *Lactobacillus*. Cependant, les proportions de ces deux genres ne sont pas identiques selon la période de récolte. En effet, nous pouvons constater que pour la plupart des échantillons de la grande récolte (février 2018), les OTU affiliées au genre *Acetobacter* sont plus abondantes de 56,2% à 75 % tandis que pour les échantillons en provenance de la petite récolte (octobre 2018) la proportion d'*Acetobacter* est de 25,5 % à 62,5 %. Cette analyse montre également la présence des genres *Komagataeibacter* et *Gluconobacter* dans la caisse 2 de la période (février 2018), au second jour de fermentation. Les bactéries appartenant à ces genres sont des BAA. Les bactéries du genre *Komagataeibacter* sont connues pour leur capacité à sécréter des exo polysaccharides (Zhang *et al.* 2017). L'espèce *Komagataeibacter hansenii* a été identifiée pour la première fois dans les fermentations de fèves de cacao réalisées au Brésil (Serra

et al. 2019). Enfin, les échantillons en provenance de la petite récolte (Oct-18) ont un nombre de genre plus important que les échantillons de la grande récolte. En effet, au 2<sup>ème</sup> de fermentation, des OTUs affiliées au genre *Weissella*, et *Leuconostoc* sont présentes contrairement à la grande récolte. Et au 4<sup>ème</sup> jour de fermentation, des OTUs affiliée au genre *Lysinibacillus* pour les deux caisses de fermentation sont identifiées, le genre *Rummeliibacillus* est présent uniquement dans la caisse 2 et les genres *Enterococcus* et *Corynebacterium* sont présents dans la caisse 3 sûrement due à des contaminations par les manipulateurs des fèves de cacao et les caisses de fermentation, par la suite cette hypothèse sera vérifiée dans la partie III.2.3.

Composition en genres bactériens (Top 20 top Genre)



**Figure 42** : Composition en genre des échantillons en provenance des deux récoltes réalisées au Mexique, grande (fevr-18) et petite récolte Oct-18).

Cette analyse démontre que les communautés microbiennes associées aux fèves de cacao en fermentation diffèrent selon l'abondance des genres principaux impliqués dans les fermentations (*Acetobacter* et *Lactobacillus*). Ce qui explique le regroupement des échantillons lors de l'analyse de la diversité  $\beta$ , ce sont les genres présents uniquement dans une période de récolte que dans

l'autre. L'hypothèse étant que le matériel employé pour réaliser les fermentations (hypothèse vérifiée dans la partie III.2) participe à l'ensemencement des fèves, il est possible qu'entre les deux périodes les manipulateurs ont été en interaction différente entre les différents éléments employés ce qui expliquerait cette différence dans le nombre de genres différents impliqués. Il est possible également qu'en février 2018, la masse de fèves mise en fermentation ait été plus aérée, avec une entrée d'air plus importante, favorisant ainsi la famille des AAB : *Acetobacter*, *Komagataibacter* et *Gluconobacter* au détriment des LAB... De même, on ne peut pas être affirmé que les fermentations se sont déroulées de la même manière : avec les mêmes manipulateurs, les mêmes caisses, etc. Pour la période d'octobre 2018, j'ai pu contrôler et vérifier certains paramètres qu'il m'a été impossible de vérifier pour la période de février 2018.

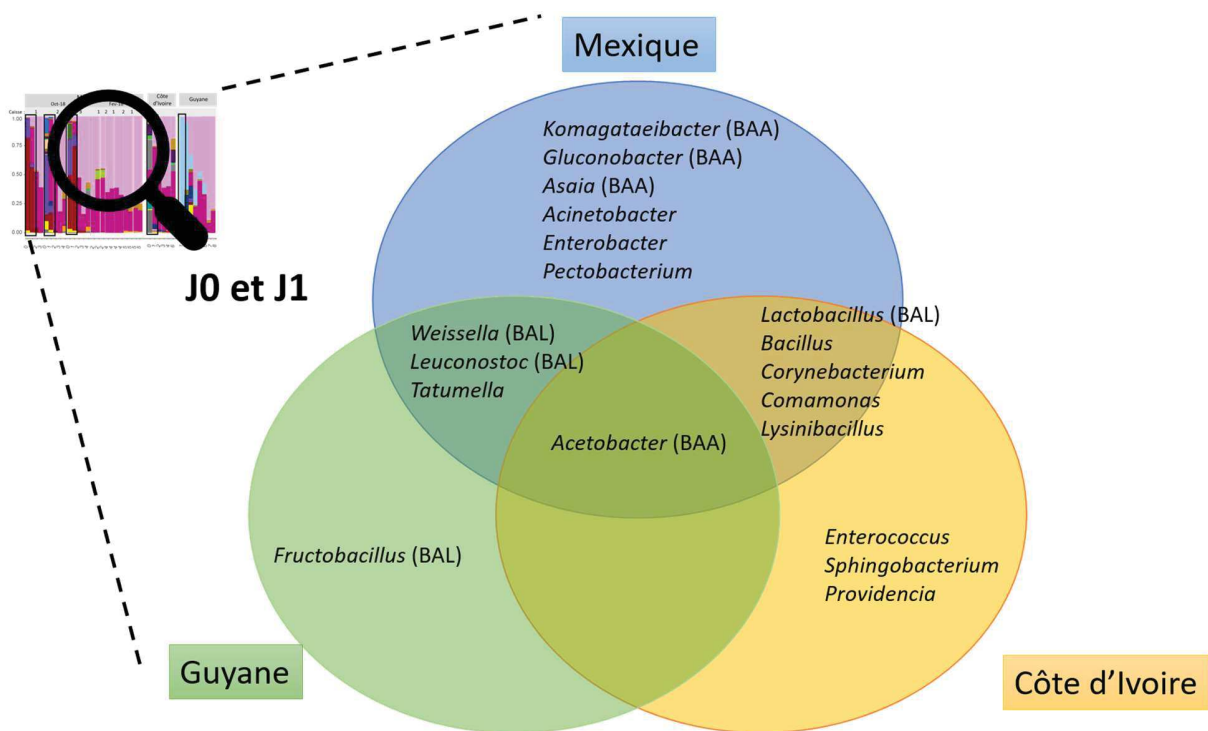
### III.1.3 Conclusion partielle

L'analyse des profils obtenus par PCR-DGGE des communautés bactériennes associées aux fèves de cacao fermentées du Mexique et Guyane nous a permis de démontrer une grande diversité au sein de ces communautés qui diminue au cours du temps. L'analyse de la diversité  $\alpha$  des échantillons de fèves de cacao indique que les communautés bactériennes associées aux fermentations au Mexique sont plus diversifiées que celle réalisée en Guyane. Cependant la difficulté d'identification des séquences des OTUs ne nous a pas permis de conclure quels sont les taxons qui expliquent cette différence.

L'analyse par metabarcoding des communautés bactériennes associées aux fèves de cacao fermentées selon les 3 pays, nous a permis de mettre en évidence une importante diversité des échantillons et plus particulièrement au début des fermentations. Les regroupements des échantillons réalisés à l'aide des indices de dissimilarité / similarité de Bray-Curtis et Jaccard, nous ont démontrés que l'ensemble des échantillons à partir du 3<sup>ème</sup> jour de fermentation avaient une communauté bactérienne similaire avec une prédominance des genres *Lactobacilles* et *Acetobacter* puis *Bacillus* (en fin de fermentation). La différence entre les communautés bactériennes observées lors de l'analyse par PCR-DGGE est donc due à la présence d'espèces

différentes au sein de ces genres principaux. En revanche, l'affiliation des OTUs obtenues par la méthode de metabarcoding n'a pas été possible au niveau de l'espèce. ; cela ne nous permet pas de conclure qu'elles sont les espèces impliquées.

Cependant, avant que ces genres bactériens ne se mettent en place et dominent les fermentations, on a pu constater que dans chaque pays, en début de fermentation, les genres bactériens retrouvés en majorité ne sont pas les mêmes et sont spécifiques à chaque pays (**Figure 43**).



**Figure 43** : Diagramme de Venn montrant la composition des communautés bactériennes en genre en début de fermentation selon les 3 pays (Mexique, Côte d'Ivoire et Guyane)

Seul le genre *Acetobacter* est retrouvé parmi les débuts de fermentation des 3 pays. Les genres *Leuconostoc*, *Weissella* et *Tatumella* sont présents au Mexique et en Guyane. Cependant, les genres *Weissella* et *Tatumella* sont prédominants au Mexique alors qu'en Guyane ils représentent une portion minoritaire des OTUs totales. Les communautés bactériennes du Mexique et de la

Côte d'Ivoire, partagent 5 genres : *Corynebacterium*, *Bacillus*, *Comamonas*, *Lactobacillus*, *Lysinibacillus*.

Cependant, certains genres ne sont retrouvés que dans les communautés bactériennes des fèves de cacao d'un seul pays. En Guyane, c'est le genre *Fructobacillus* qui est représenté en majorité, ainsi que le genre *Leuconostoc*, présent en proportion minoritaire, au Mexique, ce sont les genres *Asaia*, *Komagateibacter*, *Acinetobacter*, *Gluconobacter* et *Pectobacterium* et en Côte d'Ivoire ce sont les genres *Enterococcus*, *Sphingobacterium* et *Providencia* ainsi que d'autres genres minoritaires. Nous avons pu également démontrer un effet de la période de récolte sur les communautés bactériennes associées aux fèves de cacao au Mexique, expliqué par la présence de genre bactériens en second jour de fermentation différents selon la période. Une différence en proportions des deux genres retrouvés en majorité lors des fermentations (*Acetobacter* et *Lactobacillus*) a également été observé ; *Acetobacter* est retrouvé en de plus fortes proportions lors de la fermentation de la grande récolte.

L'analyse par metabarcoding des communautés bactériennes associées aux fèves de cacao fermentées selon deux périodes de récolte au Mexique, nous a permis de mettre en évidence une différence d'abondance des genres *Lactobacillus*, *Lysinibacillus* et *Komagateibacter* et la présence de genre différents selon la période de récolte. Cela peut être dû par des différences de pratiquement au sein d'un même centre de fermentation : avec une aération plus grande de la masse de fèves entre les deux périodes, avec également une période de pluie très importante qui a précédé les fermentations en octobre 2018. Il est impossible d'affirmer que les conditions ont été les mêmes : avec les mêmes manipulateurs, les mêmes caisses, etc. Pour la période d'octobre 2018, j'ai pu contrôler et vérifier certains paramètres qu'il m'a été impossible de vérifier pour la période de février 2018.

### III.2 Détermination de l'origine du pool de bactéries impliqué dans les fermentations de fèves de cacao : au niveau de la plantation ou de l'environnement de la fermentation du Mexique.

Lors du chapitre précédant nous avons mis en évidence que la structure des communautés bactériennes présente à la surface des fèves de cacao et de la pulpe en fermentation est similaire à partir du 3<sup>ème</sup> jour de fermentation selon le pays d'origine où la fermentation s'est déroulée. Les genres *Acetobacter* et *Lactobacillus* dominent les fermentations. Cependant, aux premiers jours et au dernier jour de la fermentation, nous avons pu constater que ces communautés ne sont pas similaires, et sont constitués de taxons spécifiques à chaque pays. A l'ouverture des cabosses, les fèves sont stériles et sont contaminées par l'environnement, mais la contribution de chaque élément du process n'est que peu répertoriée dans la littérature. Dans ce chapitre nous nous sommes intéressés aux communautés bactériennes des éléments de l'environnement du processus de récolte des cabosses et de la mise en fermentation des fèves et de la pulpe de cacao.

Cela nous permettra de répondre à la question de recherche suivante : quelle est l'origine des bactéries retrouvées dans les fermentations des fèves de cacao ?

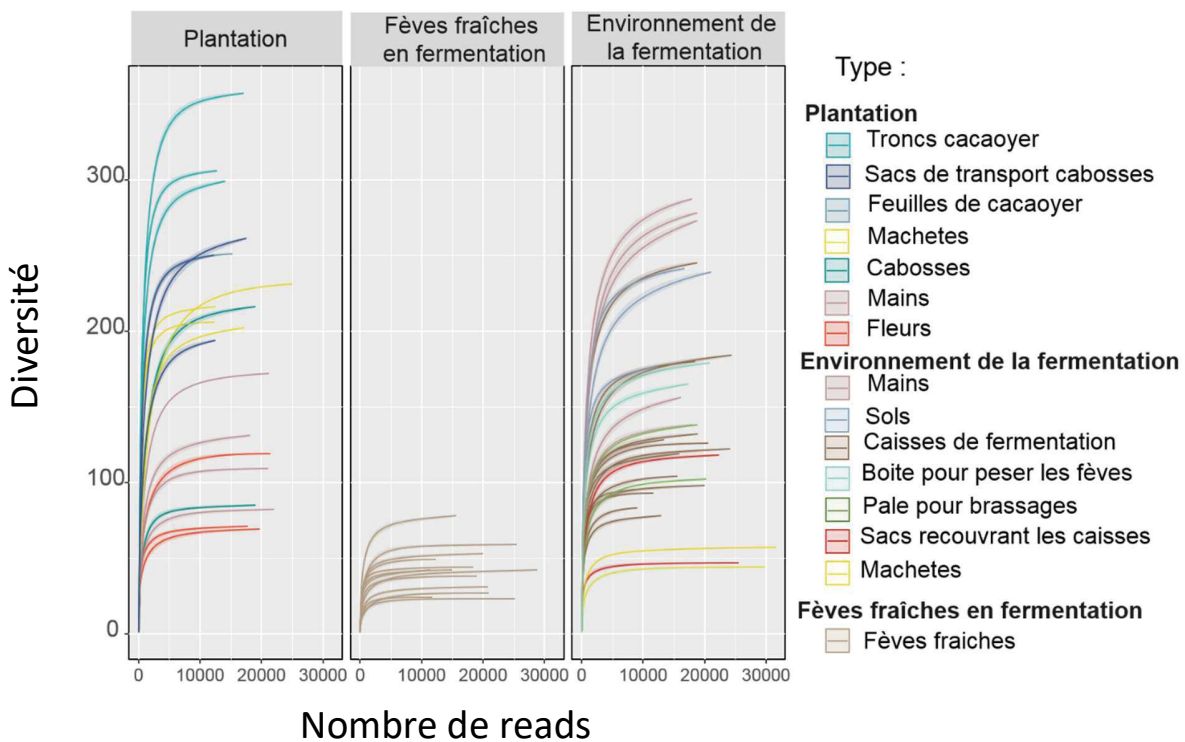
Pour répondre à cette question de recherche, des prélèvements environnementaux ont été réalisés lors de la campagne au Mexique en Octobre 2018. Vingt prélèvements provenant des surfaces (cabosses, troncs, mains, etc.) liés à la plantation et 3 prélèvements correspondant aux 3 sacs utilisés pour le transport des cabosses de la plantation à la zone de fermentation ont été réalisés. Trente-quatre prélèvements de l'environnement liés à la fermentation ainsi que 15 échantillons des fèves en fermentation du jour 0 au 4<sup>ème</sup> jour selon trois caisses différentes ont également été effectués. Ensuite, les ADN extraits par la méthode de phénol/ chloroforme de ces prélèvements ont été analysés par le séquençage haut débit de nouvelle génération (Illumina). Les régions V3-V4 du gène de l'ARNr 16S ont été amplifiées suivant le protocole énoncé dans le chapitre, Matériels et méthodes (II.3.3.1) et les amplicons ont été envoyés à la plateforme Genseq pour la PCR d'indexation, la purification et le séquençage. Les indices de diversité  $\alpha$  et  $\beta$  ont été calculés, puis la composition en phyla et genres bactériens a été déterminée.

### III.2.1 Comparaison de la diversité $\alpha$ des prélèvements environnementaux et des fèves de cacao en fermentation

Des extraits d'ADN ont été obtenus dont les concentrations varient de 0,256 à 49,1 ng/ $\mu$ L. Les concentrations les plus importantes correspondent aux extractions réalisées sur les fèves de cacao et les plus faibles sur les machettes et les mains des récolteurs et écabosseurs. L'amplification par PCR n'a pas pu être obtenue pour 6 échantillons et n'ont donc pas été inclus dans l'analyse : C1T4 (caisse 1 de fermentation au 4<sup>ème</sup> jour de fermentation), M5b (main gauche écabosseur), Fe2/Fe3 (feuilles de cacaoyer), C3a/C3d, (fond et Sud de la caisse 3 avant fermentation).

Après le nettoyage des séquences obtenues, pour éliminer les séquences de mauvaise qualité et les chimères, un total de 1 187 255 séquences a été obtenu, avec une moyenne de 17 988 séquences par échantillon. Les échantillons sont répartis en trois groupes : plantation, fèves fraîches en fermentation et environnement de la fermentation.

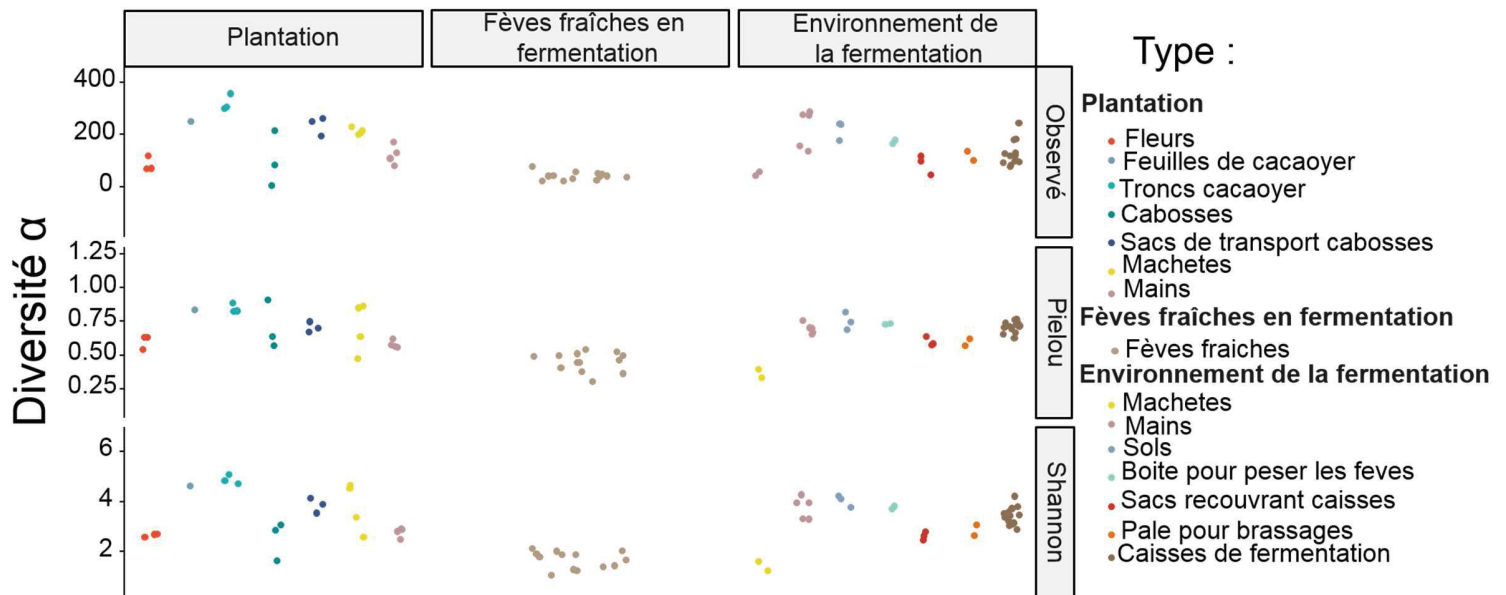
L'analyse de la raréfaction nous permet de dire si la richesse des échantillons a été entièrement séquencée, pour la plupart des échantillons, leurs courbes ont atteint une phase de saturation (asymptote) indiquant que la profondeur de séquençage a été suffisante pour décrire les communautés bactériennes associées aux échantillons (**Figure 44**).



**Figure 44** : Courbes de raréfaction pour chaque échantillon des prélèvements environnementaux et des fèves de cacao en fermentation. Les courbes représentent le nombre d'unités taxonomiques opérationnelles (OTUs) détectées par échantillon en fonction du nombre de reads.

La diversité  $\alpha$  des échantillons a été analysée, en tenant compte de la richesse observée (nombre d'OTUs) et en utilisant les indices de Shannon, Inverse Simpson et Pielou pour chaque échantillon (Figure 45).





**Figure 45 :** Indices de diversité  $\alpha$  des échantillons de fèves de cacao en fermentation en provenance de la plantation, des fèves fraîches en fermentation et de l’environnement de la fermentation.

La richesse spécifique (nombre d’OTUs présentes, observées), l’indice de Pielou (mesurer la répartition (abondance) des séquences au sein des OTUs), l’indice de Shannon (mesure la diversité spécifique des échantillons en tenant compte du nombre d’OTUs et de la distribution des séquences au sein de l’OTU), et l’indice d’inverse Simpson (mesure la probabilité que les séquences choisies aléatoirement dans la population soient de la même OTU).

Concernant, la richesse spécifique, les communautés bactériennes sont plus riches au niveau de l’environnement de la plantation, avec un nombre d’OTUs différent allant de 61 (F2, fleur) à 342 OTUs (T1, tronc) que de celles de l’environnement de la fermentation avec un nombre d’OTUs allant de 41 (Mqa, machette) à 244 (M5a, main). Les communautés bactériennes des fèves en fermentation sont moins riches, avec un nombre d’OTUs allant de 21 (C3T2) à 99 (C2T0).

Au niveau, de l’environnement de la plantation, ce sont les communautés bactériennes associées aux prélèvements réalisés sur les troncs d’arbres qui ont un plus grand nombre d’OTUs différents (289, 305 et 342 OTUs) et celles associées aux fleurs, qui ont le nombre le moins important (61 et 69 OTUs). Au niveau de l’environnement de la fermentation, les communautés bactériennes associées aux mains des écabosseurs, ont une richesse spécifique plus importante avec un

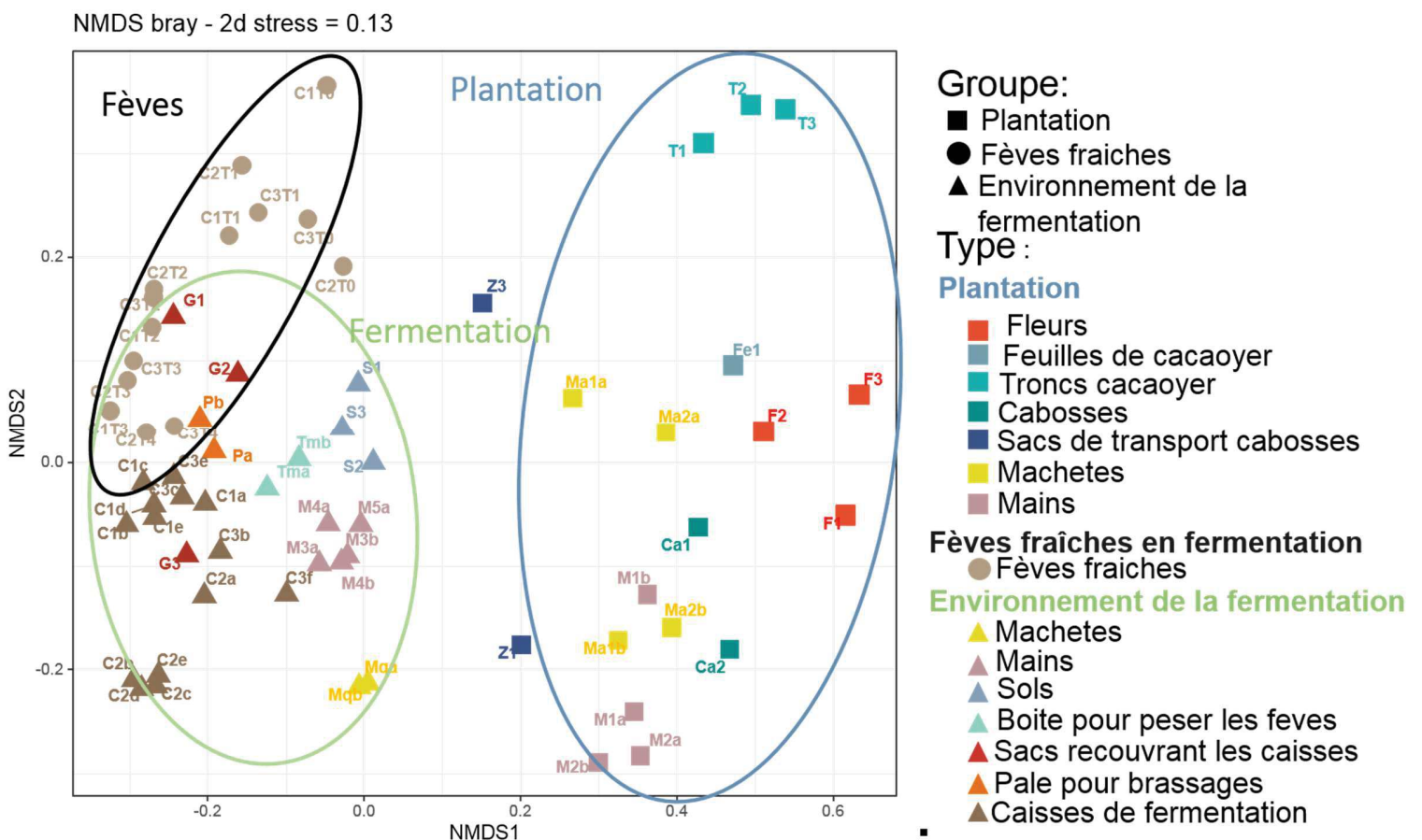
nombre d'OTUs de 113, 134, 237, 237 et 244. On peut également voir que les communautés bactériennes associées aux machettes utilisées pour la récolte des cabosses au champ ont une diversité plus importante que celle des machettes utilisées pour l'écabossage des cabosses.

Concernant, la diversité des communautés à la surface des 3 cabosses, celles-ci comportent un nombre variable d'OTUs selon la cabosse (190, 80 et 40). L'indice de diversité Pielou, permet de mesurer l'abondance des séquences au sein des OTUs. Pour la plupart des prélèvements, l'indice est supérieur à 0,5 indiquant que les séquences ont tendance à être réparties de manière homogène au sein des différentes OTUs et que la communauté bactérienne est représentée par un nombre important d'OTUs différentes. Cependant, cette répartition homogène des OTUs au sein des communautés bactériennes est plus importante au niveau des prélèvements réalisés dans l'environnement de la fermentation (hormis, la machette, la pale et les sacs recouvrant les fèves) et pour les prélèvements sur les troncs de cacaoyer, une cabosse et une machette. En effet les indices sont proches de 1 (entre 0,63 et 0,85). L'indice de Pielou pour la machette utilisée pour l'écabossage des cabosses au niveau de l'environnement de la fermentation, montre que sa communauté bactérienne n'est pas diversifiée et n'est pas homogène dans la répartition du nombre de séquences obtenues dans les différents OTUs. Avec l'indice de Shannon, la même conclusion peut être réalisée sur l'ensemble des prélèvements. Les communautés bactériennes associées présentent une grande diversité spécifique avec un indice de Shannon supérieur à 2. Cependant, les communautés bactériennes des fèves fraîches en fermentation ont une diversité spécifique plus faible que celles issues des prélèvements au niveau de la plantation et de l'environnement de la fermentation. Cela est dû au nombre d'OTUs différents moins important et la répartition non homogène des séquences au sein de ces OTUs que pour les autres communautés bactériennes.

### III.2.2 Analyse de la diversité $\beta$ des prélèvements environnementaux et des fèves de cacao en fermentation

L'analyse de la diversité  $\beta$  qui permet de comparer les échantillons entre eux a été réalisée. Pour cela les indices de diversité Bray-Curtis, Jaccard ont été appliqués.

L'indice de diversité  $\beta$ , Bray-Curtis permet la mesure de la dissimilarité entre deux échantillons. La matrice de dissimilarité obtenue permet de réaliser un positionnement multidimensionnel non-métrique afin d'obtenir le regroupement des échantillons (**Figure 46**).



**Figure 46 :** Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l'aide des indices de dissimilarité de Bray-Curtis pour les échantillons de fèves fraîche à différents temps de fermentation et des prélèvements environnementaux en provenance du Mexique.

A partir du positionnement multidimensionnel non métrique en **Figure 46**, nous pouvons voir que les communautés bactériennes issues des prélèvements environnementaux et des fèves fraîches de cacao en fermentation sont regroupées en 2 groupes. Il y a une séparation statistique entre les prélèvements réalisés au niveau de la plantation et ceux au niveau de la fermentation ainsi que des fèves. On peut voir également, que les prélèvements des sacs ayant servis au transport des cabosses de cacao, de la zone de plantation à la mise en fermentation, sont regroupés au centre, entre les prélèvements issus de la plantation et ceux issus de l'environnement de la fermentation.

La plupart des prélèvements ont été réalisés en répliques biologiques (3 pour la plupart des éléments en plantation) et sont regroupés ensemble lors du positionnement multidimensionnel. Cette séparation distincte en deux groupes nous indique que les communautés bactériennes issues des prélèvements au niveau de la plantation ne ressemblent pas aux communautés bactériennes issues des fèves fraîches en fermentation et des prélèvements issus de l'environnement de la fermentation.

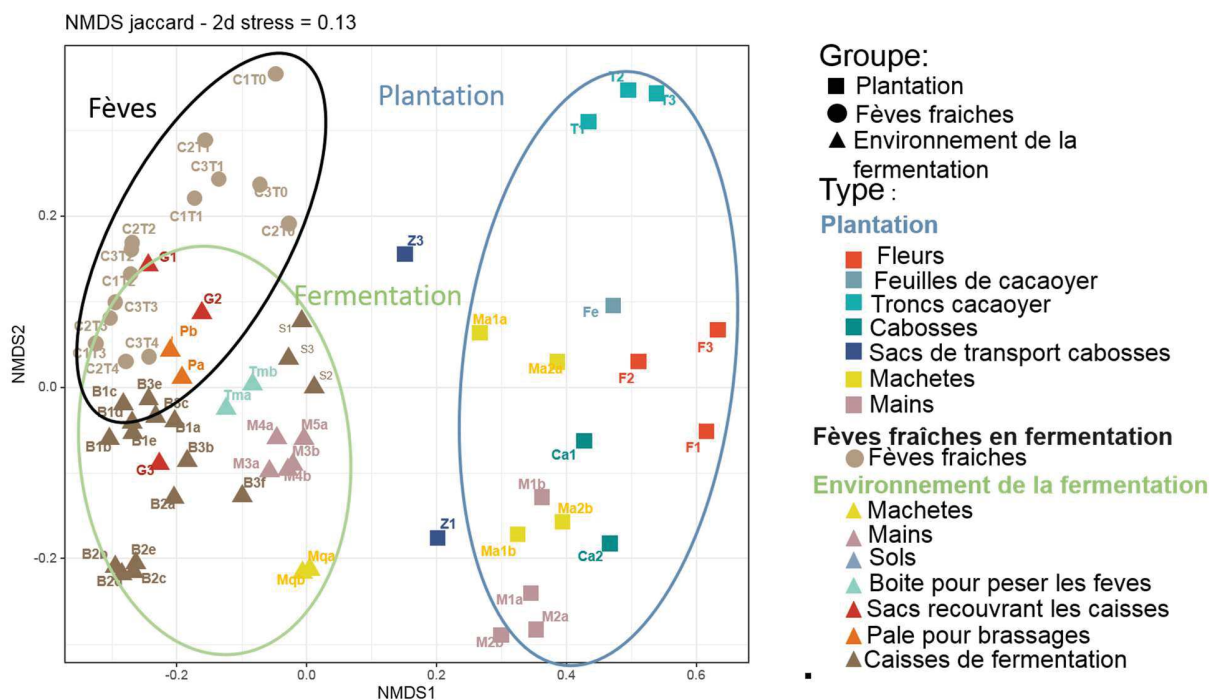
Au niveau du regroupement des fèves fraîches en fermentation et des prélèvements issus de l'environnement de la fermentation, on peut voir que les communautés bactériennes retrouvée sur les sacs recouvrant les fèves lors de la fermentation et la pale utilisée pour effectuer le brassage des fèves au cours de la fermentation se regroupent avec les communautés des fèves fraîches en fermentation, indiquant une forte similarité.

Un test statistique a été réalisé afin de confirmer cette séparation en deux des communautés bactériennes et d'identifier si l'origine du prélèvement, le groupe (plantation, fèves et environnement de la fermentation) ainsi que le type du prélèvement (feuilles, caisses, machette, etc.) ont un effet sur le regroupement des prélèvements lors du NDMS avec les indices de Bray-Curtis. Le **Tableau 22** représente les p-value obtenues lors du test statistique et le  $R^2$ , pour chacune des variables testées. Le type et le groupe ont un effet sur le regroupement avec des p-value = 0.001\*\*\*. Le type d'échantillon, explique à 32,7 % la répartition des échantillons et le groupe (fèves fraîches en fermentation, plantation ou environnement de la fermentation) à 21,4 %.

**Tableau 22 :** Résultats du test statistique PERMANOVA sur la matrice de distance réalisées avec les indices de Bray-Curtis pour les échantillons de fèves fraîches à différents temps de fermentation et des prélèvements environnementaux en provenance du Mexique

Variable	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr..F.
Groupe	2.000	5.492	2.746	12.157	0.214	0.001
Type	11.000	8.401	0.764	3.381	0.327	0.001
Interaction des deux	1.000	0.705	0.705	3.123	0.027	0.001
	49.000	11.069	0.226		0.431	
	63.000	25.669			1.000	

Afin de tester, si ces regroupements des communautés bactériens sont dus à l’abondance relative des séquences au sein des OTUs ou seulement à la diversité du nombre d’OTUs différents, l’indice de Jaccard a été employé. De la même manière qu’avec Bray-Curtis, une matrice de dissimilarité est obtenue puis une visualisation graphique est obtenue en créant une NMDS (**Figure 47**).



**Figure 47:** Positionnement multidimensionnel non-métrique à partir de la matrice de distance obtenue à l’aide des indices de similarité de Jaccard pour les échantillons de fèves fraîches à différents temps de fermentation et des prélèvements environnementaux en provenance du Mexique.

Avec cet indice, nous obtenons la même répartition des échantillons qu'avec l'indice de Bray-Curtis. Les tests statistiques, confirment également l'influence du groupe et du type du prélèvement sur la répartition des échantillons, avec des p-value = 0.001\*\*\*, **Tableau 23**. Le type d'échantillon, explique à 29,3 % la répartition des échantillons et le groupe (fèves fraîches en fermentation, plantation ou environnement de la fermentation) à 15,1 %, ces pourcentages plus faibles indiquent que les abondances relatives des OTUs apportent de l'information sur le regroupement.

**Tableau 23:** Résultats du test statistique PERMANOVA sur la matrice de distance réalisées avec les indices de Jaccard pour les échantillons de fèves fraîches à différents temps de fermentation et des prélèvements environnementaux en provenance du Mexique.

Variable	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr..F.
Groupe	2.000	4.197	2.098	7.029	0.151	0.001
Type	11.000	8.132	0.739	2.476	0.293	0.001
Interaction des deux	1.000	0.797	0.797	2.670	0.029	0.001
	49.000	14.629	0.299		0.527	
	63.000	27.755			1.000	

Le regroupement des échantillons est identique selon l'indice de Bray-Curtis et de Jaccard. Cela indique que la composition des communautés bactériennes en terme d'OTU différentes suffit à expliquer le regroupement des échantillons même si l'abondance relative des OTUs participant également à ce regroupement.

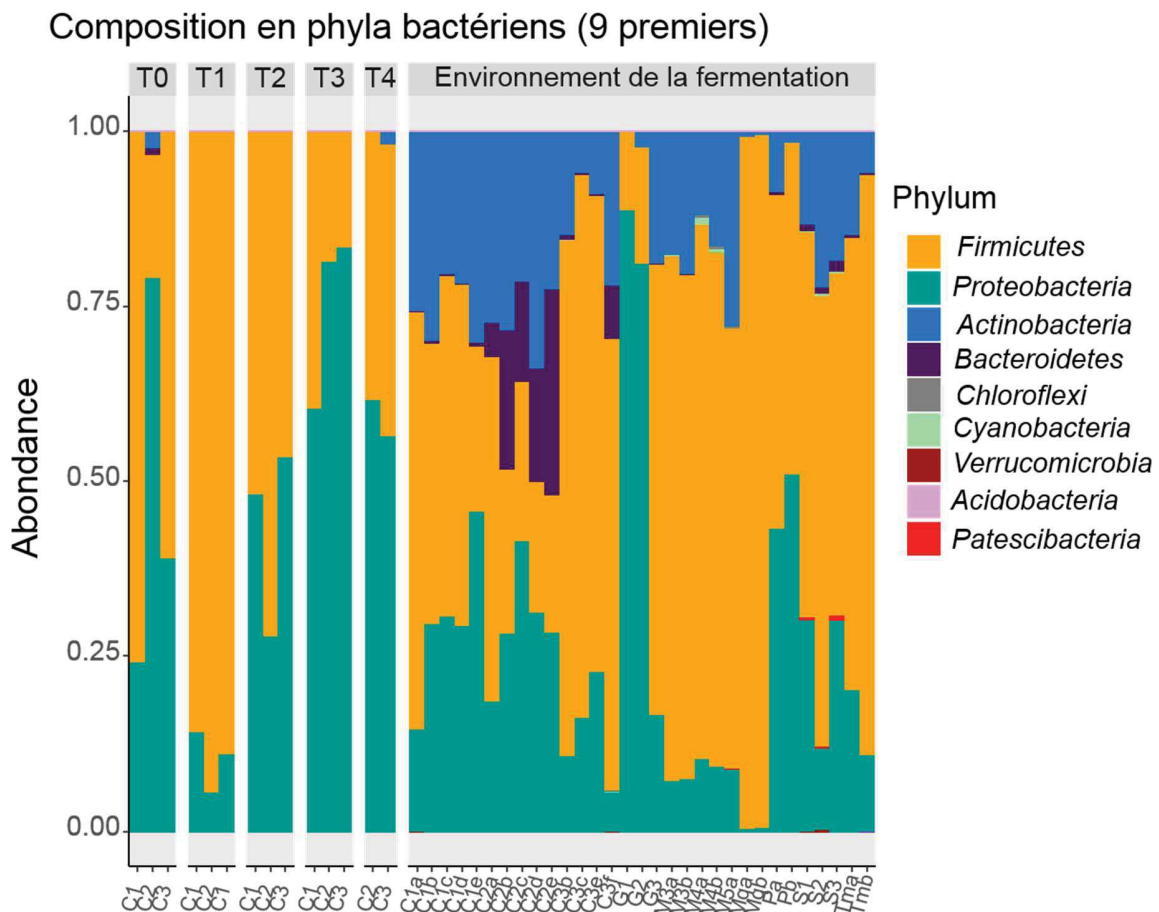
Le regroupement séparant en deux l'ensemble des échantillons selon l'environnement de la plantation et l'environnement de la fermentation, nous a permis d'exclure les éléments de l'environnement de la plantation comme étant à l'origine du pool de bactéries impliqué dans les fermentations de cacao. C'est pourquoi, par la suite de l'étude nous nous sommes intéressés à la composition des communautés bactériennes associées aux surfaces des éléments de la fermentation du cacao et des fèves de cacao en fermentation, afin d'identifier quelles surfaces sont à l'origine de la colonisation des fèves de cacao par les taxons impliqués dans les fermentations.

### III.2.3 Analyse de la composition des communautés bactériennes des fèves de cacao en fermentation et des surfaces de l'environnement de la fermentation

Les séquences des OTUs obtenues par NGS ont été comparées dans la base de données Silva 123 afin d'affilier la taxonomie. L'affiliation au niveau de l'espèce n'a pas pu être obtenue pour l'ensemble des OTUs, mais a pu être obtenue au niveau du genre.

#### Composition en phyla bactériens

Les communautés bactériennes des fèves de cacao fraîches en fermentation sont composées principalement des phyla des *Firmicutes* et des *Proteobacteria*, comme démontré dans le chapitre précédent (**Figure 48**).



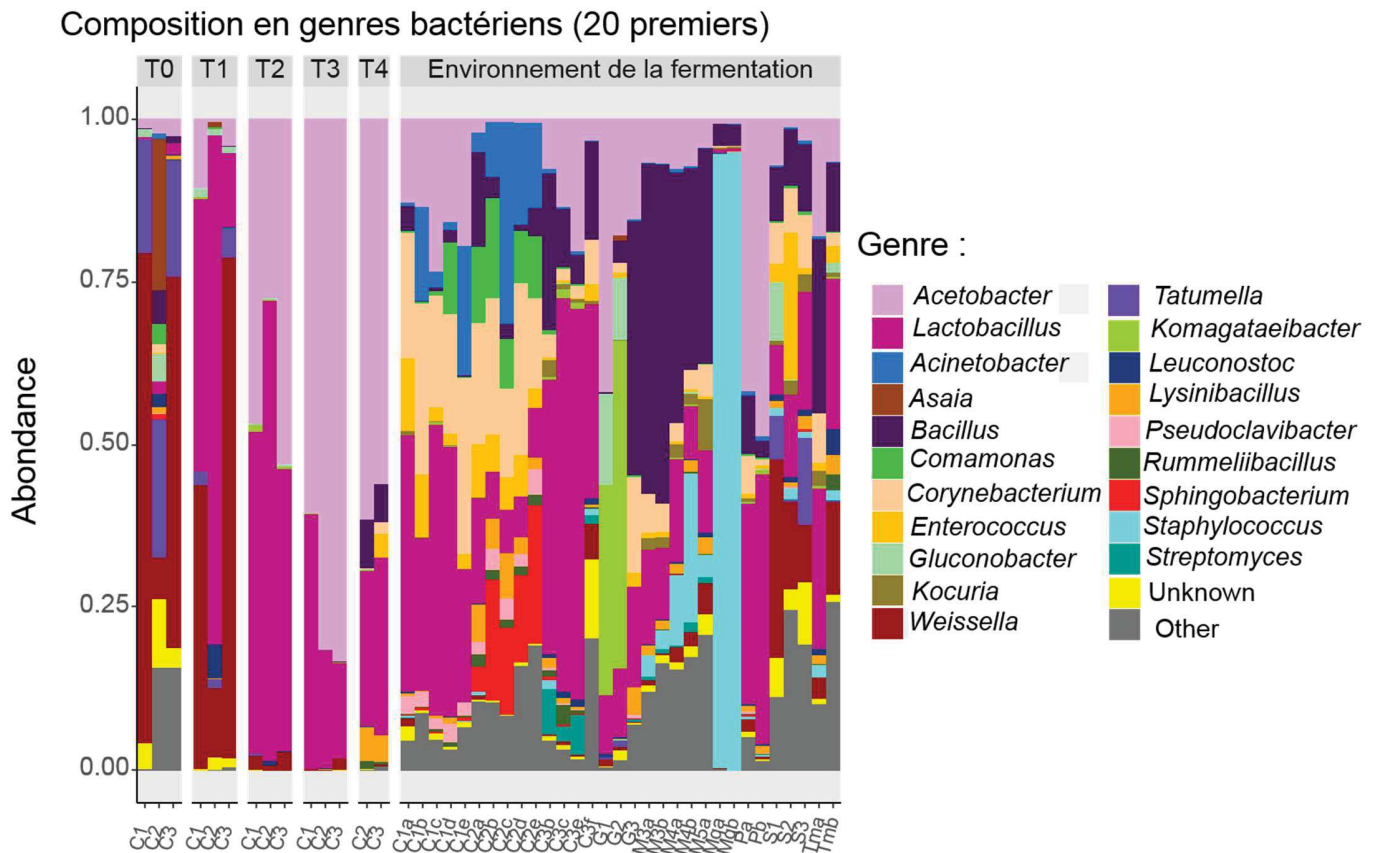
**Figure 48** : Composition des échantillons en phylum bactériens pour les prélèvements des fèves de cacao en fermentation et ceux issus de l'environnement de la fermentation.

Concernant les communautés bactériennes associées aux surfaces des éléments dans l'environnement de la fermentation, on peut voir une majorité d'OTUs affiliée aux phyla *Firmicutes*, *Proteobacteria* et *Actinobacteria*; cependant selon la surface la proportion de ces phyla est différents et d'autres phylums sont présents. Pour les caisses échantillonnées avant la mise des fèves en fermentation, une majorité d'OTUs ont été affiliées aux phyla *Firmicutes*, *Proteobacteria* et *Actinobacteria* (entre 15 et 55 %), mais également dans la caisse 2 une importante proportion (10 et 25 %) d'OTUs ont été affiliés au phylum *Bacteroidetes*. Pour les mains d'un des écabosseurs (M4), le phylum des *Firmicutes* est représenté majoritairement (75 %), le phylum *Proteobacteria* est à moins de 10 % et une faible proportion (moins de 5 %) des phylum *Chloflexi* et *Cynobacteria* sont identifiés. La machette utilisée pour écabosser les cabosses est composée principalement d'OTUs affiliées au phylum des *Firmicutes*. Le sol, en plus des phyla *Firmicutes*, *Proteobacteria* et *Actinobacteria* présents majoritairement ont également, une faible proportion d'OTUs affiliées au phyla *Patescibacteria*. La pale utilisée pour brasser les fèves est composée principalement de *Firmicutes* et *Proteobacteria* avec une faible proportion du phyla des *Actinobacteria* (moins de 10 %) comparée aux autres surfaces. La composition en genres bactériens est donnée dans la **Figure 49**.

#### Composition en genres bactériens

Lors du regroupement des échantillons à l'aide des indices de dissimilarité et de similarité de Bray-Curtis et Jaccard, il a été mis en évidence que les sacs recouvrant les fèves et les pales servant à réaliser le brassage ont été regroupés et sont proches des échantillons de fèves fermentées, indiquant que la composition de leurs communautés bactériennes est similaire (**Figure 49**).





**Figure 49 :** Composition des échantillons en genre bactériens pour les prélèvements des fèves de cacao en fermentation et ceux issus de l'environnement de la fermentation.

Nous avons démontré lors du chapitre précédant que les communautés bactériennes au 1<sup>er</sup> jour de fermentation sont plus diversifiées, puis au cours de la fermentation, les genres *Acetobacter* et *Lactobacillus* sont dominants jusqu'à la fin de la fermentation. A la fin de fermentation, une proportion importante d'OTUs affiliées au genre *Bacillus* est visualisée.

Ces 3 genres sont retrouvés dans les communautés bactériennes de la plupart des surfaces liées à l'environnement de la fermentation ainsi que les mains des écabosseurs. Excepté pour la machette, qui a une très faible proportion d'*Acetobacter* et aucune OTU affiliée au genre *Lactobacillus*. Ces résultats indiquent, que l'ensemble des surfaces de l'environnement de la fermentation permettent l'ensemencement des fèves de cacao avec les genres impliqués majoritairement dans les fermentations des fèves de cacao.

Concernant le début de fermentation, on peut voir que le genre *Weissella* est retrouvé dans les trois caisses de fermentation en de fortes proportions pour la caisse 1 et 3 (50 à 75 %) et en de plus faibles proportions dans la caisse 2 (moins de 10 %). Son abondance diminue au cours de la fermentation jusqu'à moins de 5 % car dominée par les genres *Acetobacter* et *Lactobacillus*. Dans l'environnement de la fermentation, les bactéries du genre *Weissella* sont des BAL fréquemment isolées lors de fermentation du cacao à travers différents pays (Camu *et al.* 2008; Papalexandratou *et al.* 2011c, 2013, 2019; Bortoloni *et al.* 2016; Agyirigo *et al.* 2019; Serra *et al.* 2019). Également, elles sont composées du genre *Tatumella* (20 à 25 %) ; Les bactéries du genre *Tatumella* sont importantes pour le déroulement de la fermentation car elles sont impliquées dans la dépectinisation de la pulpe en début de fermentation. Ces deux genres *Tatumella* et *Weissella* sont retrouvés dans l'environnement de la fermentation uniquement au niveau du sol et du seau pour peser les fèves, indiquant que ce sont ces surfaces qui ont permis la colonisation des fèves par ces genres. Ensuite, le genre *Gluconobacter* est également détecté dans les communautés des fèves en fermentation des caisses 1 et 2 ; Les bactéries de ce genre sont des BAA, fréquemment isolées lors de fermentations des fèves de cacao (Papalexandratou *et al.* 2011a, 2011b, 2013 ; Ho *et al.* 2014a, 2015; Agyirifo *et al.* 2019) et pouvant également être des pathogènes opportunistes pour l'Homme (Alauzet *et al.* 2010). Elles ont été identifiées également dans les communautés bactériennes de la surface des sacs recouvrant ces deux caisses (G1 et G2), du seau permettant de peser les fèves et de la pale servant aux brassages des fèves. Le genre *Acinetobacter* est retrouvé dans les communautés bactériennes des fèves en fermentation dans la caisse 2 à un moins de 5 % d'abondance et est décelé également dans la caisse 1 et 2 avant la mise en fermentation des fèves en de plus fortes proportions 10 à 25 %.

La caisse 2 présente une diversité plus riche en termes de genres différents retrouvés comme *Asaia*, *Gluconobacter*, *Leuconostoc*, *Enterococcus*, *Comamonas*, *Corynebacterium*, et *Bacillus*.

Le genre *Enterococcus* est retrouvé au niveau des communautés bactériennes des mains des ouvriers, *Corynebacterium* est retrouvé au niveau des mains mais également au niveau des caisses de fermentation et enfin le genre *Rummeliibacillus* au niveau des caisses, cela confirme notre hypothèse émise plus haut lors de l'analyse des communautés bactériennes des fèves en fermentation selon les 3 pays.

### III.2.4 Conclusion partielle

Les communautés bactériennes associées aux prélèvements réalisés sur les surfaces et les mains des travailleurs au niveau de la plantation et de l'environnement présentent une plus grande diversité spécifique que celles associées aux fèves de cacao en fermentation.

L'analyse comparative de leur composition en OTUs à l'aide des indices de diversité  $\beta$  nous indique que l'environnement de la plantation ne participe pas à la colonisation bactérienne des fèves. Ce sont les éléments de surface liés à l'environnement de la fermentation qui permettent donc l'inoculation des fèves. Les 3 genres retrouvés en majorité dans les communautés bactériennes des fèves de cacao (*Acetobacter*, *Lactobacillus* et *Bacillus*) sont également présents dans la plupart des communautés bactériennes des surfaces liées à l'environnement de la fermentation ainsi que sur les mains des écabosseurs. Ce résultat indique que chaque surface participe à l'inoculation des fèves. Le sol, le seau pour peser les fèves, la pale qui sert à brasser les fèves et la caisse de fermentation (particulièrement pour les caisses 1 et 3) sont les principales surfaces à l'origine des genres bactériens retrouvés en début de fermentation. L'analyse des pratiques artisanales montre qu'entre deux fermentations, les outils et surfaces utilisés ne sont pas nettoyés. Les microorganismes ont donc la capacité de survivre à la surface. Les surfaces des équipements agissent donc comme des réservoirs pour le transfert bactérien vers les fermentations.

### III.3– Diversité intraspécifique de souches de bactéries acétiques et lactiques en fonction de leurs origines géographiques

Dans les chapitres précédents, la composition et la structure des communautés bactériennes impliquées dans le processus de fermentation des fèves de cacao selon différentes zones géographiques ont été analysées. Les BAL et les BAA sont les deux grands groupes intervenants majoritairement. Selon, la zone géographique, les mêmes genres sont retrouvés en majorité : *Lactobacillus* et *Acetobacter*. De nombreuses études ont été réalisées dans le but de créer des starters de culture à partir des souches majoritaires isolées lors des fermentations. Ces starters de culture ont été réalisés avec des souches isolées d'un seul pays producteur sans comparer différentes souches isolées de différentes zones géographiques. Dans ce chapitre, nous nous sommes intéressés à la diversité intra spécifique de souches isolées de différentes zones géographiques afin de répondre aux questions de recherche suivantes :

- Existe-t-il des souches de BAA et de BAL spécifique à une origine géographique ?
- Est-ce que les souches d'une même espèce présentent les mêmes caractéristiques biochimiques lors de la fermentation des fèves de cacao ?

La réponse à ces questions pourrait nous permettre de proposer des pistes de souches candidates pour l'élaboration d'un starter de culture.

Nous avons également déterminé si certaines souches de BAA se distinguent des autres par leur comportement biochimique dans le but d'élaborer des starters de cultures pour la fermentation du cacao. Pour répondre à ces interrogations, nous avons dans un premier temps élaboré une collection de souches de BAA et BAL provenant de trois pays : le Mexique, la Côte d'Ivoire et la Guyane. La diversité intraspécifique des isolats des espèces majoritaires de BAA a été analysée selon leur polymorphisme génomique par « fingerprinting » et en étudiant le comportement de chaque souche dans un milieu simulant les conditions de la pulpe de cacao lors de la fermentation. Pour cette analyse, la vitesse spécifique de croissance des souches a été calculée, les consommations des principaux substrats (mannitol, éthanol et acide lactique) et la production du principal métabolite (acide acétique) par les BAA lors de la fermentation ont été analysées par HPLC.

### III.3.1 Elaboration d'une collection de souches de BAA et BAL

#### *III.3.1.1 Stratégies d'échantillonnage et de conservation des échantillons*

Pour l'obtention des échantillons de fèves de cacao, il n'a pas été possible de se rendre dans tous les pays producteurs pour réaliser nous-même l'échantillonnage. Pour la plupart des échantillons, des collaborateurs du Cirad ont réalisé les prélèvements et le transport des fèves jusqu'à Montpellier.

Les pratiques de fermentations et post-fermentations diffèrent selon les différents producteurs. Il n'a donc pas été possible de contrôler ou d'imposer une stratégie de fermentation pour obtenir une standardisation entre les différentes campagnes. Nous étions donc dépendants des pratiques employées dans chaque pays : (1) pour la quantité de fèves mise en fermentation, soit 20-25 kg (Guyane et Mexique) soit 800 kg (Côte d'Ivoire), (2) pour le mode de fermentation, réalisé en tas (Côte d'Ivoire 2017 et 2018) ou en caisses (Mexique et Guyane) et (3) pour les brassages des fèves, réalisées deux fois en Côte d'Ivoire, et toutes les 24h à partir du second jour de fermentation pour les autres campagnes.

En ce qui concerne le conditionnement, le transport et la conservation des fèves, des différences ont également été observées. En effet, pour les échantillons en provenance de Côte d'Ivoire (2017), Pérou et Nicaragua, les fèves ont été séchées puis transportées dans des sachets de manière stérile à température ambiante. Pour les campagnes d'échantillonnage, en Côte d'Ivoire (2018), en Guyane et au Mexique (Octobre 2018), les fèves fraîches ont été mises en contact avec un cryoprotectant (glycérol) puis congelées à -20 °C. Pour le transport jusqu'à Montpellier, les échantillons ont été conservés de manière réfrigérée dans des glacières puis placés à -20 °C. Pour la campagne au Mexique en février 2018, nous n'avons pas reçu les échantillons de fèves directement mais des bouillons de culture réalisés sur place et conservés dans du glycérol à -20°C.

Cependant, lors de ma thèse, j'ai pu effectuer un séjour dans un « beneficio » mexicain, une coopérative de plantations de cacao où les fèves sont fermentées, séchées puis stockées. J'ai pu réaliser les prélèvements et les isolements bactériens sur place sans délai, sans congélation des fèves. En ce qui concerne les souches de BAA isolées en Côte d'Ivoire 2016, la même stratégie a été employée, c'est-à-dire que les isolements ont été réalisés sur place par une doctorante de l'UMR QualiSud, Christelle Kouame.

### III.3.1.2 Isolement des BAA et des BAL

Les BAA sont reconnues pour être difficiles à isoler, cultiver et maintenir en culture pure (Teyssier et Hamdouche, 2016). Elles ont besoin d'une grande quantité d'oxygène pour réaliser l'oxydation de l'éthanol en acide acétique. En ce qui concerne la fermentation des fèves de cacao ce sont les brassages des fèves qui permettent cette entrée d'oxygène. En laboratoire, les cultures sont placées en aérobiose et l'agitation peut accélérer ce processus.

#### Isolement des BAL et BAA à partir des échantillons de fèves séchées

Dans un premier temps, nous disposons d'échantillons de fèves de cacao séchées à différents temps de fermentation. Les échantillons de fèves séchées réalisés en Côte d'Ivoire ont été mis en culture dans du bouillon nutritif et du bouillon MRS pour l'isolement respectif des BAA et BAL. Une faible turbidité dans les bouillons de culture a été observée. Ces précultures ont été étalés sur milieu *Acetobacter* et MRS pour l'isolement respectif des BAA et BAL après incubation, l'isolement des bactéries sur les milieux a été difficile, un faible nombre de colonies ont poussés pour l'ensemble des échantillons. Concernant l'isolement sur milieu *Acetobacter*, 20 bacilles à Gram négatif de forme ellipsoïdale ont été observées comme potentielles BAA. L'identification a pu être obtenue seulement pour 5 souches ; Elles appartiennent à la famille des *Enterobacteriaceae* (*Escherichia coli*, *Serratia marcescens*, *Enterobacter xiangfangensis*, *Lelliottia amnigena*). Ces espèces sont habituellement retrouvées dans de nombreux habitats (tube digestif des animaux, plantes, eau et le sol). Ces souches correspondent probablement à des contaminations provenant de l'environnement autour du processus de traitement post fermentation des fèves (mains des manipulateurs). Concernant l'isolement sur milieu MRS, 15 souches potentielles de BAL ont été isolées, mais l'identification de ces souches n'a pas été réalisée.

Cette difficulté à isoler des bactéries peut venir du fait que les fèves de cacao ont été séchées au soleil et que les microorganismes n'ont peut-être pas survécu à cette étape de séchage. Nous avons donc procédé à des mesures d'activité de l'eau ( $A_w$ ) dans les échantillons de fèves de cacao. Cette mesure permet d'évaluer la quantité d'eau libre dans les échantillons c'est-à-dire l'eau disponible pour la croissance des microorganismes. Cet indice est compris entre 0 et 1. Plus l' $A_w$

est élevée, plus il y a d'eau libre. Pour la croissance des microorganismes, il faut un indice supérieur à 0,7. Cet indice est la limite de croissance des moisissures même les plus résistantes. Afin que les bactéries puissent croître, il faut que l'activité de l'eau dans les fèves de cacao présente un indice supérieur à 0,9. Les Aw des échantillons de Côte d'Ivoire, du Pérou et du Nicaragua ont été mesurés et figure dans le **Tableau 24**. Tous les indices d'activités de l'eau sont inférieurs à 0,71 ce qui signifie que les microorganismes peuvent très difficilement se multiplier sur les échantillons de fèves de cacao séchées. Les faibles indices d'activité de l'eau dans les fèves expliquent la difficulté à isoler les bactéries à partir des fèves de cacao séchées. C'est pourquoi pour la réalisation de la collection de souches, nous avons décidé de travailler uniquement avec des échantillons frais de fèves de cacao avec la pulpe mucilagineuse encore présente autour des fèves.

**Tableau 24** : Mesure de l'activité de l'eau des fèves de cacao séchées en provenance de Côte d'Ivoire, du Pérou et du Nicaragua.

Ville	Temps	Site	aw	Ville	Temps	Site	aw	Ville	Echantillon	aw
Akoupé	96 h	Site 3	0,5827	Soubre	96 h	Site 3	0,5473	Nicaragua	N-1	0,5144
		Site 2	0,5504			Site 2	0,5443		N-2	0,519
		Site 1	0,5495			Site 1	0,4927		N-3	0,566
	Site 3	0,573	Site 3		0,4605	N-T	0,5313			
	72 h	Site 2	0,5517		72 h	Site 2	0,6278	Pérou	P-T	0,4589
		Site 1	0,6084			Site 1	0,5418		P-1	0,4687
		Site 3	0,5737			Site 3	0,5622		P-2	0,4383
	48 h	Site 2	0,5228		48 h	Site 2	0,6679		P-3	0,4853
		Site 1	0,5792			Site 1	0,7015	P-4	0,5225	
Site 3		0,5663	Site 3	0,5466						
Diekanou	96 h	Site 2	0,5216	Daloa	96 h	Site 2	0,5919			
		Site 1	0,4911			Site 1	0,4737			
		Site 3	0,5591			Site 3	0,5383			
	72 h	Site 2	0,4959		72 h	Site 2	0,495			
		Site 1	0,5022			Site 1	0,5109			
		Site 3	0,6494			Site 3	0,5313			
	48 h	Site 2	0,5607		48 h	Site 2	0,4994			
		Site 1	0,616			Site 1	0,5043			

### Isolement des BAL et BAA à partir des échantillons de fèves fraîches

A partir des échantillons de fèves fraîches, nous avons travaillé avec deux milieux de culture liquide (CPSM-AAB et CPSM-LAB) puis avec deux milieux sélectifs (Kneifel et MRS) pour isoler respectivement les BAA et les BAL. Ces milieux et leur utilisation sont décrits dans la partie matériels et méthodes (II.4.1). Ils ont permis d'isoler 149 souches potentielles de BAA, c'est à dire des bacilles à Gram négatif de forme ellipsoïdale et 63 souches potentielles de BAL, c'est-à-dire des bacilles à Gram positif, catalase négative. Pour les BAL, une pré-identification basée sur l'amplification d'une portion du gène de l'ARN 16S présente uniquement chez les genres *Pediococcus*, *Weissella*, *Leuconostoc* et *Lactobacillus* a confirmé leur appartenance aux BAL.

#### III.3.1.3 Identification des BAA et BAL

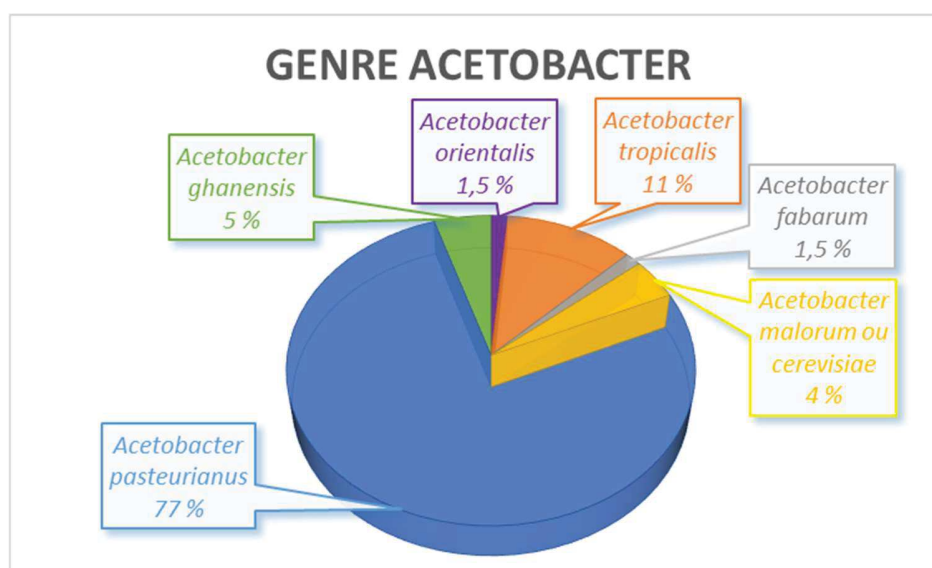
L'identification au niveau de l'espèce a été obtenue à l'aide de l'amplification et du séquençage du gène de l'ARNr 16S (1500 pb).

#### Identification des BAA

Le séquençage du gène de l'ARNr 16S a été réalisé sur les 149 souches potentielles de BAA et nous a permis d'obtenir des séquences allant de 1000 à 1404 pb (**Annexe 1**). Un fort pourcentage d'identité (id % > 99) a été obtenu avec les séquences présentes dans Genbank. Elles ont été identifiées comme des BAA. Ces résultats montrent que la stratégie employée au cours de cette étude est adaptée pour l'isolement des BAA.

Les souches de BAA ont été affiliées aux genres *Acetobacter* (n = 141) et *Gluconobacter* (n = 8). Parmi les souches d'*Acetobacter*, *A. pasteurianus* est l'espèce la plus fréquemment isolée avec 77 % des isolats, suivi de *A. tropicalis*, *A. ghanensis*, *A. fabarum* et *A. orientalis*, qui représentent respectivement 11 %, 5 %, 1,5 % et 1,5 % du total des isolats d'*Acetobacter* (**Figure 50**). Six souches d'*Acetobacter* (4 %) n'ont pas pu être affiliées au niveau de l'espèce. Effectivement, le gène de l'ARNr 16S d'*A. cerevisiae* et d'*A. malorum* sont très proches génétiquement et la taille de la séquence de l'ARNr 16S est insuffisamment informative pour les différencier (Valera *et al.* 2013). Pour les souches appartenant au genre *Gluconobacter*, 7 souches ont été identifiées comme *Gluconobacter oxydans* et une comme *Gluconobacter nephelii*.





**Figure 50** : Graphique résumant la distribution des isolats au sein des espèces appartenant au genre *Acetobacter*.

**Tableau 25**: Répartition des 149 isolats de BAA selon les trois pays (Côte d'Ivoire, Guyane et Mexique).

Identification des espèces bactériennes	Pays Total	Côte d'Ivoire		Guyane	Mexique	
		2016	2018		Grande récolte	Petite récolte
<i>A. pasteurianus</i>	108	35	15	12	15	31
<i>A. tropicalis</i>	16	5	2	2	5	2
<i>A. fabarum</i>	2	1	1			
<i>A. ghanensis</i>	7	5	2			
<i>A. orientalis</i>	2			2		
<i>A. malorum / A. cerevisiae</i>	6					6
<i>Gluconobacter sp</i>	8	2		6		
<b>Sous total</b>		<b>48</b>	<b>21</b>		<b>19</b>	<b>39</b>
<b>Total</b>	<b>149</b>	<b>68</b>		<b>22</b>	<b>59</b>	

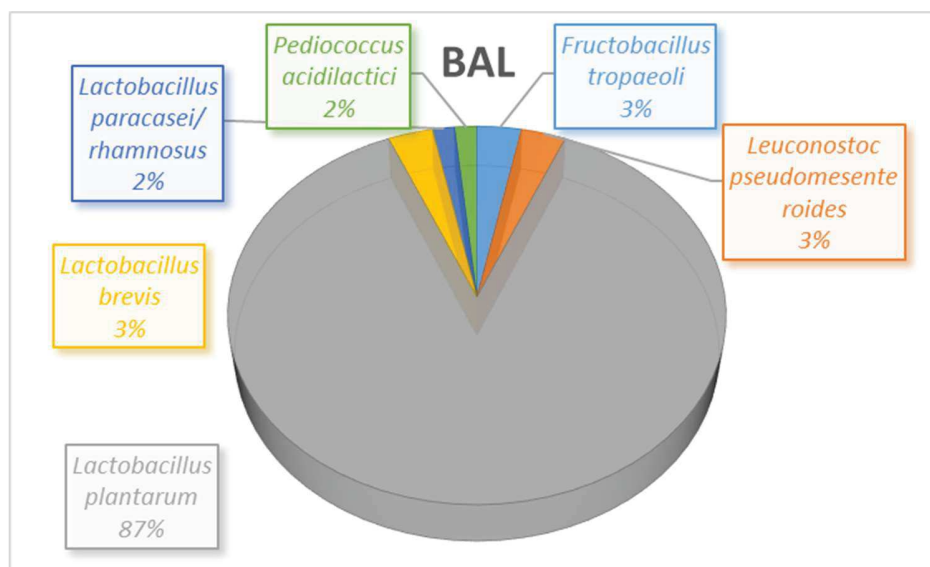
Le **Tableau 25** présente la répartition des souches de BAA selon les pays d'origine. Parmi les 149 isolats de BAA, 68 souches ont été isolées à partir de fermentations en Côte d'Ivoire, 59 souches de fermentations au Mexique et 22 souches de Guyane. Globalement, les souches d'*A. pasteurianus* et d'*A. tropicalis* ont été retrouvées dans les fermentations des 3 pays. En revanche, les souches d'*A. fabarum* et *A. ghanensis* n'ont été isolées qu'en Côte d'Ivoire ; *A. orientalis* n'a été isolée qu'en Guyane, *A. cerevisiae / malorum* qu'au Mexique. Les souches de *Gluconobacter sp* n'ont été isolées qu'en Côte d'Ivoire et qu'en Guyane.

Les espèces d'*Acetobacter* ont été fréquemment isolées lors de fermentations du cacao à travers les différents pays producteurs. L'inventaire de la diversité des BAA réalisé au cours de la thèse et résumé dans le **tableau 3 (Etat de l'art - I.6.2.3)** montre que l'espèce *A. pasteurianus* a été l'espèce majoritairement isolée des fermentations de cacao au Brésil, Cameroun, Côte d'Ivoire, Equateur, Ghana, Malaisie, République Dominicaine, Australie, Bolivie et au Honduras (Ardhana, 2003; Camu *et al.* 2007, 2008; Papalexandratou *et al.* 2011a,2011b,2011c; Ho *et al.* 2014a, 2015, ; Magalhães da Veiga Moreira *et al.* 2016; Miescher Schwenninger *et al.* 2016; Mota-Gutierrez *et al.* 2018; Romanens *et al.* 2018) suivie de *A. ghanensis*, *A. fabarum*, *A. tropicalis*. Les espèces *A. cerevisiae* ou *A. malorum* ont été peu fréquemment isolées lors des fermentations de cacao dans des pays où la production de cacao est faible notamment au Honduras (Romanens *et al.* 2018), Brésil (de Melo Pereira *et al.* 2013) et en Equateur (Papalexandratou *et al.* 2011c).

#### Identification des BAL

Le séquençage du gène de l'ARNr 16S a été réalisé sur les 63 souches potentielles de BAL et nous a permis d'obtenir des séquences allant de 725 à 1414 pb (**Annexe 2**). Un fort pourcentage d'identité (id % > 97) a été obtenu avec les séquences présentes dans la Genbank. Elles ont été identifiées comme des BAL.

*L. plantarum* est l'espèce la plus fréquemment isolée avec 87 % des isolats, suivie de *Leuconostoc pseudomesenteroides*, *Lactobacillus brevis*, *Fructobacillus tropaeoli*, *Pediococcus acidilactici* et *Lactobacillus paracasei/ rhamnosus* qui représentent respectivement 3 %, 3 %, 3 %, 2 % et 2 % du total des isolats de BAL (**Figure 51**).



**Figure 51:** Graphique résumant la distribution des isolats au sein des différents genres et espèces de BAL.

**Tableau 26 :** Répartition des 63 isolats de BAL selon les trois pays (Côte d'Ivoire, Guyane française et Mexique).

Identification des espèces bactériennes	Pays	Côte d'Ivoire		Guyane		Mexique	
		Total	2018	2017	Grande récolte	Petite récolte	
<i>L. plantarum</i>	55	18	3	8	26		
<i>Leuconostoc pseudomesenteroides</i>	2		2				
<i>L. brevis</i>	2			2			
<i>Fructobacillus tropaeoli</i>	2			2			
<i>Pediococcus acidilactici</i>	1	1					
<i>L. paracasei/rhamnosus</i>	1			1			
<b>Sous total</b>				<b>11</b>	<b>26</b>		
<b>Total</b>	<b>63</b>	<b>19</b>	<b>7</b>	<b>37</b>			

Le **Tableau 26** présente la répartition des souches de BAL selon les pays d'origine. Parmi les 63 isolats de BAL, 19 souches ont été isolées à partir de fermentations de Côte d'Ivoire, 37 souches

à partir de fermentations au Mexique et 7 souches à partir de fermentations en Guyane. Seules, les souches de *L. plantarum* ont été retrouvées dans les fermentations des 3 pays. Cependant, les souches de *Leuconostoc pseudomesenteroides* et *Fructobacillus tropaeoli* n'ont été isolées qu'en Guyane; *L. brevis* et *L. paracasei/rhamnosus* qu'au Mexique et *Pediococcus acidilactici* qu'en Côte d'Ivoire.

Ces espèces ont été déjà identifiées comme intervenants dans la fermentation des fèves de cacao, d'après l'inventaire de la diversité des BAL (**Tableau 4- I.4.3.3**). *L. plantarum* est l'une des espèces majoritairement identifiées comme dominant les fermentations au Brésil, Cameroun, Côte d'Ivoire, Equateur, Ghana, Malaisie, République Dominicaine, Australie, Bolivie et au Honduras (Ardhana, 2003; Camu *et al.* 2007, 2008; Papalexandratou *et al.* 2011a,2011b,2011c; Ho *et al.* 2014b, 2015, ; Magalhães da Veiga Moreira *et al.* 2016; Miescher Schwenninger *et al.* 2016; Mota-Gutierrez *et al.* 2018; Romanens *et al.* 2018).

#### III.3.1.4 Conclusion partielle

Cette partie a mis en évidence l'importance du conditionnement des échantillons de fèves de cacao pour l'isolement des BAA et BAL. Il a été nécessaire d'avoir des échantillons frais de fèves de cacao avec la pulpe encore présente autour afin d'optimiser l'isolement. La collection de souches a été réalisée à partir d'échantillons frais de fèves de cacao en provenance de trois zones géographique le Mexique, La Côte d'Ivoire et la Guyane.

Notre stratégie d'isolement des souches s'est avérée efficace, car toutes les souches isolées à l'aide des milieux spécifiques ont été identifiées comme des souches de BAA et BAL. Cependant, la stratégie employée pour l'isolement des BAL comporte plusieurs biais pour l'isolement des espèces minoritaires. Seuls les bacilles à Gram positif ont été sélectionnés et les cocci à Gram positif n'ont pas été pris en compte pour la réalisation de la collection de souches de BAL. L'utilisation des amorces spécifiques à 4 genres (*Weissella*, *Leuconostoc*, *Pediococcus* et *Lactobacillus* de BAL pour la pré-identification ne permet pas d'accéder à la totalité des genres de BAL.

*A. pasteurianus* et *L. plantarum* sont les espèces majoritairement isolées des fermentations de fèves de cacao à travers les 3 zones géographiques, avec 107 isolats soit 73 % du total de souches isolées pour *A. pasteurianus* et 55 isolats soit 87 % du total des souches de BAL isolées pour *L. plantarum*. Ce résultat est en adéquation avec les données de la littérature, *A. pasteurianus* et *L. plantarum* étant les espèces majoritaires, dominant les fermentations et formant le consortium microbien. L'espèce *A. tropicalis* a aussi été retrouvée dans les 3 zones mais à un plus faible pourcentage (11 %) soit 16 souches.

Pour la suite de l'étude, nous avons décidé de nous intéresser à l'analyse de la diversité intraspécifique des souches d'*A. pasteurianus* et d'*A. tropicalis* qui sont les espèces retrouvées majoritairement dans les fermentations de fèves de cacao des trois pays échantillonnés.

### III.3.2 Analyse de la diversité intraspécifique des souches d'*A. pasteurianus*

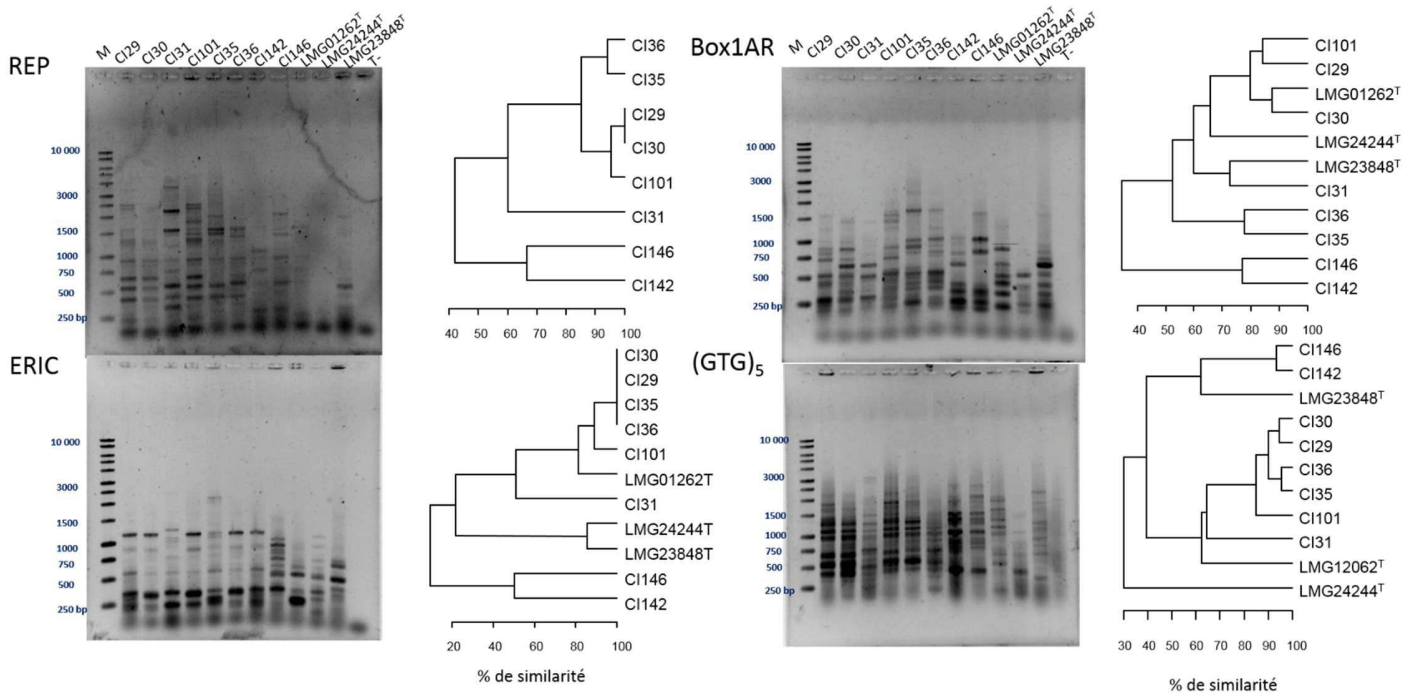
#### III.3.2.1 Analyse du polymorphisme génomique

##### III.3.2.1.1 Mise au point de la méthode

L'analyse de la diversité intraspécifique des souches de d'*A. pasteurianus* a été réalisée en étudiant leur polymorphisme génomique à l'aide de l'amplification de portions d'ADN situées entre des séquences répétées largement distribuées dans le génome des bactéries. Nous disposons de 4 jeux d'amorces différents qui ciblent respectivement les séquences Box, REP, ERIC et (GTG)<sub>5</sub>.

#### a) Choix des amorces des séquences cibles

Dans un premier temps chaque jeu d'amorce (BoxA1R, REP, ERIC et (GTG)<sub>5</sub>) a été testé sur 6 souches d'*A. pasteurianus* (CI29, CI30, CI31, CI35, CI36 et CI101), 2 souches d'*A. ghanensis* (CI142 et CI146) et trois souches types (*A. pasteurianus* LMG1262<sup>T</sup>, *A. ghanensis* LMG23848<sup>T</sup> et *A. fabarum* LMG24244<sup>T</sup>). Le traitement d'image des gels et l'analyse statistique nous permet d'obtenir des dendrogrammes (**Figures 52 et 53**).

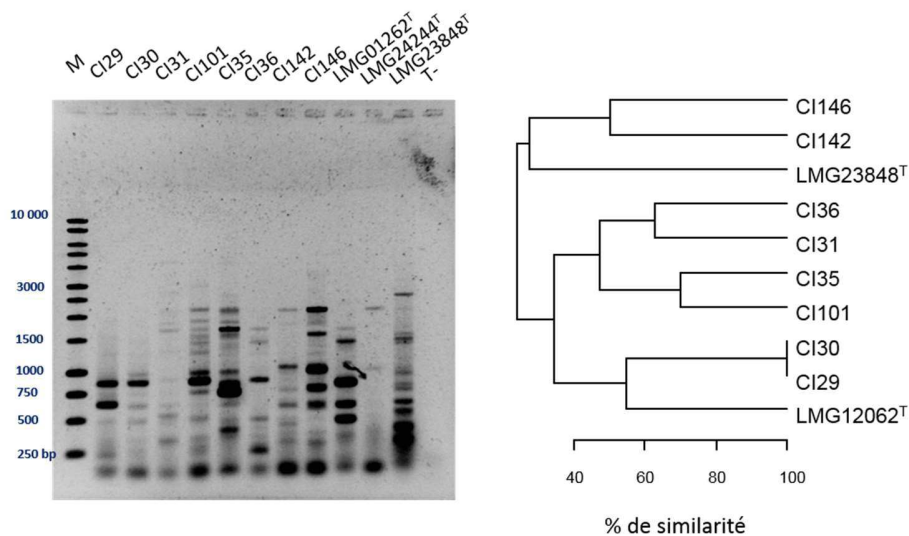


**Figure 52** : Gels d'électrophorèse et regroupement des souches en dendrogramme à partir des profils génomiques des souches de BAA après amplification par PCR à l'aide des amorces de séquences répétées (REP, Box1AR, ERIC et (GTG)<sub>5</sub>). L'indice de Sørensen-Dice a été utilisé pour analyser la similarité et un algorithme de regroupement UGPMa a été appliqué pour regrouper les souches. M = marqueur de poids moléculaire 1 kb. T- = témoin négatif.

Le dendrogramme obtenu à partir des profils génomiques réalisés à l'aide des amorces ERIC montre qu'il existe un pourcentage élevé de similarité entre 4 souches d'*A. pasteurianus* (CI30, CI29, CI35 et CI36) qui ont exactement le même profil génomique (**Figure 52**). C'est pourquoi les amorces ERIC ne sont pas suffisamment discriminantes pour l'analyse de la diversité intraspécifique des *A. pasteurianus*. Le dendrogramme obtenu à partir des profils génomiques réalisées à l'aide de l'amorce Box1AR, montre que la souche d'*A. fabarum* est regroupée avec des souches d'*A. pasteurianus* et la souche d'*A. ghanensis* est regroupée avec une souche d'*A. pasteurianus* (CI31) (**Figure 52**). Cette amorce n'est également pas suffisamment discriminante puisqu'elle ne permet pas de discriminer les espèces au sein du même genre.

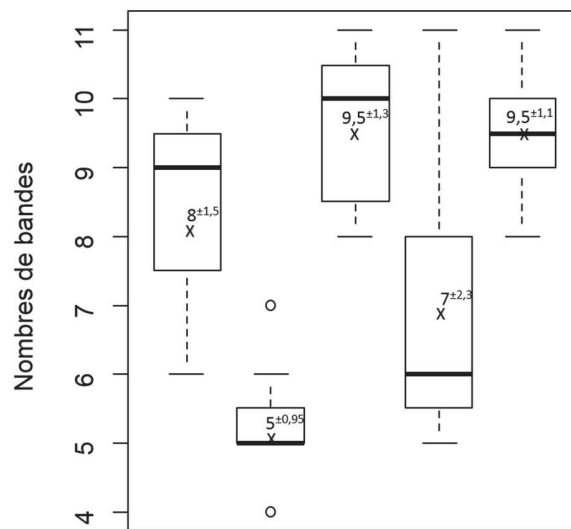
Le dendrogramme obtenu à partir des profils génomiques réalisés à l'aide des amorces REP (**Figure 52**), a permis de regrouper les souches d'*A. pasteurianus*, d'*A. ghanensis* en deux groupes distincts. Les profils complets n'ont pas pu être obtenus pour les 3 souches types, elles n'ont donc

pas été introduites dans l'analyse. Enfin, le dendrogramme obtenu à partir des profils génomiques réalisés à l'aide de l'amorce (GTG)<sub>5</sub> montre que les souches d'*A. pasteurianus* sont regroupées dans un seul et même groupe partageant 60 % de similarité, les souches d'*A. ghanensis* également et la souche d'*A. fabarum* qui a 30 % de similarité avec les deux grands groupes d'*A. pasteurianus* et d'*A. ghanensis*. Ces deux jeux d'amorces (REP et (GTG)<sub>5</sub>) sont discriminants pour l'analyse de la diversité interspécifique et intraspécifique.



**Figure 53:** Gel d'électrophorèse et regroupement des souches en dendrogramme sur les profils génomiques des souches de BAA après amplification par PCR par la méthode de multiplexe (amorces : Box1AR, REP et ERIC). L'indice de Sørensen-Dice a été utilisé pour analyser la similarité et un algorithme de regroupement UGPMA a été appliqué pour regrouper les souches. M = marqueur de poids moléculaire 1 kb. T- = témoin négatif.

Le dendrogramme obtenu à partir de la PCR multiplexe (**Figure 53**) montre 3 grands groupes : deux composés de souches d'*A. pasteurianus* avec moins de 40 % de similarité entre les deux. Le troisième groupe est composé de souches d'*A. ghanensis* et de la souche type *A. fabarum* avec moins de 40 % de similarité avec les deux groupes d'*A. pasteurianus*. Des dimères d'amorces ont été détectés en bas du gel pour chaque souche, indiquant qu'une grande quantité des amorces ne se sont pas hybridées avec l'ADN génomique des souches.



**Figure 54** : Boîtes à moustache des nombres de bandes des profils génomiques selon les différents amorces (Box1AR, ERIC, (GTG)<sub>5</sub>, Multiplexe et REP).

La moyenne est indiquée en plus de la médiane.

L'analyse comparative du nombre de bandes a été réalisée sur les 6 souches d'*A. pasteurianus* et la souche type LMG12062<sup>T</sup> d'*A. pasteurianus* (profils complets avec tous les jeux d'amorces). La **Figure 54** représente le nombre de bandes médian et moyen obtenu avec chaque jeu d'amorces. Selon le jeu d'amorce utilisé les profils génomiques contiennent un nombre variable de bandes. Le test statistique d'ANOVA a confirmé la significativité de ces résultats avec une p-value < 0.05, après vérification de la normalité des données (Shapiro-Wilk test = p value > 0.05). Cependant, la comparaison deux à deux, a montré que la moyenne du nombre de bandes entre REP et (GTG)<sub>5</sub> n'est significativement pas différente (p-value > 0.05). Ces deux amorces permettent d'obtenir un profil génomique avec un nombre de bandes plus important (9,5 en moyenne) que les autres.



D’après l’analyse des profils génomiques selon leur nombre de bandes et leur regroupement en dendrogrammes, l’utilisation des amorces (GTG)<sub>5</sub> semble être adéquate pour notre étude.

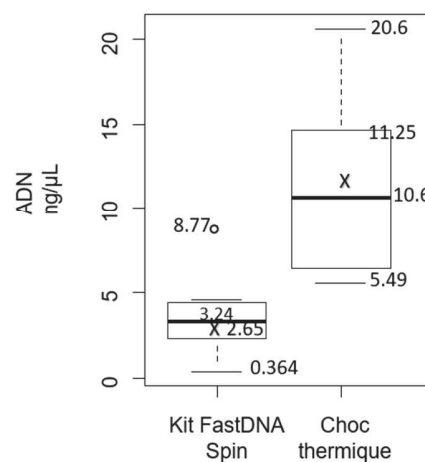
b) Mise au point de la (GTG)<sub>5</sub> rep-PCR

Une mise au point a été effectuée afin de s’assurer de l’efficacité de l’extraction et de l’amplification d’ADN. Pour cela deux méthodes d’extraction d’ADN ont été comparées : méthode d’extraction rapide par choc thermique et extraction à l’aide du kit FastDNA Spin (**Figure 55, Tableau 27**).

L’analyse à l’aide du fluoromètre (Qubit) nous indique que les quantités d’ADN extraites avec la méthode par choc thermique sont plus élevées qu’avec le kit (**Tableau 27**), avec une moyenne de 10.6 ng/μL et 2.65 ng/μL pour la méthode rapide par choc thermique et par le Kit FastDNA Spin respectivement (Figure 6A). Le test statistique nous a permis de confirmer la significativité de ces résultats (p-value < 0.05), après vérification de la normalité des données (Shapiro-Wilk test = p value < 0.05)

**Tableau 27** : Comparaison de la quantité d'ADN extraite de 6 souches d'*A. pasteurianus* et la souche type LMG12062<sup>T</sup> d'*A. pasteurianus* par les deux méthodes d'extraction : par choc thermique et avec le kit FastDNA spin

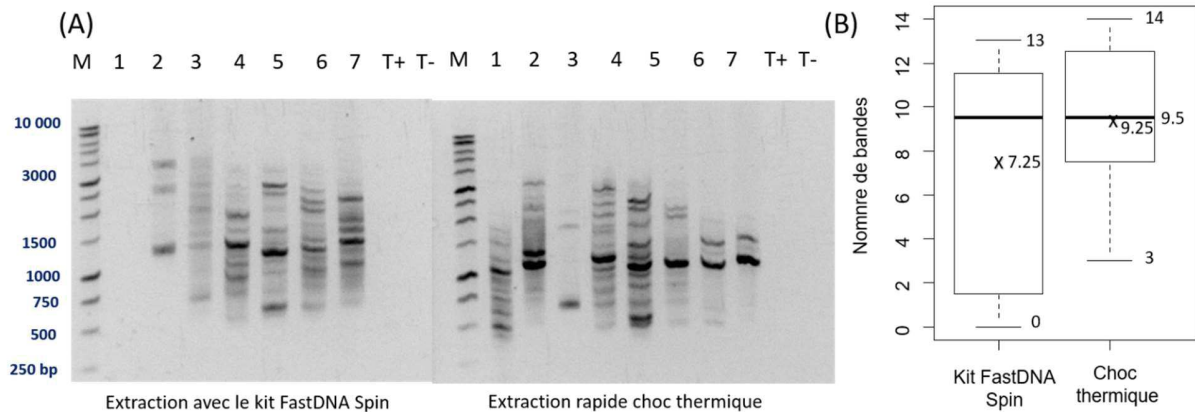
Souches	Méthode d'extraction	
	Kit FastDNA Spin (ng/μL)	Rapide par choc thermique (ng/μL)
CI32 (1)	8.77	20.6
CI1 (2)	1.97	5.49
CI98 (3)	0.364	18.4
CI33 (4)	4.56	10.8
CI37 (5)	2.61	6.37
CI38 (6)	3.24	10.6
LMG 01262 <sup>T</sup> (7)	4.22	6.49



**Figure 55:** Boîtes à moustaches des quantités d’ADN des extractions réalisées par les deux méthodes (choc thermique et le kit Fast DNA Spin) des 7 souches d'*A. pasteurianus*.

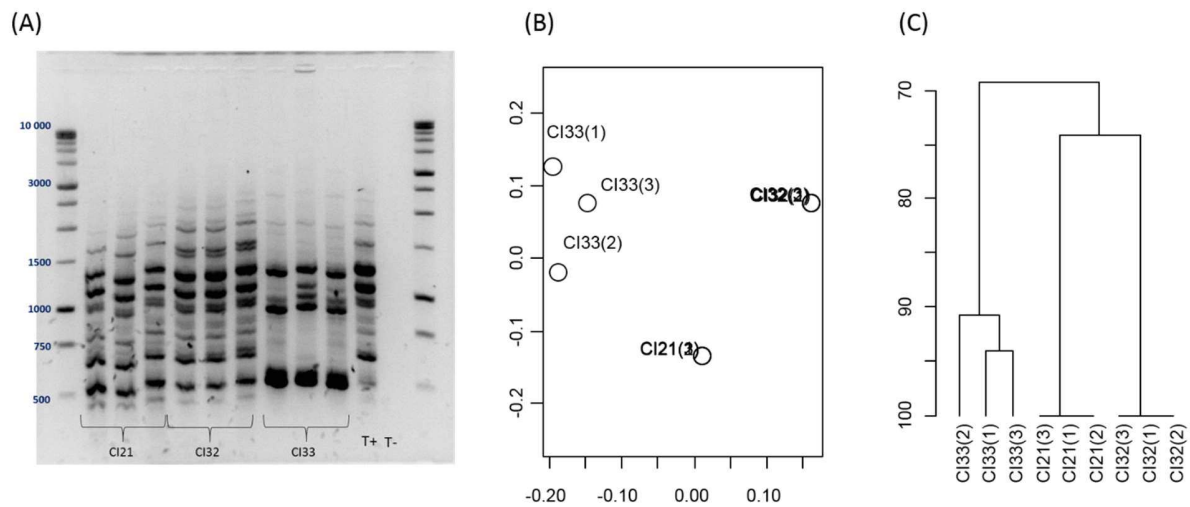
La **Figure 56A** correspond aux profils génomiques (GTG)<sub>5</sub> obtenus à partir des extraits d’ADN soit par choc thermique soit avec le Kit FastDNA. Le nombre médian de bandes obtenues pour les

profils génomiques est similaire entre les deux méthodes d'extraction (**Figure 56B**), le test de Student a confirmé la significativité de ce résultat avec une p-value < 0.05.



**Figure 56** : (A) Profils génomiques obtenus à l'aide de l'amorce (GTG)<sub>5</sub> et (B) Boîtes à moustache des nombres de bandes obtenus sur les profils génomique des souches d'*A. pasteurianus*. Durée de migration 85 min

Afin de valider la méthode extraction d'ADN par choc thermique, nous avons voulu vérifier la reproductibilité des profils de bandes selon des extractions réalisées en triplicata sur 3 souches d'*A. pasteurianus* (CI21, CI32 et CI33) (**Figure 57A**). Le temps de migration a également été mis au point, afin d'avoir une meilleure séparation des bandes. Les amplicons ont migré pendant 2h au lieu de 85 min. Visuellement, les profils génomiques obtenus pour les 3 extractions comportent le même nombre de bandes (12 pour CI21 et 16 pour CI32) sauf pour CI33, le premier réplica n'a pas la bande à 1100 pb et le second celle à 2750 pb.



**Figure 57** : (A) Gels d'électrophorèse des profils  $(GTG)_5$  (B) Positionnement multidimensionnel non-métrique (NMDS) et (C) dendrogramme pour le regroupement des souches selon leur profil génomique. Les extraits d'ADN ont été réalisés par choc thermique.

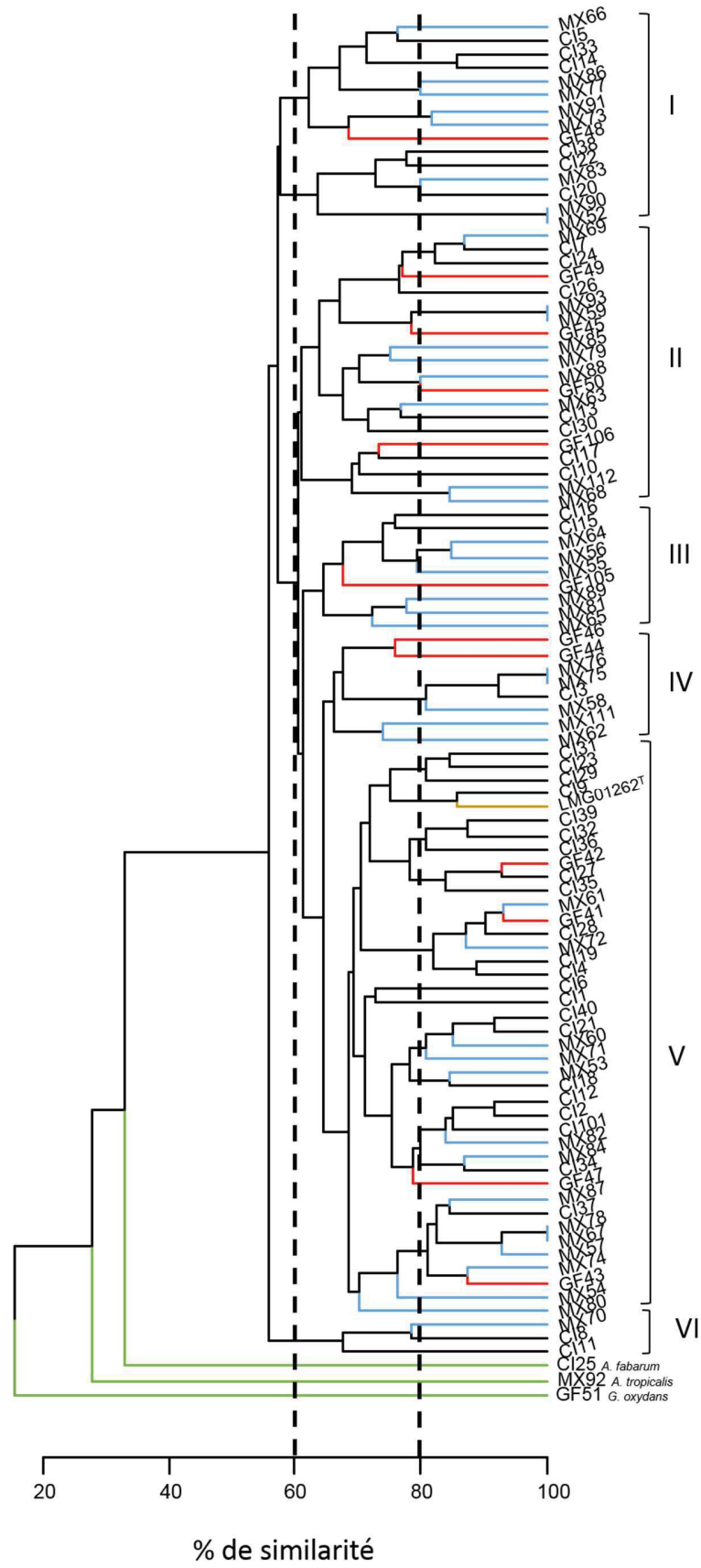
L'indice de Sørensen-Dice a été utilisé pour analyser la similarité, un positionnement multidimensionnel non-métrique (NMDS) et un algorithme de regroupement UGPA a été appliqué pour regrouper les souches.

L'indice de similarité de Sørensen-Dice a été calculé sur la matrice créée à partir de l'absence (0) ou la présence (1) des bandes. A partir de la matrice de distance ainsi obtenue avec les indices de similarité, un positionnement multidimensionnel (**Figure 57B**) et un dendrogramme (**Figure 57C**) ont été réalisés pour le regroupement des souches. Statistiquement, les réplicas des souches sont regroupés ensemble, avec des différences pour la souche CI32 dues aux deux bandes absentes dans deux des profils. L'absence d'une bande augmente de 5 à 10 % la dissimilarité entre deux profils. Ces résultats confirment la reproductibilité de la méthode d'extraction afin d'obtenir les mêmes profils génomiques. C'est pourquoi, pour l'analyse de la diversité intraspécifique de toutes les souches d'*A. pasteurianus*, le choix de l'extraction d'ADN s'est porté sur la méthode rapide par choc thermique et l'amplification d'ADN avec les amorces  $(GTG)_5$  a été employé.

### III.3.2.1.2 Analyse du polymorphisme génomique à l'aide des amorces $(GTG)_5$

L'analyse de la diversité intraspécifique de l'ensemble des souches d'*A. pasteurianus* a été réalisée en étudiant leur polymorphisme génomique à l'aide des amorces (GTG)<sub>5</sub>. Les amplifications d'ADN ont été réalisées en duplicatas afin de s'assurer de la reproductibilité de la méthode. Cent huit souches d'*A. pasteurianus* ont été analysées. Trois souches appartenant à des espèces différentes ont été introduites à l'analyse en tant que « out group » : une souche d'*A. fabarum* (CI25), une d'*A. tropicalis* (MX92) et une souche de *G. oxydans* (GF51). La souche type d'*A. pasteurianus* LMG 01262<sup>T</sup>a également été introduite dans l'analyse. Après amplification et électrophorèse sur gel d'agarose pour les deux répliques, les profils génomiques des souches obtenus présentent de légères différences : absence / présence de quelques bandes entre les deux profils, notamment, les bandes présentant un faible poids moléculaire (entre 600 et 100 pb) et la bande à 3000 pb. C'est pourquoi seules les bandes communes aux deux répliques ont été conservées pour l'analyse du polymorphisme génomique. Les profils (GTG)<sub>5</sub> des souches contiennent entre 7 et 23 fragments d'ADN de 350 à 3000 pb (**Annexe 3**). Visuellement, les profils des souches d'*A. pasteurianus* partagent plusieurs bandes à 350, 380, 480, 980 and 1100 pb. Le profil (GTG)<sub>5</sub> n'a pas pu être obtenu pour 13 souches d'*A. pasteurianus* : CI94, CI95, CI96, CI97, CI98, CI99, CI100, CI102, CI103, CI104, MX107, MX108, MX109. Ces souches ont donc été exclues de l'analyse.

Après analyse numérique, les distances de migration (Rf) des amplicons d'ADN ont été normalisées entre les différents gels à l'aide de deux marqueurs de taille déposés aux extrémités des gels. La présence ou l'absence de chaque bande (1= présence, 0= absence) est notée et utilisée afin de réaliser une matrice en deux dimensions en code binaire pour toutes les souches. Les résultats de l'analyse numérique et statistique des profils génomiques des souches sont présentés dans le dendrogramme **Figure 58**. Les souches d'*A. pasteurianus* ont été regroupées en 6 groupes (I à VI) ayant un minimum de 55 % de similarité et tous les groupes contiennent des souches d'au moins 2 pays.



**Figure 58 :** Dendrogramme généré après analyse statistique de la similarité des profils génomiques obtenus par PCR (GTG)<sub>5</sub>.

Sur 95 souches d'*Acetobacter pasteurianus*, 3 AAB ; *Acetobacter fabarum* (CI25), *Acetobacter tropicalis* (MX92) et *Gluconobacter oxydans* (GF51) isolées lors de fermentations spontanées de cacao dans trois régions du monde et une souche type *A. pasteurianus* LMG0162<sup>T</sup>. L'indice de Sørensen-Dice a été utilisé pour analyser la similarité et un algorithme de regroupement UGPMA a été appliqué pour regrouper les données. Les branches des arbres sont colorées selon l'origine de la souche, le Mexique (bleu), la Côte d'Ivoire (noir), la Guyane française (rouge), la souche type (brun) et les trois souches AAB en vert.

La majorité des souches ont été regroupées en 4 groupes présentant une similarité égale ou supérieure à 60 % entre les groupes avec un nombre variable d'isolats pour chaque groupe. Les groupes II, III, IV, et V présentent 20, 9, 8, et 40, isolats respectivement. Dans les sous-groupes, la plupart des souches présentent un pourcentage de similarité entre les profils d'environ 80 %, indiquant un faible polymorphisme génomique intraspécifique. De plus, au sein des sous-groupes, quatre paires de souches isolées au Mexique présentent une similarité de 100 % : MX75 et MX76 isolées dans le même temps de fermentation et la même caisse de fermentation ; MX90 et MX52 isolées à deux temps de fermentation différents (4 et 5 jours respectivement); MX78 et MX67 isolées au troisième jour de fermentation mais dans 2 caisses de fermentations différentes et MX93 et MX59, isolées à deux temps de fermentation différents. Trois paires de souches isolées en Côte d'Ivoire, présentent une similarité élevée, supérieure ou égale à 90 % : les souches CI12 et CI2, isolées à deux temps de fermentation, les souches CI40 et CI21, isolées au même moment de la fermentation mais en deux caisses différentes et les souches CI19 et CI4, isolées en 2016 et 2018 à deux temps de fermentation différents. En ce qui concerne ces souches présentant des similarités élevées (plus de 90 %), on peut supposer qu'il s'agit des mêmes souches. En revanche, d'autres paires de souches, isolées dans différents pays, ont été regroupées avec un pourcentage élevé de similarité (supérieur à 90 %) pour CI27/GF42, MX61/GF41/CI28 et MX84/CI34 respectivement. Ces résultats indiquent que de mêmes souches peuvent être trouvées dans deux ou trois pays différents.

Deux autres groupes de souches (groupes I et VI) présentent une similarité inférieure à 60 % avec les autres groupes. Le groupe I, avec 58 % de similarité avec les autres groupes, est composé de 8 isolats du Mexique, 6 de Côte d'Ivoire et 1 de Guyane française. Le groupe VI avec 55 % de

similarité avec les autres groupes, est composé de 3 souches, l'une isolée au Mexique (MX70) et deux en Côte d'Ivoire (CI8/CI11). Comme attendu, les 3 souches n'appartenant pas à l'espèce *A. pasteurianus* (CI25, MX92 et GF51) n'ont pas été regroupées avec les souches *A. pasteurianus* et ont montré moins de 35 % de similarité. *A. fabarum* (CI25) étant la souche la plus proche et *G. oxydans* (GF51) la plus éloignée.

Afin d'évaluer l'influence des pays d'origine des souches ou du jour de fermentation de l'isolement de la souche sur le regroupement, des tests statistiques ont été effectués en analysant les matrices de distance utilisant la variance créée avec l'indice Sørensen-Dice. Il semble que le pays d'origine ait un effet important sur le regroupement ( $p$ -value=0,001\*\*) alors que le jour de fermentation de l'isolement n'en a pas ( $p$ -value > 0,05).

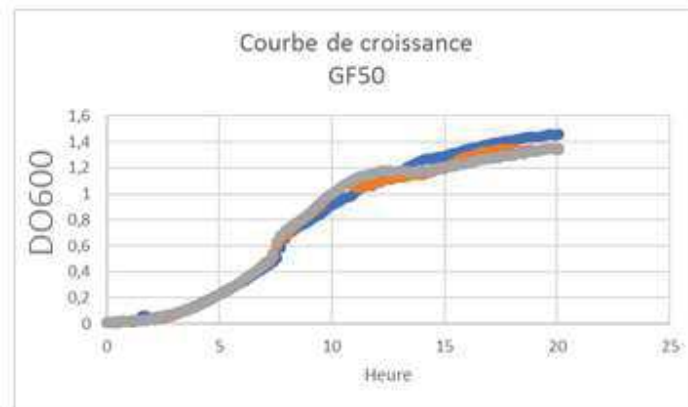
#### *III.3.2.2 Etude du comportement des souches d'*A. pasteurianus* dans le milieu simulant la pulpe de cacao*

Afin d'étudier la diversité biochimique intraspécifique des souches d'*A. pasteurianus* isolées pendant la fermentation du cacao de 3 pays producteurs, leur métabolisme primaire a été étudié dans un milieu simulant les conditions de la pulpe de cacao au second jour de fermentation, le CPSM-AAB (Moens *et al.* 2014).

Cette analyse a été réalisée sur les 95 souches d'*A. pasteurianus* dont le profil (GTG)<sub>5</sub> a pu être obtenu, sur la souche type d'*A. pasteurianus* LMG01262<sup>T</sup> et 3 souches de BAA, *A. fabarum* (CI25), *A. tropicalis* (MX92) et *G. oxydans* (GF51).

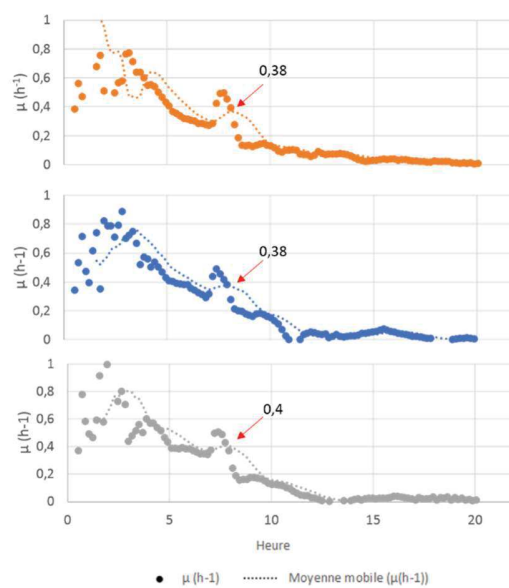
##### *III.3.2.2.1 Détermination du taux de croissance spécifique des souches*

Un suivi de la cinétique de croissance des souches a été réalisé à l'aide d'un spectrophotomètre Spart 10 M (TECAN) pendant 20 h. Après soustraction du blanc (milieu de croissance) les courbes de croissance sont obtenues. Un exemple de courbes de croissance pour la souche GF50 (isolée en Guyane) est montré dans la **Figure 59**.



**Figure 59** : Exemple de courbe de croissance obtenue après le suivi cinétique au spectrophotomètre.

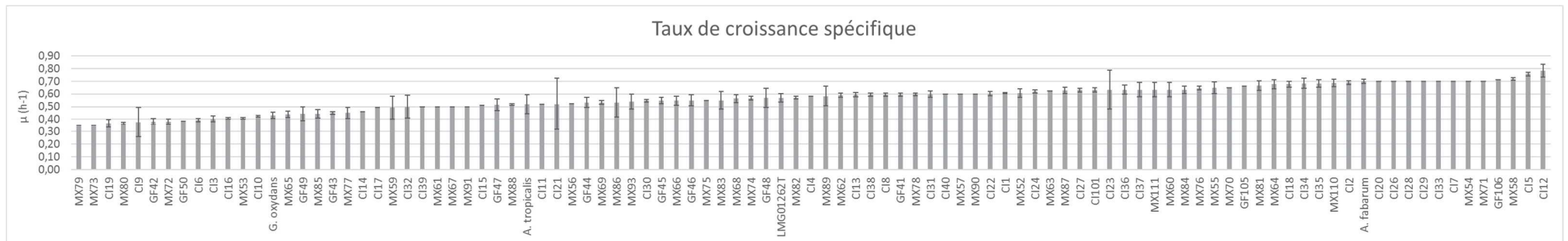
Ensuite, le taux de croissance spécifique est calculé en extrayant la pente de  $\ln(\text{OD600}) = f(\text{temps})$  réalisé sur une moyenne mobile de 10 points et en prenant la valeur maximale (**Figure 60**).



**Figure 60**: Exemple de la détermination du taux de croissance spécifique à l'aide de moyenne mobile sur 10 points.

Les 3 courbes représentent les 3 réplicas, avec le taux de croissance pour chaque courbe représenté par la flèche.



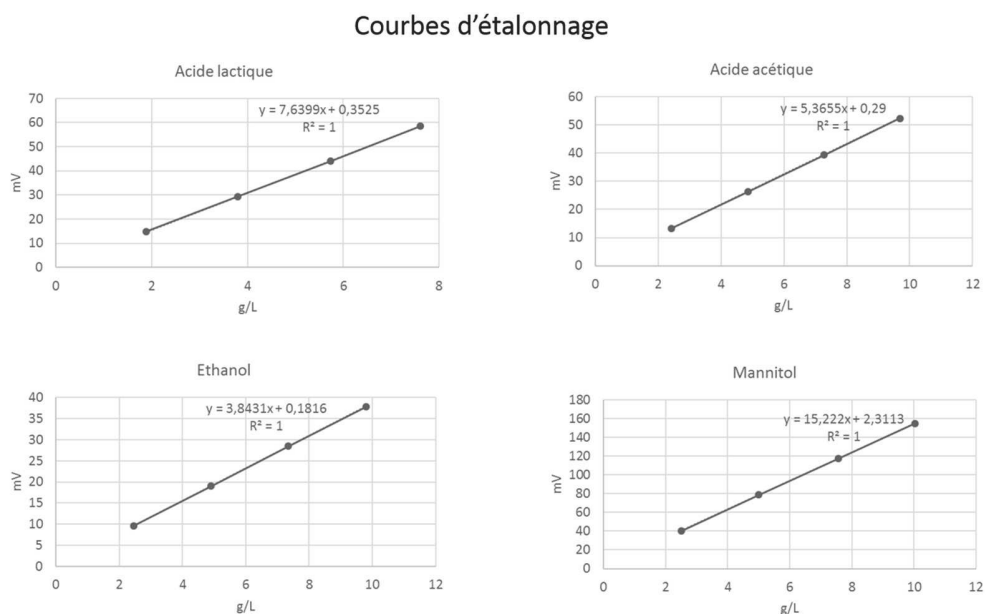


**Figure 61:** Histogramme des taux de croissance spécifique des 95 souches d'*A. pasteurianus*, de la souche type LMG01262<sup>T</sup> et des trois BA : *A. fabarum* (CI25), *A. tropicalis* (MX92) et *G. oxydans* (GF51)

Les taux de croissance spécifiques des 99 souches de BAA allaient de 0.35 (MX79) à 0.78 h<sup>-1</sup>(CI12) (**Figure 61**). Le test non paramétrique de Kruskal-Wallis réalisé après avoir vérifié que les données ne suivent pas une loi normale (Shapiro-Wilk : p-value = 0.015) démontre que le pays d'origine des souches n'a pas d'effet sur la variabilité des valeurs du taux de croissance spécifique (p-value = 0.472).

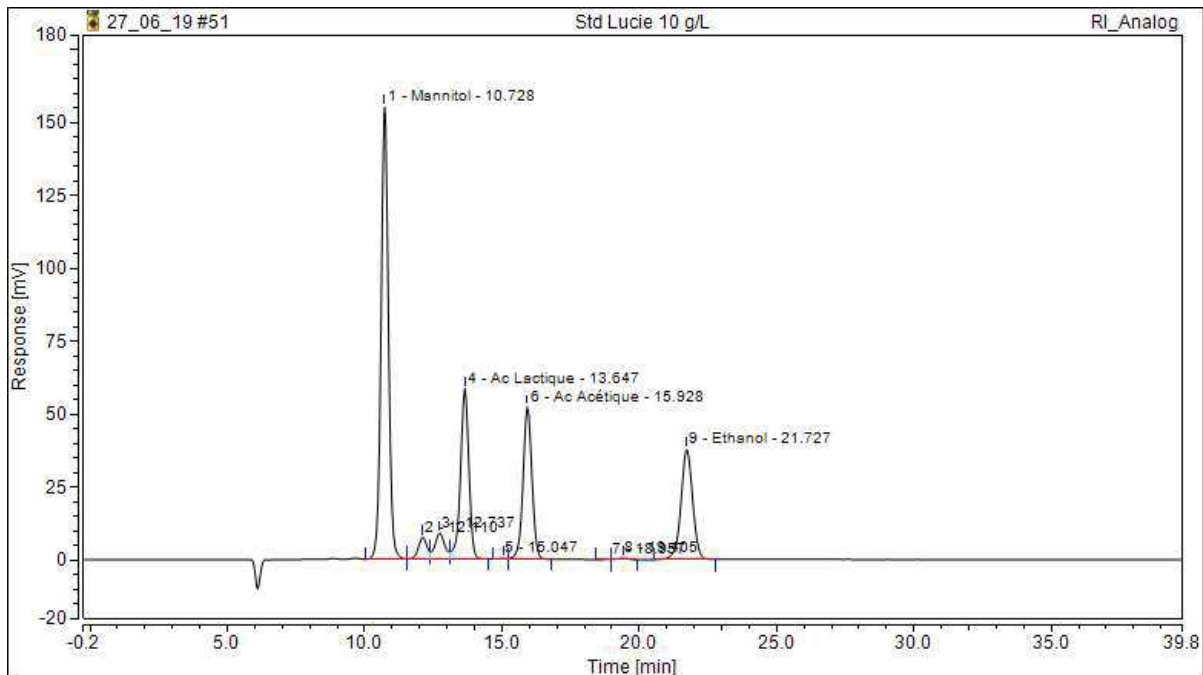
### III.3.2.2.2 Courbes d'étalonnage et analyse des chromatogrammes à l'HPLC

Les courbes d'étalonnage pour la quantification des 3 substrats (mannitol, éthanol, acide lactique) et du principal métabolite produit par les souches d'*A. pasteurianus* ont été construites en utilisant une solution mère à 100 g/L. Puis 5, 10, 15 et 20 µL ont été prélevés et injectés à l'HPLC afin d'avoir une concentration à 2.5, 5, 7.5 et 10 g/L respectivement. Les courbes d'étalonnage ont donc été construites en traçant l'aire des pics en fonction de la concentration (**Figure 62**). Pour chaque courbe d'étalonnage le coefficient de corrélation R<sup>2</sup> obtenu est de 1.



**Figure 62** : Courbes d'étalonnage de l'acide lactique, de l'acide acétique, de l'éthanol et du mannitol.

Le chromatogramme obtenu après l'injection à l'HPLC pour la solution à 10 g/L est montré dans la **Figure 63**. Les temps de rétention pour les 4 composés sont de 10.7, 13.6, 15.9 et 21.7 pour le mannitol, l'acide lactique, l'acide acétique et l'éthanol respectivement.



**Figure 63** : Chromatogramme obtenu après l'analyse à l'HPLC pour la solution étalon à 10 g/L.

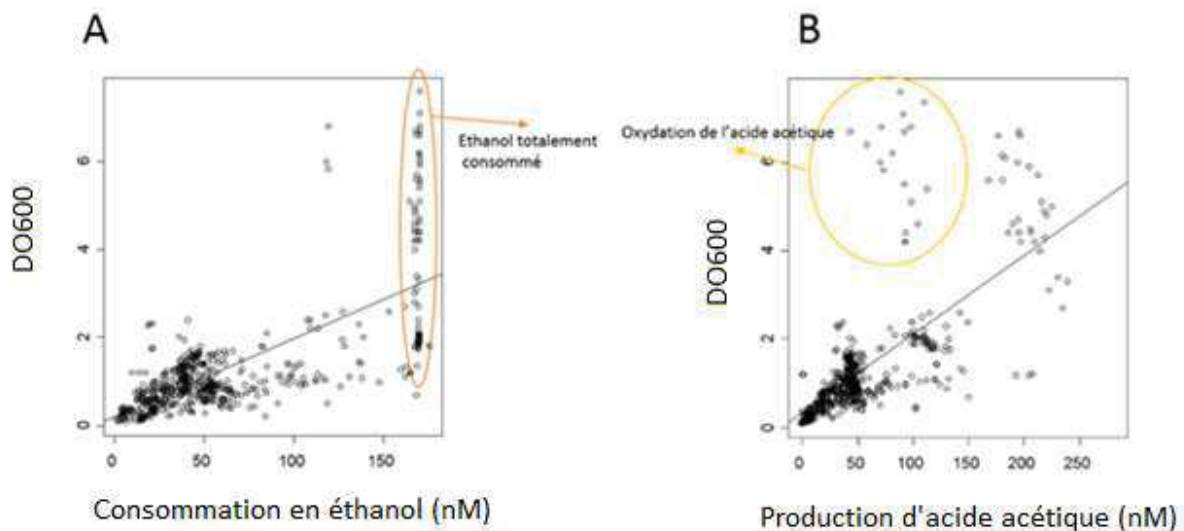
### III.3.2.2.3 Analyse comparative des souches et regroupement en dendrogramme

La consommation des substrats et la production d'acide acétique ont été analysées et différents paramètres ont été calculés :

- Le rendement en biomasse ( $Y_{X/S}$ ) pour déterminer la productivité de biomasse à partir des deux substrats principaux (éthanol et acide lactique),
- L'efficacité de conversion de l'éthanol en acide acétique ( $Y_{AA/ETOH}$ ), qui correspond au rendement expérimental de l'éthanol en acide acétique sur le rendement théorique. Une conversion de 1 mole d'acétique par mole d'éthanol a été considérée.
- Le rendement en biomasse produit par mole d'acide acétique produite ( $Y_{X/AA}$ )

- Le ratio de consommation en acide lactique sur la consommation en éthanol ( $R_{LA/ETOH}$ ). Ce ratio permet de déterminer l'affinité relative pour l'éthanol et/ou pour l'acide lactique comme source de carbone pour leur croissance et/ou pour produire de l'acide acétique.

Afin que toutes les souches soient en phase de croissance exponentielle, nous n'avons considéré que les points entre 0.5 et 3 pour une longueur égale à 600 nm. En effet, à une DO supérieure à 3, une suroxydation de l'acide acétique couplée à une consommation totale de l'éthanol a été observée (**Figure 64A et 64B**) (Raspor et Goranovič, 2008). Pour une DO inférieure à 0.5, la production d'acide acétique et la consommation d'éthanol sont faibles indiquant que les souches ne sont pas phase de croissance.



**Figure 64** : Droite de corrélation entre la DO et la consommation d'éthanol (A) et la production d'acide acétique (B).

Six souches n'ont pas été incluses dans l'analyse : CI101, GF105, GF106, MX93, MX111, MX112. Les DO (600 nm) des souches GF105, MX110 et MX111 étaient inférieures à 0.5 lors des précultures et lors des cultures statiques en pot. La souche GF106, a présenté des DO à 600 nm inférieures à 0.5 ou supérieures à 3. Pour la souche CI101, ses DO 600 étaient inférieures à 0.5 à 18h et pour les répliques à 40h, une surestimation de la consommation d'éthanol a été observée lors de l'analyse HPLC rendant impossible sa comparaison avec les autres souches. Finalement, la souche MX93 avait le même comportement que la souche de BA, *A. tropicalis* (MX92) analysée lors de cette étude, une probable contamination croisée lors de la réalisation des cultures pourrait

être à l'origine de ces comportements similaires. Au total, 89 souches de *A. pasteurianus*, ainsi que la souche type LMG01262<sup>T</sup> et 3 souches de BAA (CI25, MX92 et GF51) ont été analysées.

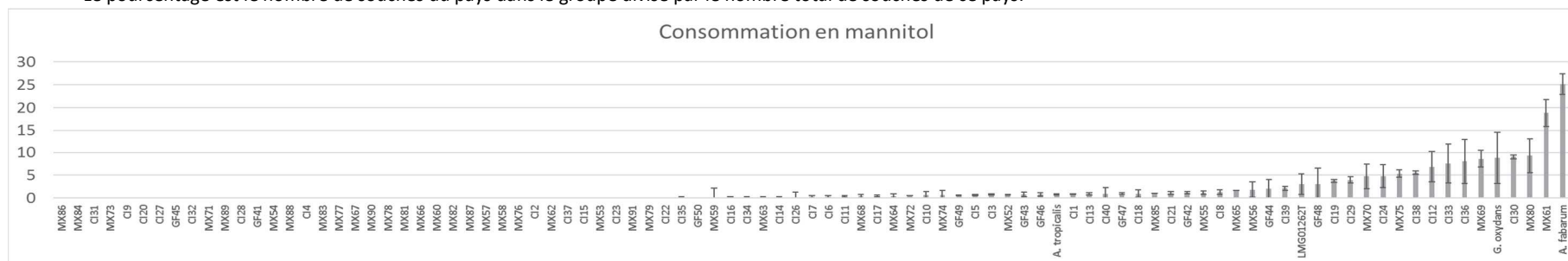
La plupart des souches n'ont pas ou peu consommé le mannitol présent dans le milieu dans nos conditions expérimentales de 0 (MX86) à 20 (*A. fabarum*) mM/UOD, **Figure 65**. Quarante-six pourcent des souches n'ont pas consommé le mannitol, 27 % ont consommé de 0.1 à 1 mM, 15% de 1 à 5 mM, 10 % de 5 à 9 mM et seulement 2 % plus de 10 mM. Ce substrat n'a donc pas été pris en compte dans l'analyse comparative des souches. Une étude réalisée sur une souche d'*A. pasteurianus* a montré que le mannitol n'est consommé que lorsque l'éthanol est complètement épuisé et que le lactate reste en très faible quantité (<2g/L), une condition qui n'est jamais atteinte dans les expériences considérées pour nos analyses (toutes les DO600 < 3 et l'éthanol restant dans le milieu). Il a déjà été démontré qu'*A. pasteurianus* oxyde habituellement le sucre très lentement jusqu'à épuisement de l'éthanol (Sakurai *et al.* 2011 ; Moens *et al.* 2014).

Les différents paramètres sélectionnés pour la comparaison des souches ont été compilés afin de les regrouper et une vue d'ensemble est montrée dans le dendrogramme et la heatmap (**Figure 66**). L'utilisation du Z-score permet de comparer les souches entre-elles puisque le z-score exprime l'écart-type des valeurs par rapport à la moyenne de l'ensemble des souches. Selon l'histogramme des données de Z-scores (**Figure 66**), la plupart des valeurs sont regroupées entre -1 et 1 indiquant une forte similarité entre les différentes valeurs des paramètres des souches. Si l'on regarde le dendrogramme, plusieurs sous-groupes, qui présentent une forte similarité entre souches avec une distance Euclidienne inférieure à 2 sur une échelle de 8, sont composés de souches d'une seule origine. Afin d'évaluer l'influence du pays d'origine et le jour de fermentation sur le regroupement des souches, des tests statistiques ont été effectués en analysant les matrices de distance Euclidiennes. Il semble que le pays d'origine ait un effet sur le regroupement (p-value=0,028\*) alors que le jour de la fermentation n'en a pas (p-value>0,05). Les souches d'*A. pasteurianus* ont été regroupées en deux grands groupes présentant une forte similarité entre les souches (**Figure 66**). La répartition des souches dans les deux groupes est détaillée dans le **Tableau 28**.

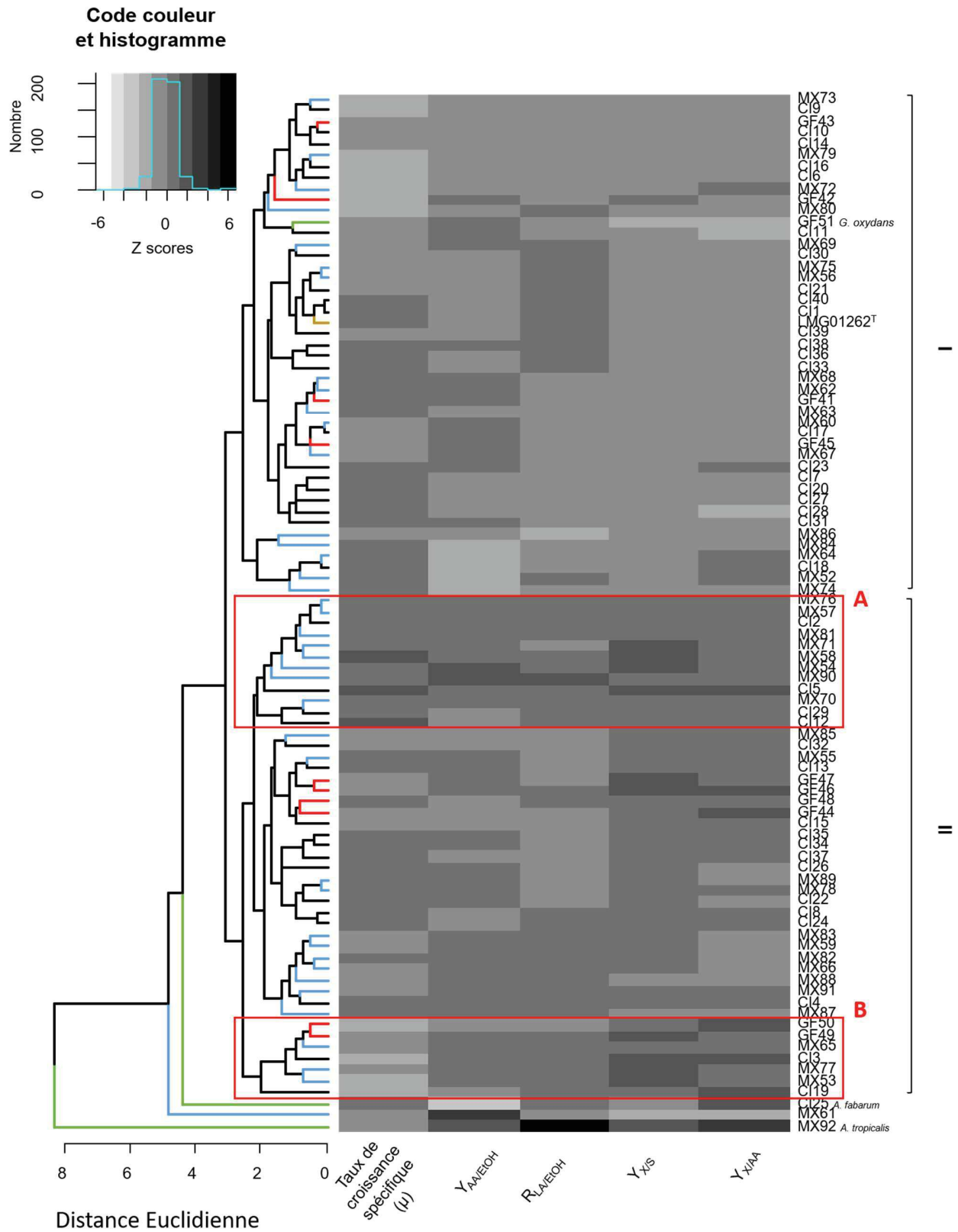
**Tableau 28** : Répartition des souches dans les deux grands groupes du dendrogramme et de la heatmap **Figure 66**.

Groupe métabolique	Nombre d'isolats	Isolats	Pays d'origine
I	45	LMG01262 <sup>T</sup> ; CI1; CI10; CI11; CI14; CI16; CI17; CI18; CI20; CI21; CI23; CI27; CI28; CI30; CI31; CI33; CI36; CI38; CI39; CI40; CI6; CI7; CI9; GF41; GF42; GF43; GF45; MX52; MX56; MX60; MX62; MX63; MX64; MX67; MX68; MX69; MX72; MX73; MX74; MX75; MX79; MX80; MX84; MX86; GF51	Côte d'Ivoire (22/39, 56.5 %*) & Mexique (17/40, 42.5 %*) and Guyane (4/10, 40 %*)
II	45	CI12; CI13; CI15; CI19; CI2; CI22; CI24; CI26; CI29; CI3; CI32; CI34; CI35; CI37; CI4; CI5; CI8; GF44; GF46; GF47; GF48; GF49; GF50; MX53; MX54; MX55; MX57; MX58; MX59; MX65; MX66; MX70; MX71; MX76; MX77; MX78; MX81; MX82; MX83; MX85; MX87; MX88; MX89; MX90; MX91	Mexique (22/40, 55 %*), Côte d'Ivoire (18/39, 43.5 %*) et Guyane (6/10, 60 %*)

\* Le pourcentage est le nombre de souches du pays dans le groupe divisé par le nombre total de souches de ce pays.



**Figure 65** : Consommation en mannitol des 89 souches d'*A. pasteurianus*, et 3 BAA : *G. oxydans*, *A. fabarum* et *A. tropicalis*



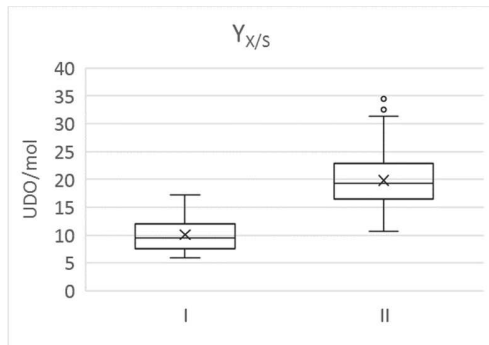
**Figure 66** : Relation biochimique et regroupement en dendrogramme des 89 souches d'*Acetobacter pasteurianus*, la souche de LMG01262<sup>T</sup> et les 3 BAA.

Les mesures ont été converties en Z-scores, les similarités par paires ont été calculées par distance euclidienne et un algorithme de regroupement UGPMA a été appliqué pour regrouper les données. Avec  $Y_{\text{EtOH/AA}}$  = efficacité de conversion de l'éthanol en acide acétique,  $R_{\text{LA/EtOH}}$  = correspond à la consommation d'acide lactique divisée par la consommation d'éthanol,  $Y_{\text{X/S}}$  = rendement biomasse,  $Y_{\text{X/AA}}$  = rendement biomasse produit par mol d'acide acétique produit et taux de croissance spécifique ( $\mu$ ). Où X est la biomasse, S = consommation de substrats, LA = acide lactique, AA = acide acétique. EtOH = éthanol. Les branches des arbres sont colorées selon l'origine de la souche, le Mexique (bleu), la Côte d'Ivoire (noir), la Guyane française (rouge), la souche type (brun) et les trois souches AAB en vert.

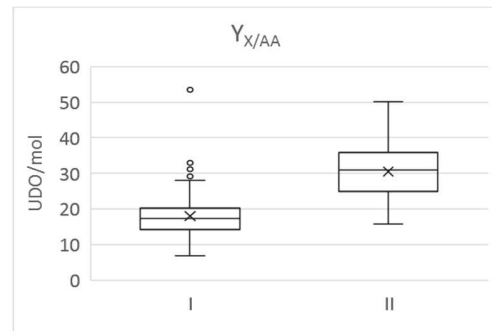
Le groupe métabolique I (GI) est composé de 45 souches dont 22 isolats de Côte d'Ivoire (soit 56,5 % des souches ivoiriennes), 17 isolats du Mexique (soit 42,5 % des souches mexicaines), 4 souches de Guyane (soit 40 % des souches guyanaises), la souche type LMG01262<sup>T</sup> et *G. oxydans* (GF51). Des valeurs moyennes de  $Y_{\text{X/S}}$ ,  $Y_{\text{X/AA}}$  et  $Y_{\text{AA/EtOH}}$  ont été observées pour les souches appartenant à ce groupe. Le groupe II (GII) est composé de 45 souches, dont 22 du Mexique (soit 55 % des souches mexicaines), 17 isolats de Côte d'Ivoire (soit 43,5 % des souches ivoiriennes), et 6 de Guyane (soit 60 % des souches guyanaises). Les souches de ce groupe ont été caractérisées par des valeurs plus élevées, avec un z-score supérieur à 1 de  $Y_{\text{AA/EtOH}}$ ,  $R_{\text{LA/EtOH}}$ ,  $Y_{\text{X/S}}$ , et  $Y_{\text{X/AA}}$  que les souches du GI. La souche MX61 n'est pas regroupée avec les souches d'*A. pasteurianus* et présente la valeur  $Y_{\text{AA/EtOH}}$  la plus élevée (z-score = 3,8) avec une faible valeur pour  $Y_{\text{X/S}}$  (z = -1,52) et  $Y_{\text{X/AA}}$  (z = -2,03) comparativement aux autres souches. La souche de *G. oxydans* (GF51) est regroupée avec les souches d'*A. pasteurianus*, cependant avec un rendement moyen de  $Y_{\text{AA/EtOH}}$  (z = 0,9) et un rendement faible de  $Y_{\text{X/AA}}$  (z = -1,87). Les 2 souches *A. fabarum* (CI25) et *A. tropicalis* (MX92) ne sont pas regroupées avec les souches d'*A. pasteurianus*. L'efficacité de conversion de l'éthanol en acide acétique d'*A. fabarum* est inférieure à celle d'*A. tropicalis* comme il a été démontré dans une précédente étude (Moens *et al.* 2014). Cette dernière se caractérise par un  $R_{\text{LA/EtOH}}$  et  $Y_{\text{X/AA}}$  plus élevé par rapport aux souches d'*A. pasteurianus*, z = 0,86 et z = 1,96 respectivement.



Les souches du GI métabolique présentent des valeurs plus faibles de  $Y_{X/S}$  (médiane = 9,38 UDO/mol) (**Figure 67**) et  $Y_{X/AA}$  (médiane = 17,5 UDO/mol) (**Figure 68**) que les souches de GII avec  $Y_{X/S}$  (médiane = 19,35 UDO600/mol) et  $Y_{X/AA}$  (médiane = 31 UDO600/mol). Révélant une plus grande efficacité de croissance pour les souches de GII, les souches nécessitent moins de substrat pour produire de la biomasse que les souches de GI.

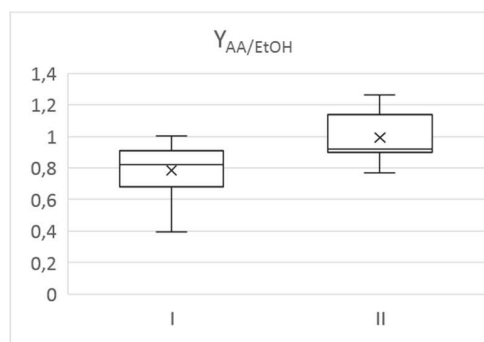


**Figure 67** : Boîtes à moustaches des rendements de biomasse en mol par unité de DO pour les deux groupes obtenus dans le dendrogramme et heatmap **Figure 66**



**Figure 68** : Boîtes à moustaches des rendements de biomasse, en mol par unité de DO, produite par mole d'AA produite pour les deux groupes obtenus dans le dendrogramme et heatmap **Figure 66**

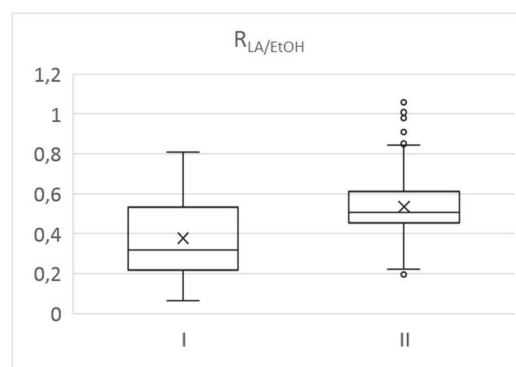
Les deux groupes ont cependant une grande efficacité de conversion de l'éthanol en acide acétique ( $Y_{AA/EtOH}$ ), GI=0,83 et GII=0,94 valeur médiane (**Figure 69**).



**Figure 69** : Boîtes à moustaches pour l'efficacité de conversion de l'éthanol en acide acétique pour les deux groupes obtenus dans le dendrogramme et heatmap **Figure 66**.

GII présente une valeur  $Y_{AA/ETOH}$  supérieure, expliquée par certaines souches qui ont des valeurs supérieures à 1 tandis que dans GI, certaines souches ont des valeurs inférieures à 0,5. Pour les souches qui ont des valeurs inférieures à 1, voire à 0,5, cela suggère une suroxydation de l'acide acétique produit à partir de l'éthanol. La suroxydation de l'acide acétique a été rapporté comme étant provoquée par une activité accrue des enzymes du cycle de l'acide tricarboxylique (TCA) et de l'acétyl-CoA synthétase pour l'approvisionnement énergétique. Cependant, chez les BAA, cette suroxydation de l'acide acétique est réalisée normalement en absence d'éthanol, mais ce mécanisme de commutation entre l'oxydation incomplète et la suroxydation assimilatoire dans le TCA, et le contrôle du métabolisme de l'énergie et du carbone reste méconnu (Adler *et al.* 2014; Saeki *et al.* 1999; Sakurai *et al.* 2011; Yang *et al.* 2019).

Les deux groupes ont une préférence pour l'éthanol plutôt que pour l'acide lactique ; en effet, leurs  $R_{AL/ETOH}$  sont inférieurs à 1 (**Figure 70**). Ces données sont confirmées par l'analyse de la consommation spécifique d'éthanol et d'acide lactique des 89 souches d'*A. pasteurianus*. Elles ont consommé généralement plus d'éthanol que d'acide lactique, entre  $25,58 \pm 3,49$  et  $132,15 \pm 8,34$  mM/UDO (**Figure 71**) et  $7,5 \pm 0,79$  et  $56,71 \pm 5,07$  mM/UDO (**Figure 72**) respectivement (valeur moyenne pour chaque souche) ce qui est en adéquation avec les précédentes études réalisées sur les souches d'*A. pasteurianus* (Lefeber *et al.* 2010; Moens *et al.* 2014).



**Figure 70** : Boîtes à moustaches des ratios de consommation de l'acide lactique sur la consommation de l'éthanol pour les deux groupes obtenus dans le dendrogramme et heatmap **Figure 66**.

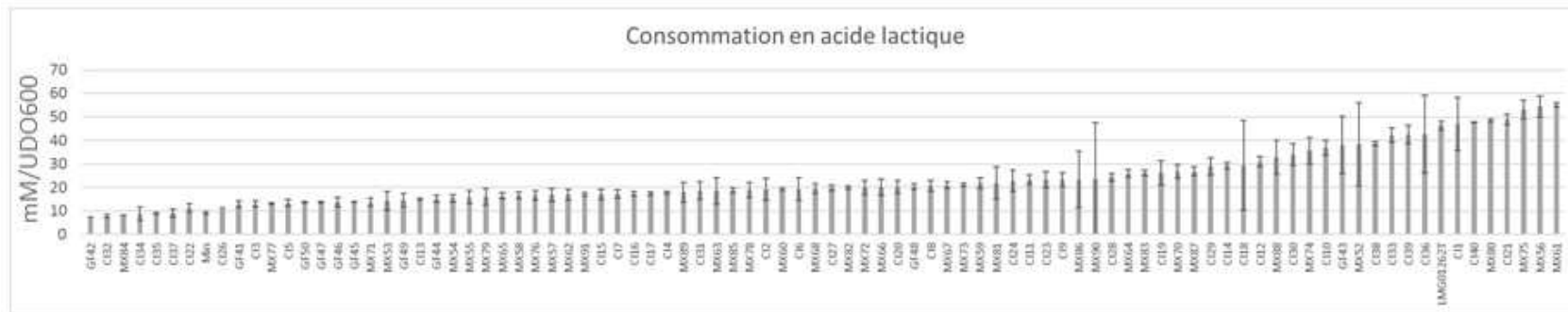


Figure 71 : Histogramme de la consommation en acide lactique selon les 89 souches d'*Acetobacter pasteurianus* et la souche type LMG 01262<sup>T</sup>

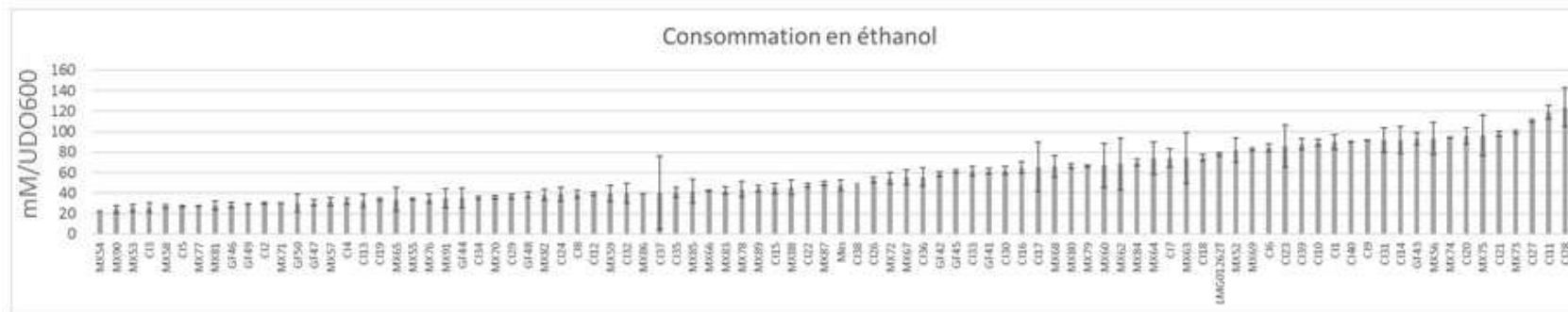
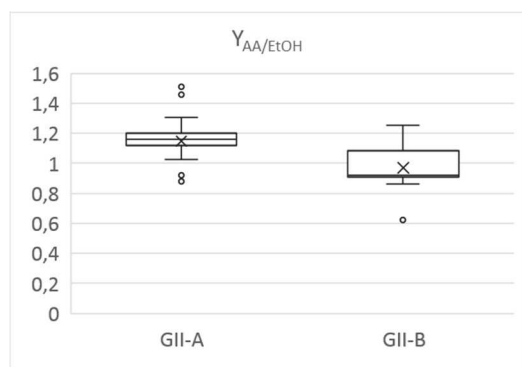


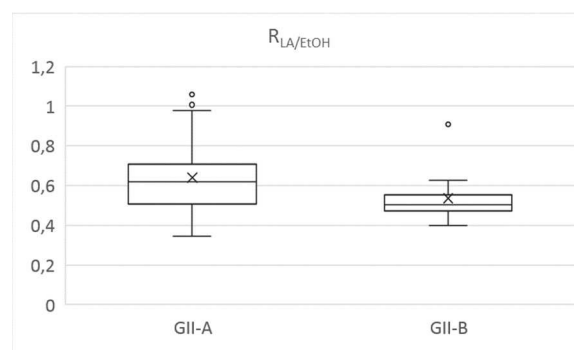
Figure 72 : Histogramme de la consommation en éthanol selon les 89 souches d'*Acetobacter pasteurianus* et la souche type LMG01262<sup>T</sup>

Cependant, les souches de GII ont un  $R_{LA/ETOH}$  plus élevé, avec une valeur médiane de 0,5 alors que le groupe GI avait un  $R_{LA/ETOH}$  médian de 0,31 indiquant une affinité supérieure pour l'acide lactique.

En effet, si on s'intéresse de plus près au GII, deux sous-groupes, GII-A et de GII-B (**Figure 66**) semblent avoir une plus grande consommation de l'acide lactique avec une valeur médiane élevée pour le  $R_{LA/ETOH}$  que le GI (GII-A = 0,61, GII-B= 0.5, **Figure 74**) et également une valeur médiane élevée pour  $Y_{AA/ETOH}$  (GII-A= 1.16 et GII-B=0.91) (**Figure 73**) dont le  $R_{LA/ETOH}$  de certaines souches était égal ou proche de 1, comme les souches MX90, CI19, MX81, suggérant une grande affinité pour l'acide lactique. Ces 3 souches avaient également un  $Y_{AA/ETOH}$  proche ou supérieur à 1. Ces valeurs élevées, indiquaient que pour la production d'acide acétique, la quantité d'éthanol n'était pas suffisante et souligne l'implication potentielle de l'acide lactique. Ces résultats suggèrent que les souches pourraient utiliser l'acide lactique pour produire de l'acide acétique ou que l'acide lactique est consommé en même temps que l'éthanol. Ces sous-groupes avaient également une productivité élevée en biomasse, entraînant un rendement de production d'acide acétique élevé. Les souches du groupe I n'ont alors utilisé que peu d'acide lactique pour produire de l'acide acétique et auraient utilisé plus l'acide lactique pour la biomasse. Dans la fermentation du cacao, l'acide lactique est un métabolite indésirable, car contrairement à l'acide acétique, qui est volatil, l'acide lactique ne peut être éliminé des fèves lors des étapes de séchage (Barel 2013) et provoque un goût désagréable au cacao final (Holm *et al.* 1993).

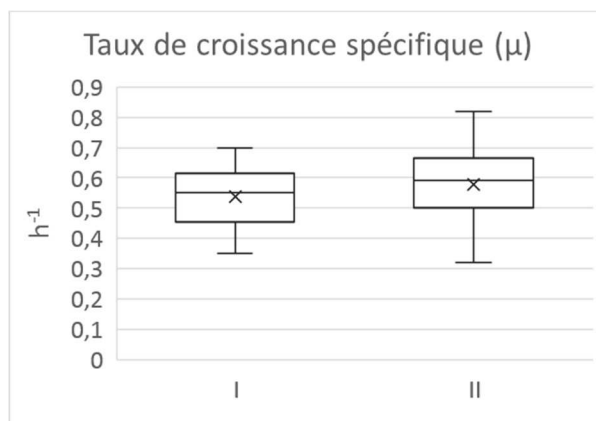


**Figure 73** : Boîtes à moustaches pour l'efficacité de conversion de l'éthanol en acide acétique pour les deux sous-groupes de GII.



**Figure 74** : Boîtes à moustaches des ratios de consommation de l'acide lactique sur la consommation de l'éthanol pour les deux sous-groupes de GII.

Concernant le taux de croissance spécifique, la valeur médiane pour chaque groupe est d'environ  $0,6 \text{ h}^{-1}$ , les tests statistiques montrent une différence non significative entre les 2 groupes avec une valeur de p-value à 0,091 (**Figure 75**).



**Figure 75** : Boîtes à moustaches pour les taux de croissance spécifique pour les deux groupes obtenus dans le dendrogramme et heatmap **Figure 66**.

### II.3.2.3 Conclusion partielle

L'analyse du polymorphisme génomique a mis en évidence une importante similarité entre les souches d'*A. pasteurianus* isolées des trois pays. Les principaux groupes obtenus étant composés de souches des trois pays. Cependant, le test statistique réalisé sur la matrice de distance (Sørensen-Dice) souligne un effet important de l'origine sur le regroupement des souches (p-value=0,001\*\*). En effet, la plupart des souches regroupées deux à deux et présentant un pourcentage élevé de similarité (> 80 %) appartiennent à la même origine géographique ce qui explique la significativité du test statistique.

L'analyse du comportement des souches dans le milieu CPSM-AAB, a également montré une importante similarité entre elles. Avec deux grands groupes composés de souches en provenance des 3 pays, le test statistique réalisé sur la matrice de distance (Euclidienne) souligne un effet important de l'origine sur le regroupement des souches (p-value=0,001\*\*) expliqué par le fait que de nombreux sous-groupes qui ont une distance euclidienne inférieures à 2 sont principalement composés d'une seule origine. Cependant, ces analyses ont mis en évidence une affinité différente vis-à-vis de l'acide lactique entre les souches et ce indépendamment du pays d'isolement d'origine. Certaines souches, notamment celles dans

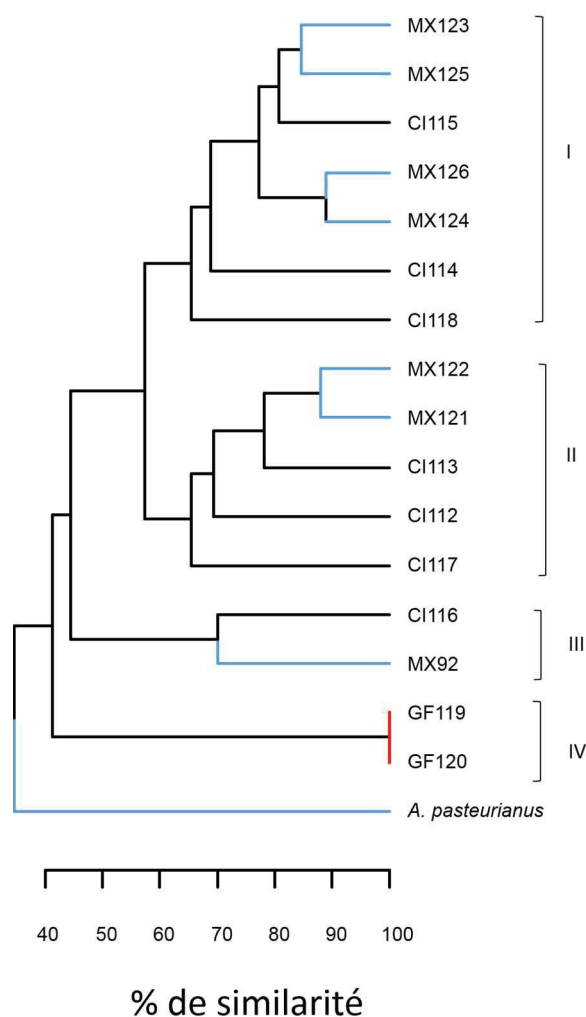
le GII, ont consommé plus d'acide lactique, comme le prouve leur ratio  $R_{LA/ETOH}$ . Une autre caractéristique a été mise en évidence, c'est que les souches dans le GI, avaient probablement suroxydé l'acide acétique. Ces deux caractéristiques sont à la base du regroupement en deux grands groupes. La souche d'*A. tropicalis* est regroupée à part des souches d'*A. pasteurianus*, elle diffère étant donné sa grande efficacité d'oxydation de l'éthanol et de l'acide lactique en acide acétique et une très grande affinité pour ce dernier.

Cependant, les deux analyses ne sont pas corrélées. Dans l'analyse du polymorphisme génomique, seules deux paires de souches ayant obtenu un pourcentage élevé de similarité sont regroupées dans le même groupe dans le dendrogramme basé sur l'analyse de leur comportement en milieu CPSM. L'utilisation de la (GTG)<sub>5</sub> rep PCR ne semble pas suffisamment discriminante pour différencier les souches ayant des profils métaboliques différents.

### III.3.3 Analyse du polymorphisme génomique des souches d'*A. tropicalis*

La même stratégie a été employée pour analyser la diversité intraspécifique des 16 souches d'*A. tropicalis*. Leur polymorphisme génomique a été déterminé à l'aide des amorces (GTG)<sub>5</sub>. Les amplifications d'ADN ont été réalisées également en duplicatas afin de s'assurer de la répétabilité de la méthode. Une souche d'*A. pasteurianus* isolée pendant la thèse a été ajoutée à l'analyse en tant que « out group ». Après amplification et électrophorèse sur gel d'agarose pour les deux amplifications, les profils génomiques des souches obtenus présentaient de légères différences, absence / présence de quelques bandes entre les deux profils, notamment, les bandes entre 3000 et 2000 bp. C'est pourquoi seules les bandes communes aux deux amplifications ont été conservées pour l'analyse du polymorphisme génomique. Les profils (GTG)<sub>5</sub> des souches contiennent entre 6 et 17 fragments d'ADN de 350 à 3000 pb. Visuellement, les profils des souches d'*A. pasteurianus* partageaient plusieurs bandes à 350, 1300 and 1400 pb 3000 pb (**Annexe 4**).

Les résultats de l'analyse numérique et statistique des profils des génomiques des souches sont présentés dans le dendrogramme **Figure 76**.



**Figure 76** : Dendrogramme généré après analyse statistique de la similarité des profils génomiques obtenus par PCR (GTG)<sub>5</sub>.

Sur 16 souches d'*Acetobacter tropicalis* et une souche d'*Acetobacter pasteurianus* ont été isolées lors de fermentations spontanées de cacao dans trois pays. L'indice de Sørensen-Dice a été utilisé pour analyser la similarité et un algorithme de regroupement UGPMa a été appliqué pour regrouper les données. Les branches des arbres sont colorées selon l'origine de la souche, le Mexique (bleu), la Côte d'Ivoire (noir), la Guyane française (rouge), la souche d'*A. pasteurianus* en vert.

Les souches d'*A. tropicalis* ont été regroupées en 4 groupes (I à IV, **Figure 76**) ayant un minimum de 42.5 % de similarité. Les deux souches isolées en Guyane (GF119 et GF120) sont regroupées ensemble dans le groupe IV à 42.5% de similarité avec les autres souches et sont à 100% similaires. Elles ont été isolées lors du second jour de fermentation, il s'agit de la même souche isolée deux fois. Concernant les trois autres groupes, le groupe I, était composé de 4 souches isolées au Mexique et de 3 souches isolées en Côte d'Ivoire, avec une similarité à 60

% avec le groupe II. Les deux paires de souches du Mexique, MX123 / MX125 et MX124 / MX126 sont regroupées ensemble avec une similarité à plus de 80 %. Les souches des deux paires ont été isolées lors des deux campagnes réalisées au Mexique à deux temps de fermentation différents, cela suggère qu'au cours d'une campagne à l'autre la même souche peut être isolée deux fois. Les souches en provenance de Côte d'Ivoire n'ont pas été regroupées ensemble dans le groupe I mais ont une similarité supérieure à 60% avec les souches de Mexique et entre elles. Le groupe II, était composé de 2 souches du Mexique et 3 souches de Côte d'Ivoire. Les deux souches isolées au Mexique, de la même campagne au même temps de fermentation, sont regroupées avec un pourcentage élevé de similarité à plus de 85 %, il s'agit de la même souche isolée deux fois. Comme dans le groupe I, les souches isolées en Côte d'Ivoire n'étaient pas regroupées entre elles mais présentaient une similarité élevée supérieure à 60 % avec les autres souches. Finalement, le groupe III, regroupées avec moins de 50% de similarité avec les autres groupes, était composé de deux souches, une isolée au Mexique et une isolée en Côte d'Ivoire, Ces deux souches sont regroupées avec un pourcentage de 70 % d'identité. Suggérant, qu'une souche peut être isolée dans deux pays différents. Comme attendu, la souche d'*A. pasteurianus* ajoutée en tant qu'out group pour enraciner l'arbre n'est pas regroupée avec les souches d'*A. tropicalis*, avec un pourcentage de similarité inférieur à 40 % avec les autres souches. Afin d'évaluer l'influence du pays d'origine et le jour de fermentation sur le regroupement des souches, des tests statistiques ont été effectués en analysant les matrices de distance. Il semble que le pays d'origine ait eu un effet sur le regroupement (p-value=0,018\*) alors que le jour de la fermentation n'en a pas eu (p-value>0,05). Cette significativité est due aux souches du Mexique et de la Guyane qui sont regroupées ensemble avec un fort pourcentage de similarité.

### III.3.4 Conclusion partielle

Pour conclure ce chapitre, l'utilisation d'échantillons frais de fèves avec la pulpe encore présente a été un point clé pour l'élaboration de la collection de souches. L'utilisation de milieux spécifiques pour l'isolement des bactéries acétiques a été concluant, et nous a permis d'isoler 149 BAA avec une majorité de souches d'*A. pasteurianus* (108) et *A. tropicalis* (16) et 63 BAL avec une majorité de souches de *L. plantarum* (55).



Pour les souches d'*A. pasteurianus*, l'analyse de leurs polymorphismes génomiques nous a indiqué que les souches peuvent être retrouvées dans plusieurs pays différents, même si le test statistique nous indique une forte influence de l'origine géographique sur le regroupement. Cela est expliqué par le fait que les souches qui avaient un pourcentage élevé de similarité, supérieur à 90 %, appartenaient à la même origine. Sinon, tous les groupes sont composés de souches en provenance de 2 ou 3 pays. Les souches isolées ont un degré élevé de similarité entre elles, plus de 60 %, dans la littérature il a été démontré que le génome des souches d'*A. pasteurianus* avaient un fort degré d'homologie. Plus de 68 % des protéines prédites sont partagées et forment le core génome de l'espèce (Illegheems *et al.* 2013). Ce fort degré d'homologie a également été observé lors de la comparaison du génome de souches d'*A. pasteurianus* provenant de différentes matrices alimentaires, comme le vinaigre.

En ce qui concerne, les souches d'*A. tropicalis*, il semble que l'origine géographique ait un impact sur leur diversité. Les deux souches isolées en Guyane sont regroupées entre elles avec un pourcentage faible de similarité avec les autres souches, moins de 50 % et les souches isolées au Mexique sont regroupées en paire avec un fort pourcentage de similarité, suggérant que c'est la même souche. D'un point de vue du polymorphisme génomique, les souches d'*A. pasteurianus* ne semblent pas être différentes selon leur origine géographique contrairement aux souches d'*A. tropicalis*. *A. pasteurianus* étant l'espèce dominante des fermentations, il se pourrait qu'au cours du temps, une souche ait dominé les autres par une plus forte résistance à l'environnement acide (à un pH de l'ordre de 4,5) et chaud (45-50°C) lors des fermentations. Contrairement à *A. tropicalis*, cette espèce n'est pas retrouvée dans toutes les fermentations de cacao, il se pourrait donc qu'il n'y a pas de souche particulière qui serait dominante.

Concernant, l'analyse du comportement des souches d'*A. pasteurianus* dans le milieu CPSM, l'influence de l'origine géographique sur le regroupement des souches semble être plus importante que lors de l'analyse du polymorphisme génomique.

L'utilisation de la méthode de (GTG)<sub>5</sub>, ne semble pas adaptée pour distinguer les différences métaboliques obtenues. Cette analyse nous a permis de mettre en évidence deux caractéristiques importantes dans le choix d'une souche en vue du développement d'un starter de culture.

La première est que certaines souches semblent avoir une affinité plus grande pour l'acide lactique. L'acide lactique est un métabolite produit par les BAL en début de fermentation, il n'est pas volatil contrairement à l'acide acétique et ne pourra pas être éliminé lors des étapes de séchages des fèves. Il n'est pas un métabolite désiré lors de la fermentation des fèves, car il provoque des notes désagréables au cacao final. Sélectionner des souches qui consomment plus l'acide lactique que d'autres semble être un critère important pour l'élaboration d'un starter de cultures. La seconde caractéristique est que certaines souches semblent oxyder l'acide acétique produit alors que l'éthanol est encore présent dans le milieu. L'acide acétique est un métabolite essentiel au cours de la fermentation du cacao, car il permet d'activer des enzymes endogènes de la fève pour initier la production d'arôme. Les souches comportant cette caractéristique ne sont pas intéressantes dans le but d'élaborer un starter de culture pour améliorer et contrôler la qualité organoleptique du cacao.

Dans la collection de souches établie lors de cette étude, des souches se distinguent donc des autres et pourrait être sélectionnées pour l'élaboration d'un starter de culture. Ces souches ont une affinité supérieure pour l'acide lactique, une production en biomasse efficace, elles consomment moins de substrats que les autres pour produire leur biomasse. Ces souches proviennent de deux pays différents, une de la Côte d'Ivoire (CI19) et deux souches du Mexique (MX81, MX90).

## Conclusion et perspectives

Ce travail a consisté à étudier la composition des communautés bactériennes associées aux fèves de cacao en fermentation d'un point de vue inter et intra spécifique selon différentes origines géographiques (Mexique, Côte d'Ivoire et Guyane) et d'identifier quelle est l'origine technologique/géographique du pool de bactéries impliqué dans les fermentations réalisées en caisse au Mexique.

Nos travaux ont permis de mettre en évidence une plus grande diversité des communautés bactériennes aux premiers jours de fermentation avec un impact majeur de l'origine géographique sur leurs compositions en genres. La communauté bactérienne de la Guyane en début de fermentation est moins diversifiée avec un genre qui domine la fermentation : *Fructobacillus*. Il appartient à la famille des BAL et a déjà été isolé dans des fermentations réalisées dans la même zone géographique (Amérique du sud) en Equateur, au Brésil, au Honduras et au Nicaragua (Papalexandratou *et al.* 2011a, 2013, 2019; Romanes *et al.* 2018; Serra *et al.* 2019) et faiblement détecté dans une fermentation réalisée au Ghana (Agyirifo *et al.* 2019). Les bactéries du genre *Fructobacillus* pourraient être spécifiques aux fermentations réalisées en Amérique du Sud. Pour les fermentations de Côte d'Ivoire, 3 genres bactériens environnementaux ont été détectés : *Enterococcus*, *Sphingobacterium* et *Providencia*. Ils sont connus pour être des pathogènes opportunistes de l'Homme et ne sont normalement pas impliqués dans la fermentation des fèves de cacao. Ce sont des contaminations provenant principalement des manipulateurs et des insectes qui ont été en contact avec les fèves pendant la mise en route du processus de fermentation. Enfin au Mexique, une plus grande diversité a été observée. Trois genres de BAA ont été détectés : *Gluconobacter*, *Asaia* et *Komagataeibacter*. *Asaia* est le genre de BAA qui n'est pas capable d'oxyder l'éthanol en acide acétique (Yamada *et al.* 2000). Il a déjà été isolé lors de fermentation du cacao (Pereira *et al.* 2013) mais son rôle n'est pas connu à ce jour. Certaines espèces d'*Asaia* et de *Gluconobacter* sont connus pour être des pathogènes opportunistes de l'Homme (Snyder *et al.* 2004; Alauzet *et al.* 2010). Des espèces du genre *Gluconobacter* sont connues pour être impliquées dans les fermentations des fèves de cacao à travers les différents pays du globe, Brésil, Equateur, Ghana, Malaisie, Nicaragua (Papalexandratou *et al.* 2011a, 2011b, 2013, 2019; Agyirifo *et al.*

2019; Serra *et al.* 2019). Le genre *Komagataeibacter* a été isolé de fermentation de fèves de cacao pour la première fois au Brésil en 2019 (Serra *et al.* 2019). Son implication dans la fermentation du cacao reste à déterminer, mais ce genre pourrait être un marqueur bactérien des fermentations de cacao en Amérique latine. Deux genres appartenant à la famille des *Enterobacteriaceae* sont identifiés. Le genre *Enterobacter* dont les bactéries qui le composent sont principalement retrouvées dans l'intestin humain. Sa présence au début de la fermentation est un indicateur de la mauvaise hygiène des pratiques post-fermentation. Le genre *Pectobacterium*, contient des espèces phytopathogènes, qui produisent des enzymes pectinolytiques qui dégradent la paroi cellulaire de leur hôte.

Nos résultats indiquent qu'un genre de BAA (*Komagataeibacter*) est retrouvé uniquement dans les fermentations au Mexique et un genre de BAL (*Fructobacillus*) dans les fermentations en Guyane et ces genres pourraient être donc spécifiques à une origine géographique. La présence des autres genres retrouvés uniquement dans un pays est la conséquence des différentes pratiques locales. Il a été démontré que le traitement post-récolte du cacao surpasse l'impact de la variété de cacao sur la diversité des communautés microbiennes (Papalexandratou *et al.* 2019). Les pratiques de mise en fermentation entre les trois pays ont été différentes : caisse en bois, caisse en polystyrène et fermentation en tas sur des feuilles de bananiers et pourraient être à l'origine de ces différences. J'ai pu observer lors de la fermentation que j'ai menée au Mexique, que le choix du matériau pour recouvrir les fèves lors de la fermentation avait un impact sur la fermentation. Nous avons utilisé deux sacs en propylène et une en toile de jute pour recouvrir les fèves en fermentation dans les caisses. La caisse de fermentation recouverte avec la toile de jute s'est déroulée plus rapidement que les autres ; Elle s'est terminée un jour plus tôt et un développement de moisissures a été observé. La présence du genre *Bacillus*, à de plus fortes proportions que dans les autres caisses peut être responsable de la fermentation plus rapide. En effet, les bactéries du genre *Bacillus* peuvent participer à la dépectinisation de la pulpe, grâce à l'action d'une enzyme pectinolytique (Ouattara *et al.* 2008). La pulpe est ainsi liquéfiée, permettant une entrée d'air dans les fèves plus importante ; Ceci favorise le développement des bactéries acétiques.

La période de récolte des cabosses a également un effet sur les communautés bactériennes. Un genre de BAA est retrouvé uniquement lors de la grande récolte : *Gluconobacter* ; un autre genre de BAA, *Komagataeibacter* est retrouvé en de plus fortes proportions lors de la grande

récolte. La faible quantité de la pulpe lors de la petite récolte combinée à la présence du genre *Tatumella* au premier jour de fermentation pourrait être à l'origine de la courte durée de la fermentation (4 jours au lieu de 5 jours en février 2018). En effet, les bactéries du genre *Tatumella* peuvent participer à la dépectinisation de la pulpe, permettant une entrée d'air dans les fèves plus importante. Cette entrée d'air a peut-être favorisé le genre *Acetobacter*. Sa population aurait pris le dessus sur celle de *Komagataibacter* et *Gluconobacter*. Cependant, nous ne disposons pas des prélèvements de fèves de cacao aux deux premiers jours de fermentation pour la grande récolte, donc il est difficile de confirmer si la rapide fermentation lors de la petite récolte est due à la combinaison de la faible quantité de la pulpe et de la présence du genre *Tatumella*, ou seulement de la faible quantité de pulpe.

En ce qui concerne l'origine du pool de bactéries impliquées dans les fermentations du cacao au Mexique, l'analyse par metabarcoding nous a permis d'éliminer l'environnement de la plantation comme source principale du pool de bactéries. Ce sont les surfaces liées à l'environnement de la fermentation qui inoculent majoritairement les fèves. Nos résultats indiquent que pratiquement toutes les surfaces (et les mains) qui sont en contact avec les fèves de cacao possèdent les genres *Acetobacter*, *Lactobacillus* et *Bacillus*. La machette et la caisse sont les seules surfaces à ne pas présenter le genre *Acetobacter* dans leurs communautés bactériennes. Cependant, différentes surfaces sont à l'origine des genres trouvés en début de fermentation. Le genre *Tatumella* est présent dans les communautés bactériennes des 3 prélèvements du sol où étaient disposées les fèves. Le sol est donc la principale source de ce genre pour les fermentations de fèves de cacao. Concernant le genre *Weissella*, qui est une BAL connue pour son rôle dans les fermentations de fèves de cacao, il est retrouvé au niveau des communautés bactériennes du sol et au aussi du seau qui a servi pour peser les fèves avant la mise en fermentation. Dans les caisses 1 et 2, le genre *Gluconobacter* est présent, il est retrouvé en de fortes proportions au niveau des deux sacs de polypropylène qui recouvraient les fèves en fermentation et de la pale (qui sert pour le brassage des fèves) ; mais également en de moindres proportions au niveau du seau pour peser les fèves. Enfin, le genre *Acinetobacter*, qui est décelé dans une des caisses de fermentation mais à une faible proportion est également décelé dans les caisses de fermentation, cependant, sa faible proportion indique que les bactéries n'ont pas pu se développer au sein des fèves de cacao. Nous avons pu déterminer que les surfaces ont donc

une contribution différente pour l'inoculation des fèves de cacao. En œnologie, des études ont démontré que certains microorganismes, notamment *Saccharomyces cerevisiae* et des levures du genre *Hanseniaspora*, étaient déjà présents à la surface des équipements dans les caves avant la récolte (Bokulich *et al.* 2013; Grangeteau *et al.* 2015). Les surfaces des mains et des outils agissent comme des réservoirs pour le transfert bactérien vers les fermentations. Les espèces impliquées dans les fermentations sont sélectionnées depuis des années, il est difficile de dire quelle est l'origine exacte de ces espèces.

Comme il a été également démontré dans la littérature, nous avons donc confirmé que les genres de BAA et BAL sont majoritaires dans les fermentations. Avec les méthodes utilisées au cours de la thèse, la PCR-DGGE et le NGS, il n'a pas été possible d'identifier les OTUs au niveau de l'espèce, c'est le principal inconvénient avec ces méthodes; Les séquences générées sont trop courtes pour permettre l'identification des OTUs jusqu'au niveau de l'espèce. C'est pourquoi nous avons couplés ces résultats obtenus avec les méthodes indépendantes de la culture aux méthodes dépendantes de la culture. Au cours de la thèse nous avons produit une collection de souches de BAL et de BAA.

Nous avons voulu déterminer la diversité intra spécifique des souches majoritaires de BAL et BAA selon les 3 origines géographiques. Concernant la réalisation de la collection de souches de BAA, 149 souches ont été obtenues avec 68 souches (48 en 2016, 15 en 2018) isolées de fermentations en Côte d'Ivoire, 59 souches (14 en Février et 31 en octobre 2018) de fermentations au Mexique et 22 souches de Guyane. Cette différence en nombre d'isolats selon les pays peut s'expliquer par le fait que lors de la campagne réalisée en Côte d'Ivoire en 2016 (par une doctorante de l'UMR QualiSud, Christelle Kouame) et celle réalisée par moi-même au Mexique en Octobre 2018, les isolements ont été réalisés directement après le prélèvement. Pour les autres campagnes, l'isolement a été réalisé après que les fèves aient été conservées dans du glycérol à -20°C ; Cette étape de conservation a pu avoir un impact sur la survie des bactéries. Sur les 149 souches de BAA, une majorité de souches d'*Acetobacter pasteurianus* (108/149) et d'*Acetobacter tropicalis* (16/148) ont été isolées. Cette importante proportion d'*A. pasteurianus* peut être due à des biais de manipulations, impliquant un isolement sélectif de cette espèce. En effet, avant l'isolement sur le milieu spécifique Kneifel, une étape d'enrichissement en milieu liquide est réalisée dans le milieu CPSM-AAB, milieu qui simule les conditions de la pulpe au deuxième jour de fermentation. Les souches d'*A.*

*pasteurianus* qui étaient majoritaires à la surface des fèves de cacao pourraient avoir pris le dessus en masquant les espèces minoritaires. De plus, l'utilisation du milieu Kneifel peut aussi avoir favorisé les *A. pasteurianus* au détriment des espèces minoritaires. L'isolement, la cultivation et le maintien en culture pure est difficile à obtenir pour les souches de BAA (Teyssier et Hamdouche 2016), la combinaison de différents milieux de culture pourrait être adéquat afin de parvenir à isoler les espèces minoritaires également. Pour l'isolement des BAL, 63 souches ont été obtenues, avec une grande majorité de *Lactobacillus plantarum*, 55 isolats. Deux biais ont également été introduits : le premier est que seuls les bacilles à Gram positif ont été sélectionnés et le second, l'utilisation d'amorces spécifiques aux genres *Weissella*, *Lactobacillus*, *Pediococcus* et *Leuconostoc* lors de la pré-identification sous estimant la diversité réelle des BAL.

L'analyse de la diversité intraspécifique des souches d'*A. tropicalis* montre un effet de l'origine géographique. Les souches Guyanaises sont regroupées à part des autres souches, et les souches Mexicaines sont regroupées en paire. En revanche, cet effet n'est pas visible pour les souches d'*A. pasteurianus*, les 107 souches sont regroupées en 6 groupes, chaque groupe est composé de souches en provenance de 2 ou 3 pays. Un pourcentage de similarité de 60 % est observé pour l'ensemble des souches. Illegheems *et al.* (2013), ont séquencé le génome complet d'une souche d'*A. pasteurianus* isolée lors de fermentation de cacao (*A. pasteurianus* 386B) et l'ont comparé avec le génome de 5 autres souches d'*A. pasteurianus* dont la souche type LMG01262<sup>T</sup>, également introduite lors de notre étude. Ils ont démontré que 68 % des protéines prédites est partagé entre eux formant le « core » génome de l'espèce. Une équipe a comparé la séquence du chromosome et du plasmide de 11 souches d'*A. pasteurianus* isolées de différentes matrices (cacao, vinaigre, etc.). Ils ont montré que les chromosomes étaient conservés au cours de l'évolution, partageant un haut degré d'homologie (Wang *et al.* 2015). Les homologies élevées entre les chromosomes des souches d'*A. pasteurianus* expliquent les 60 % de similarité entre les souches de notre étude.

Nous avons étudié le comportement des souches d'*A. pasteurianus* dans un milieu simulant les conditions de la pulpe de cacao à 2 jours de fermentations afin de comparer leurs caractéristiques physiologiques comme leur taux de croissance spécifique, leur consommation spécifique d'éthanol et d'acide lactique, leur production spécifique d'acide acétique et leurs préférences relatives pour les sources de carbone.

Nos résultats ont montré une importante similarité entre le comportement de toutes les souches. Elles ont été séparées en deux grands groupes selon leur affinité à l'acide lactique et la suroxydation de l'acide acétique.

L'acide acétique est produit par la réaction d'oxydation de l'éthanol produit par les levures lors de la fermentation des fèves. C'est un métabolite clé pour la qualité organoleptique finale du cacao marchand. En effet, cet acide va pénétrer dans les cotylédons, tuer l'embryon et initier la production de précurseurs d'arômes nécessaires à la qualité organoleptique du cacao. Lors de notre analyse, certaines souches ont probablement sur-oxydé l'acide acétique en CO<sub>2</sub> et H<sub>2</sub>O produit à partir de l'éthanol. Cette suroxydation est reportée dans la littérature comme étant provoquée par une activité accrue des enzymes du cycle de l'acide tricarboxylique (TCA) et de l'acétyl-CoA synthétase pour un apport énergétique supplémentaire (Saeki *et al.* 1999; Sakurai *et al.* 2011; Adler *et al.* 2014) et est impliqué dans le mécanisme de résistance des *A. pasteurianus* à l'acide acétique. L'éthanol est le premier apport énergétique pour les bactéries acétiques, la production d'acide acétique peut causer un stress pour les bactéries. La suroxydation d'une petite fraction d'acide acétique via le cycle TCA peut être un apport d'énergie mais également permettrait de réduire donc le stress produit (Yang *et al.* 2019). Cependant, cette suroxydation de l'acide acétique semble être plus importante pour certaines souches de notre collection, c'est une caractéristique que l'on veut éviter dans le cadre de la sélection de souches pour la production d'un starter de culture.

La plupart des souches d'*A. pasteurianus* analysées dans notre étude ont préféré l'éthanol à l'acide lactique, avec des ratios de consommation d'acide lactique inférieurs à 1 confirmant les résultats des études précédentes. Dans la fermentation du cacao, les souches d'*A. pasteurianus* utilisent l'acide lactique et l'éthanol pour la production d'acide acétique (Camu *et al.* 2007, 2008; Lefeber *et al.* 2010; Moens *et al.* 2014) et ont montré un métabolisme séparé pendant la co-consommation de ces deux substrats. Selon Adler *et al.* (2014), la production d'acide acétique provient principalement de l'éthanol et très peu de l'acide lactique. L'acide lactique est quant à lui utilisé principalement pour la production d'acétoïne et de biomasse. Cependant, on a pu démontrer que certaines souches ont consommé plus d'acide lactique que les autres. Cette consommation est souhaitée car l'acide lactique est considéré comme un métabolite indésirable dans le processus de fermentation car il procure des notes



désagréables au cacao final (Holm *et al.* 1993 ; Camu *et al.* 2007 ; Lefeber *et al.* 2012; Barel 2013).

Ces deux caractéristiques (suroxydation de l'acide acétique et affinité à l'acide lactique) nous permettent de sélectionner des souches adéquates en vue de les intégrer dans l'élaboration d'un starter de culture. Trois souches présentant la plus forte affinité pour l'acide lactique, ont démontré avoir un rendement en biomasse plus élevé couplé à une production d'acide acétique (par mole de biomasse produite) également plus élevée. Ces trois souches d'*A. pasteurianus* semblent être de bonnes candidates pour le starter de culture, deux ont été isolées au Mexique : MX81 et MX90 (une dans chaque campagne) et une autre souche isolée en Côte d'Ivoire (2016) : CI19.

Cependant, nos données indiquent que le regroupement de l'information phénotypique n'est pas lié au regroupement génomique des souches d'*A. pasteurianus*. Les analyses montrent des similarités génomiques et phénotypiques élevées entre les souches, mais avec un regroupement différent. La plupart des souches qui présentent une similarité élevée de leur profils génomiques (> 90 %) ne sont pas similaires d'un point de vue phénotypique. Suggérant qu'une souche peut se comporter différemment dans un même milieu et/ou que la méthode employée pourrait ne pas être suffisante pour la discrimination de la variabilité intra spécifique entre les souches notamment lorsqu'il faut faire une sélection en fonction de leur potentiel fermentaire.

Les souches d'*Acetobacter* sont reportées dans la littérature comme étant hyper mutables, elles sont capables de modifier leurs génomes pour s'adapter à leur environnement (Takemura *et al.* 1991; Beppu *et al.* 1993; Sokollek *et al.* 1998; Steiner et Sauer 2011; Azuma *et al.* 2009; Wang *et al.* 2015). Le génome complet d'une souche d'*A. pasteurianus* a été séquencé par Azuma *et al.* 2009, une quantité importante de gènes qui encodent des transposases (9%) a été identifiée ; Le transposon IS1380 étant le plus abondant dans le génome. La présence de répétition en tandem hyper mutable (HTR, hyper mutable tandem repeat) a également été identifiée. Selon Beppu (1993) et Matsutani *et al.* (2013), des modifications phénotypiques au niveau de l'oxydation de l'éthanol peuvent se produire par l'insertion d'un élément de séquence codant la transposase de ce type.

Les souches d'*Acetobacter* ont un taux élevé de mutation (Beppu *et al.* 1993) et semblent être sensible à l'appauvrissement nutritionnel, ce qui peut augmenter le taux de mutation (Azuma *et al.* 2009). Lors de nos expériences, les souches d'*A. pasteurianus* ont été conservées dans un bouillon de culture additionné de glycérol à -80°C. Le nombre de repiquage effectué à partir du premier isolement a sûrement dû causer un stress chez certaines souches, entraînant des mutations ; ce qui pourrait expliquer les différences obtenues entre nos différentes souches.

Notre travail représente une des premières études sur l'analyse de la diversité intra spécifique des souches de bactéries acétiques impliquées dans la fermentation du cacao. Nos résultats ont mis en évidence que les souches présentaient des différences d'affinité pour l'acide lactique ceci indépendamment de l'origine géographique. Une haute efficacité de conversion de l'éthanol en acide acétique et une forte affinité pour l'acide lactique pourraient être utilisées comme critères de sélection pour la sélection des starters BAA.

## Perspectives

A l'issue de ce travail, des perspectives de recherche ont été dégagées :

- Comparaison des voies métaboliques des souches d'*A. pasteurianus* à l'aide de l'étude de l'expression des gènes impliqués (transcriptomique), notamment le métabolisme de l'éthanol et de l'acide lactique.
- Réaliser des fermentations de cacao avec les souches sélectionnées (groupe métabolique II, C119, MX81 et MX90 ) pour déterminer leurs contributions sur la production d'arôme.
- Etudier la diversité intra spécifique des souches d'*Acetobacter pasteurianus* isolées d'autres origines géographiques ; notamment dans les pays en Asie, en Amérique du Sud et Afrique centrale.
- Etudier les interactions entre les groupes microbiens retrouvés dans la fermentation du cacao (levures, bactéries lactiques, bactéries acétiques et les *Bacillus*).

Listes des communications :

Communications écrites (posters) :

**- Metabarcoding analysis of environmental samples collected during cocoa beans fermentation in Mexico**

L. Farrera, B. Guibert, F. Constancias, S. Hernández Medina, Y. Hamdouche, M. Suárez Quiroz, D. Montet and C. Teyssier  
FEMS 2019 – 7-11 juillet 2019 – Glasgow, Ecosse.

**- Analysis of bacterial communities involved in cocoa beans fermentation in Mexico and French Guiana**

L. Farrera, B. Guibert, I. Piro-Metayer, D. Montet, C. Teyssier  
MicrobiOccitanie 2019 – 18-20 février 2019- Montpellier, France.

**- Genetic diversity of bacterial strains associated to cocoa beans fermentation in three geographical areas**

L. Farrera, B. Guibert, I. Piro-Metayer, Y. Hamdouche, N. Durand, D. Montet, C. Teyssier  
Fourth International Conference on Cocoa Coffee and Tea – 25-28 June 2017- Turin, Italy

Publication :

Soumise - **Farrera, L.**, Colas de la Noue, A. Strub, C., Guibert, B., Kouame, C., Grabulos, J., Montet, D., et Teyssier, C., (2019). Role and intraspecific diversity of Acetic Acid Bacteria in cocoa beans fermentation from three different geographic origins.

## Références bibliographiques

- Adler, P., Frey, L.J., Berger, A., Bolten, C.J., Hansen, C.E., and Wittmann, C. 2014. « The Key to Acetate: Metabolic Fluxes of Acetic Acid Bacteria under Cocoa Pulp Fermentation-Simulating Conditions ». *Applied and Environmental Microbiology*.
- Agyirifo, S., Wamalwa, M., Otwe, E., Galyuon, I., Runo, S., Takrama, J., and Ngeranwa, J. 2019. « Metagenomics Analysis of Cocoa Bean Fermentation Microbiome Identifying Species Diversity and Putative Functional Capabilities ». *Heliyon*.
- Altschul, S. 1997 « Gapped BLAST and PSI-BLAST: a new generation of protein database search programs ». *Nucleic Acids Research*.
- Amann, R.L., Ludwig, W., and Schleifer, K.H. 1995. « Phylogenetic Identification and in Situ Detection of Individual Microbial Cells without Cultivation ». *Microbiol.*
- Ampe, F., Ben Omar, N., Moizan, C., Wachter, C., and Guyot, J. P. 1999. « Polyphasic study of the spatial distribution of microorganisms in Mexican pozol, a fermented maize dough, demonstrated the need for cultivation-independent methods to investigate traditional fermentations ». *Applied and Environmental Microbiology*.
- Arana-Sánchez, A., L. E. Segura-García, M. Kirchmayr, I. Orozco-Ávila, E. Lugo-Cervantes, and A. Gschaedler-Mathis. 2015. « Identification of Predominant Yeasts Associated with Artisan Mexican Cocoa Fermentations Using Culture-Dependent and Culture-Independent Approaches ». *World Journal of Microbiology and Biotechnology*.
- Ardhana, M. 2003. « The Microbial Ecology of Cocoa Bean Fermentations in Indonesia ». *International Journal of Food Microbiology*.
- Ardui, S., Ameer, A., Vermeesch, J.R., Hestand, M.S. 2018. « Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics ». *Nucleic Acids Research*.
- Attchelouwa, Constant K., Florent K. N'guessan, Francine M. D. Aké, and Marcellin K. Djè. 2018. « Molecular Identification of Yeast, Lactic and Acetic Acid Bacteria Species during Spoilage of Tchapano, a Traditional Sorghum Beer from Côte d'Ivoire ». *World Journal of Microbiology and Biotechnology*.
- Azuma, Y., Hosoyama, A., Matsutani, M., Furuya, N., Horikawa, H., Harada, T., Hirakawa, H., Kuhara, S., Matsushita, K., Fujita, N., and Shirai, M. 2009. « Whole-Genome Analyses Reveal Genetic Instability of *Acetobacter Pasteurianus* ». *Nucleic Acids Research*.
- Barel, M. 1998. « Première transformation du cacao. Formation de l'arôme cacao. In J. Pontillon. Cacao, chocolat, production, utilisation, caractéristiques ». Paris: Techniques and documentation, Lavoisier.
- Barel, M. 2013. *Qualité du cacao: L'impact du traitement post-récolte*. Quae éditions.
- Belkum, A.V and Scherer, S. 1998. « Short-Sequence DNA Repeats in Prokaryotic Genomes ». *Microbiol. Mol. Biol. Rev.*
- Beppu, T. 1994. « Genetic Organization Of *Acetobacter* for Acetic Acid Fermentation ». *Antonie van Leeuwenhoek*.

- Biehl, B., Passern, D., and Sagemann, W. 1982. « Effect of Acetic Acid on Subcellular Structures of Cocoa Bean Cotyledons ». *Journal of the Science of Food and Agriculture*.
- Bonnin, R. A., T. Naas, et Dortet, L. 2017. « Impact du séquençage d'ADN à haut débit sur la surveillance des épidémies de bactéries multi-résistantes aux antibiotiques ». *Feuillets de Biologie*.
- Bortolini, C., Patrone, V., Puglisi, E., and Morelli, L. 2016. « Detailed Analyses of the Bacterial Populations in Processed Cocoa Beans of Different Geographic Origin, Subject to Varied Fermentation Conditions ». *International Journal of Food Microbiology*.
- Bukin, Yu. S., Yu. P. Galachyants, I. V. Morozov, S. V. Bukin, A. S. Zakharenko, and T. I. Zemskaya. 2019. « The Effect of 16S rRNA Region Choice on Bacterial Community Metabarcoding Results ». *Scientific Data*.
- Camu, N., T. De Winter, K. Verbrugghe, I. Cleenwerck, P. Vandamme, J. S. Takrama, M. Vancanneyt, and L. De Vuyst. 2007. « Dynamics and Biodiversity of Populations of Lactic Acid Bacteria and Acetic Acid Bacteria Involved in Spontaneous Heap Fermentation of Cocoa Beans in Ghana ». *Applied and Environmental Microbiology*.
- Camu, N., A. Gonzalez, T. De Winter, A. Van Schoor, K. De Bruyne, P. Vandamme, J. S. Takrama, S. K. Addo, and L. De Vuyst. 2008. « Influence of Turning and Environmental Contamination on the Dynamics of Populations of Lactic Acid and Acetic Acid Bacteria Involved in Spontaneous Cocoa Bean Heap Fermentation in Ghana ». *Applied and Environmental Microbiology*.
- Cleenwerck, I., N. Camu, K. Engelbeen, T. De Winter, K. Vandemeulebroecke, P. De Vos, and L. De Vuyst. 2007. « *Acetobacter ghanensis* sp. nov., a Novel Acetic Acid Bacterium Isolated from Traditional Heap Fermentations of Ghanaian Cocoa Beans ». *International Journal of Systematic And Evolutionary Microbiology*.
- Cleenwerck, I., A. Gonzalez, N. Camu, K. Engelbeen, P. De Vos, and L. De Vuyst. 2008. « *Acetobacter fabarum* sp. nov., an Acetic Acid Bacterium from a Ghanaian Cocoa Bean Heap Fermentation ». *International Journal of Systematic And Evolutionary Microbiology*.
- Cleenwerck, I., and De Vos, P. 2008. « Polyphasic Taxonomy of Acetic Acid Bacteria: An Overview of the Currently Applied Methodology ». *International Journal of Food Microbiology*.
- Cleenwerck, I., P. De Vos, and L. De Vuyst. 2010. « Phylogeny and Differentiation of Species of the Genus *Gluconacetobacter* and Related Taxa Based on Multilocus Sequence Analyses of Housekeeping Genes and Reclassification of *Acetobacter xylinus* subsp. *Sucrofermentans* as *Gluconacetobacter sucrofermentans* (Toyosaki and al. 1996) sp. nov., comb. nov. ». *International Journal of Systematic And Evolutionary Microbiology*.
- Crafack, M., Mikkelsen, MB., Saerens, S., Knudsen, M., Blennow, A., Lowor, S., Takrama, J., Swiegers, JH., Petersen, GB., Heimdal, H., and Nielsen, DS. 2013. « Influencing Cocoa Flavour Using *Pichia kluyveri* and *Kluyveromyces marxianus* in a Defined Mixed Starter Culture for Cocoa Fermentation ». *International Journal of Food Microbiology*.
- Dan, T., Liu, W., Song, Y., Xu, H., Menghe, B., Zhang, H., and Sun, Z. 2015. « The Evolution and Population Structure of *Lactobacillus fermentum* from Different Naturally Fermented Products as Determined by Multilocus Sequence Typing (MLST) ». *BMC Microbiology*.
- Daniel, HM., Vrancken, G., Takrama, JF., Camu, N., De Vos, P., and De Vuyst, L. 2009. « Yeast Diversity of Ghanaian Cocoa Bean Heap Fermentations ». *FEMS Yeast Research*.

- De Bruyne, K., N. Camu, L. De Vuyst, and P. Vandamme. 2009. « *Lactobacillus fabifermentans* sp. nov. and *Lactobacillus cacaonum* sp. nov., Isolated from Ghanaian Cocoa Fermentations ». *International Journal of Systematic And Evolutionary Microbiology*.
- De Roos, J., and De Vuyst, L. 2018. « Acetic Acid Bacteria in Fermented Foods and Beverages ». *Current Opinion in Biotechnology*.
- De Vuyst, L., Camu, N., Winter, T., Vandemeulebroecke, K., Vincent Van de Perre, Marc Vancanneyt, Paul De Vos, and Ilse Cleenwerck. 2008. « Validation of the (GTG)<sub>5</sub>-Rep-PCR Fingerprinting Technique for Rapid Classification and Identification of Acetic Acid Bacteria, with a Focus on Isolates from Ghanaian Fermented Cocoa Beans ». *International Journal of Food Microbiology*.
- De Vuyst, L. and S. Weckx. 2016. « The Cocoa Bean Fermentation Process: From Ecosystem Analysis to Starter Culture Development ». *Journal of Applied Microbiology*.
- Djedjro, C., Akmel, N., Assidjo, E., and Yao, B. 2008. « Effets des dispositifs de séchage à l'air libre sur la qualité des fèves de cacao bord champ ».
- Dupont C, Michon A-L, Jumas-Bilak, E. 2015. « Intrapatient diversity of *Achromobacter* spp. involved in chronic colonization of Cystic Fibrosis airways ». *Infection, Genetics and Evolution*
- Fernández-Pérez, R., Torres, C., Sanz, S., and Ruiz-Larrea, F. 2010. « Strain Typing of Acetic Acid Bacteria Responsible for Vinegar Production by the Submerged Elaboration Method ». *Food Microbiology*.
- Figuroa-Hernández, C., Mota-Gutierrez, J., Ferrocino, I., Hernández-Estrada, ZJ., González-Ríos, O., Cocolin, L., and Suárez-Quiroz, ML. 2019. « The Challenges and Perspectives of the Selection of Starter Cultures for Fermented Cocoa Beans ». *International Journal of Food Microbiology*.
- Gallais, A., A. Charrier, J. C. Courduroux, L. Decourtye, and A. B. Eskes. 1991. « (*Theobroma cacao* L.) Influence sur le nombre de graines par fruit ».
- Garcia-Armisen, T., Papalexandratou, Z., Hendryckx, H., Camu, N., Vrancken, G. De Vuyst, L., and Cornelis, P. 2010. « Diversity of the Total Bacterial Community Associated with Ghanaian and Brazilian Cocoa Bean Fermentation Samples as Revealed by a 16S RNAr Gene Clone Library ». *Applied Microbiology and Biotechnology*.
- Gilet, F. 2006. « Le cacao : des Olmèques au XXIème siècle ». Thèse pour le diplôme d'état de docteur en pharmacie. Université de Nantes, Faculté de pharmacie, Nantes.
- Gomes, RJ., de Fatima Borges, M., de Freitas Rosa, M., Hernan Castro-Gómez, RJ., and Aparecida Spinosa, W. 2018. « Acetic Acid Bacteria in the food industry: systematics, characteristics and applications ». *Food Technology and Biotechnology*.
- Hamdouche, Y., Guehi, T., Durand, N., Kedjebo, KBD., Montet, D., and Meile, JC. 2015. « Dynamics of Microbial Ecology during Cocoa Fermentation and Drying: Towards the Identification of Molecular Markers ». *Food Control*.
- Hamdouche, Y., Meile, JC., Lebrun, M., Guehi, T., Boulanger, R., Teyssier, C., and Montet, D. 2019. « Impact of Turning, Pod Storage and Fermentation Time on Microbial Ecology and Volatile Composition of Cocoa Beans ». *Food Research International*.

- Handelsman, J, Rondon, MR., Brady, SF., Clardy, J., and Goodman, RM. 1998. « Molecular Biological Access to the Chemistry of Unknown Soil Microbes: A New Frontier for Natural Products ». *Chemistry & Biology*.
- Heard, G. M., et G. H. Fleet. 1985. « Growth of Natural Yeast Flora during the Fermentation of Inoculated Wines ». *Applied and Environmental Microbiology*.
- Hii, C. L., C. L. Law, M. Cloke, and S. Suzannah. 2009. « Thin Layer Drying Kinetics of Cocoa and Dried Product Quality ». *Biosystems Engineering*.
- Ho, VTT., Zhao, J., and Fleet., G. 2014a. « Yeasts Are Essential for Cocoa Bean Fermentation ». *International Journal of Food Microbiology*.
- Ho, VTT., Zhao, J., and Fleet., G. 2014b. « The Effect of Lactic Acid Bacteria on Cocoa Bean Fermentation ». *International Journal of Food Microbiology*.
- Illegghems, K., De Vuyst, L., Papalexandratou, Z., and Weckx, S. 2012. « Phylogenetic Analysis of a Spontaneous Cocoa Bean Fermentation Metagenome Reveals New Insights Into Its Bacterial and Fungal Community Diversity » édité par M. R. Liles. *PLoS one*.
- Illegghems, K., De Vuyst, L., and Weckx, S. 2013. « Complete genome sequence and comparative analysis of *Acetobacter pasteurianus* 386B, a strain well-adapted to the cocoa bean fermentation ecosystem ». *BMC genomics*.
- Illegghems, K., Weckx, S., and De Vuyst, L. 2015. « Applying Meta-Pathway Analyses through Metagenomics to Identify the Functional Properties of the Major Bacterial Communities of a Single Spontaneous Cocoa Bean Fermentation Process Sample ». *Food Microbiology*.
- Illegghems, K. Pelicaen, R., De Vuyst, L., and Weckx, S. 2016. « Assessment of the Contribution of Cocoa-Derived Strains of *Acetobacter ghanensis* and *Acetobacter senegalensis* to the Cocoa Bean Fermentation Process through a Genomic Approach ». *Food Microbiology*.
- Ivey, ML., and Phister, TG. 2011. « Detection and Identification of Microorganisms in Wine: A Review of Molecular Techniques ». *Journal of Industrial Microbiology & Biotechnology*.
- Jamili, J. 2016. « Diversity and the role of yeast in spontaneous cocoa bean fermentation from Southeast Sulawesi, Indonesia ». *Biodiversitas, Journal of Biological Diversity*.
- Jojima, Y. 2004. « *Saccharibacter floricola* gen. nov., sp. nov., a Novel Osmophilic Acetic Acid Bacterium Isolated from Pollen ». *International Journal of Systematic And Evolutionary Microbiology*.
- Khandelwal, P., Gaspar, F., Crespo Barreto, MT., and Upendra, R. 2015. « Lactic Acid Bacteria General Characteristics, Food Preservation and Health Benefits ». in *Fermented foods: Part I : Biochemistry and Biotechnology*.
- Kirk, JL., Beaudette, LA., Hart, M., Moutoglis, P., Klironomos, JN., Lee, H., and Trevors, JT. 2004. « Methods of Studying Soil Microbial Diversity ». *Journal of Microbiological Methods*.
- Koff, O., Samagaci, L., Goualie, B., and Niamke, S. 2017. « Diversity of Yeasts Involved in Cocoa Fermentation of Six Major Cocoa-Producing Regions in Ivory Coast ». *European Scientific Journal, ESJ*.
- Koné, MK., Guéhi, T., Durand, N., L. Ban-Koffi, Berthiot, L., Fontana Tachon, A., Brou, K., Boulanger, R., and Montet, D. 2016. « Contribution of Predominant Yeasts to the Occurrence of Aroma Compounds during Cocoa Bean Fermentation ». *Food Research International*.

- Kongor, JE., Hinneh, M. Van de Walle, D., Afoakwa, EO., Boeckx, P., and Dewettinck, K. 2016. « Factors Influencing Quality Variation in Cocoa (*Theobroma cacao* L) Bean Flavour Profile — A Review ». *Food Research International*.
- Konstantinidis, KT., Ramette, A., and Tiedje, JM. 2006. « The Bacterial Species Definition in the Genomic Era ». *Philosophical Transactions of the Royal Society B: Biological Sciences*.
- Kostinek, M., Ban-Koffi, L., Ottah-Atikpo, M., Teniola, D., Schillinger, U., Holzapfel, WH., and M. A. P. Franz, C. 2008. « Diversity of Predominant Lactic Acid Bacteria Associated with Cocoa Fermentation in Nigeria ». *Current Microbiology*.
- Lagunes Gálvez, S, Loiseau, G., Paredes, JL., Barel, M., and Guiraud, JP. 2007. « Study on the Microflora and Biochemistry of Cocoa Fermentation in the Dominican Republic ». *International Journal of Food Microbiology*.
- Lefeber, T., Gobert, W., Vrancken, G., Camu, N., and De Vuyst, N. 2011. « Dynamics and Species Diversity of Communities of Lactic Acid Bacteria and Acetic Acid Bacteria during Spontaneous Cocoa Bean Fermentation in Vessels ». *Food Microbiology*.
- Lefeber, T., Papalexandratou, Z., Gobert, W., Camu, N., and De Vuyst, L. 2012. « On-Farm Implementation of a Starter Culture for Improved Cocoa Bean Fermentation and Its Influence on the Flavour of Chocolates Produced Thereof ». *Food Microbiology*.
- Li, W., Raoult, D., and Fournier, PE. 2009. « Bacterial Strain Typing in the Genomic Era ». *FEMS Microbiology Reviews*.
- Li, L., A. Wieme, F. Spitaels, T. Balzarini, O. C. Nunes, C. M. Manaia, A. Van Landschoot, L. De Vuyst, I. Cleenwerck, and P. Vandamme. 2014. « *Acetobacter sicerae* sp. nov., Isolated from Cider and Kefir, and Identification of Species of the Genus *Acetobacter* by DnaK, GroEL and RpoB Sequence Analysis ». *International Journal of Systematic and Evolutionary Microbiology*.
- Lima, LJR., Almeida, H., Rob Nout, MJ., and Zwietering, MH. 2011. « *Theobroma Cacao* L., “The Food of the Gods”: Quality Determinants of Commercial Cocoa Beans, with Particular Reference to the Impact of Fermentation ». *Critical Reviews in Food Science and Nutrition*.
- Lino, T, Suzuki, R., Kosako, Y., Ohkuma, M., Komagata, K., and Uchimura, T. 2012. « *Acetobacter okinawensis* sp. nov., *Acetobacter papayae* sp. nov., and *Acetobacter persicus* sp. nov.; Novel Acetic Acid Bacteria Isolated from Stems of Sugarcane, Fruits, and a Flower in Japan ». *The Journal of General and Applied Microbiology*.
- Longin, C. 2016. « Développement de méthodes permettant la détection et la quantification de microorganismes d’altération du vin: étude de facteurs de développement ». Microbiologie et Parasitologie. Thèse de Doctorat pour obtenir le diplôme d’état de docteur de l’Université de Bourgogne – Franche Comté.
- Ludlow, CL., Cromie, GA., Garmendia-Torres, C., Sirr, A., Hays, M., Field, C., Jeffery, EW., Fay, JC., and Dudley, AM. 2016. « Independent Origins of Yeast Associated with Coffee and Cacao Fermentation ». *Current Biology*.
- Macauley, S., McNeil, B., and Harvey, LM. 2001. « The Genus *Gluconobacter* and Its Applications in Biotechnology ». *Critical Reviews in Biotechnology*.



- Magalhães da Veiga Moreira, I., da Cruz Pedrozo Migue, MG., Lacerda Ramos, C., Ferreira Duarte, W., Efraim, P., and Freitas Schwan, R. 2016. « Influence of Cocoa Hybrids on Volatile Compounds of Fermented Beans, Microbial Diversity during Fermentation and Sensory Characteristics and Acceptance of Chocolates: Cocoa Hybrids Influenced Chocolate Quality ». *Journal of Food Quality*.
- Mahazar, N. H., N. F. Suian, A. S. Meor Hussin, H. Norhayati, M. Mathawan, and Y. Rukayadi. 2015. « *Candida sp.* as a starter culture for cocoa (*Theobroma cacao* L.) beans fermentation. » *International Food Research Journal*.
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. 2014. « Swarm: Robust and Fast Clustering Method for Amplicon-Based Studies ». *PeerJ*.
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. 2015. « Swarm v2: Highly-Scalable and High-Resolution Amplicon Clustering ». *PeerJ*.
- Maiden, M. C. J., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt. 1998. « Multilocus Sequence Typing: A Portable Approach to the Identification of Clones within Populations of Pathogenic Microorganisms ». *Proceedings of the National Academy of Sciences*.
- Marchesi, JR., Sato, T., Weightman, AJ., Martin, TA., Fry, JC., Hiom, SJ., and Wade, WG. 1998. « Design and Evaluation of Useful Bacterium-Specific PCR Primers That Amplify Genes Coding for Bacterial 16S RNAr ». *Applied and Environmental Microbiology*.
- Matsushita, K., Toyama, H., and Adachi, O. 1994. « Respiratory Chains and Bioenergetics of Acetic Acid Bacteria ». *Advances in Microbial Physiology*.
- Mayr, E. 1964. « *Systematics and the Origin of Species: From the Viewpoint of a Zoologist* ». Dover.
- Meersman, E., Steensels, J., Mathawan, M., Wittocx, PJ., Saels, V., Struyf, N., Bernaert, H., Vrancken, G., and Verstrepen, KJ. 2013. « Detailed Analysis of the Microbial Population in Malaysian Spontaneous Cocoa Pulp Fermentations Reveals a Core and Variable Microbiota » *PLoS ONE*.
- Miescher S., Leischtfeld, SF., and Gantenbein-Demarchi, C. 2016. « High-Throughput Identification of the Microbial Biodiversity of Cocoa Bean Fermentation by MALDI-TOF MS ». *Letters in Applied Microbiology*.
- Moens, F., T. Lefeber, and De Vuyst, L. 2014. « Oxidation of Metabolites Highlights the Microbial Interactions and Role of *Acetobacter pasteurianus* during Cocoa Bean Fermentation ». *Applied and Environmental Microbiology*.
- Montel, F. 2018. « Séquençage de l'ADN par nanopores: Résultats et perspectives ». *Médecine/sciences*.
- Moore, W. E. C., E. Stackebrandt, O. Kandler, R. R. Colwell, M. I. Krichevsky, H. G. Truper, R. G. E. Murray, L. G. Wayne, P. A. D. Grimont, D. J. Brenner, M. P. Starr, and L. H. Moore. 1987. « Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics ». *International Journal of Systematic and Evolutionary Microbiology*.
- Moreira, I., da Veiga, M. da Cruz Pedrozo Miguel, MG., Ferreira Duarte, W., Ribeiro Dias, D., and Schwan, RF. 2013. « Microbial Succession and the Dynamics of Metabolites and Sugars during the Fermentation of Three Different Cocoa (*Theobroma cacao* L.) Hybrids ». *Food Research International*.

Moreira, I., de Figueiredo Vilela, L., da Cruz Pedroso, M., Cledir Santos, M., Lima, N., and Freitas Schwan, R. 2017. « Impact of a Microbial Cocktail Used as a Starter Culture on Cocoa Fermentation and Chocolate Flavor ». *Molecules*.

Mota-Gutierrez, J., Botta, C., Ferrocino, I., Giordano, M., Bertolino, M., Dolci, P., Cannoni, M., and Cocolin, L. 2018. « Dynamics and Biodiversity of Bacterial and Yeast Communities during Fermentation of Cocoa Beans ». *Applied and Environmental Microbiology*.

Nielsen, D. S., U. Schillinger, C. M. A. P. Franz, J. Bresciani, W. Amoa-Awua, W. H. Holzapfel, and M. Jakobsen. 2007. « *Lactobacillus ghanensis* sp. nov., a Motile Lactic Acid Bacterium Isolated from Ghanaian Cocoa Fermentations ». *International Journal of Systematic And Evolutionary Microbiology*.

Nielsen, D.S., O. D. Teniola, L. Ban-Koffi, M. Owusu, T. S. Andersson, and W. H. Holzapfel. 2007. « The Microbiology of Ghanaian Cocoa Fermentations Analysed Using Culture-Dependent and Culture-Independent Methods ». *International Journal of Food Microbiology*.

Organisme international de la vigne and du vin (OIV). 2010. « Analyse microbiologique des vins and des moûts : détection, différenciation and dénombrement des microorganismes ». in *Recueil des méthodes internationales d'analyses*.

Ouattara, HG., Koffi, BL., Karou, GT., Sangaré, A., Niamke, SL., and Diopoh, JK. 2008. « Implication of Bacillus Sp. in the Production of Pectinolytic Enzymes during Cocoa Fermentation ». *World Journal of Microbiology and Biotechnology*.

Ouattara, H.G., Reverchon, S., Niamke, S.L., Nasser, W., 2010. « Biochemical properties of pectate lyases produced by three different Bacillus strains isolated from cocoa fermentation and characterization of their cloned genes ». *Appl. Environ. Microbiol.*

Ozturk, G., and Young, GM. 2017. « Food Evolution: The Impact of Society and Science on the Fermentation of Cocoa Beans ». *Comprehensive Reviews in Food Science and Food Safety*.

Papalexandratou, Z., Cleenwerck, I., De Vos, P., and De Vuyst, L., 2009. « (GTG)<sub>5</sub> -PCR Reference Framework for Acetic Acid Bacteria ». *FEMS Microbiology Letters*.

Papalexandratou, Z., Camu, N., Gwen Falony, and Luc De Vuyst. 2011a. « Comparison of the Bacterial Species Diversity of Spontaneous Cocoa Bean Fermentations Carried out at Selected Farms in Ivory Coast and Brazil ». *Food Microbiology*.

Papalexandratou, Z., Gino Vrancken, Katrien De Bruyne, Peter Vandamme, and Luc De Vuyst. 2011b. « Spontaneous Organic Cocoa Bean Box Fermentations in Brazil Are Characterized by a Restricted Species Diversity of Lactic Acid Bacteria and Acetic Acid Bacteria ». *Food Microbiology*.

Papalexandratou, Z., G. Falony, E. Romanens, J. C. Jimenez, F. Amores, H. M. Daniel, and L. De Vuyst. 2011c. « Species Diversity, Community Dynamics, and Metabolite Kinetics of the Microbiota Associated with Traditional Ecuadorian Spontaneous Cocoa Bean Fermentations ». *Applied and Environmental Microbiology*.

Papalexandratou, Z., Timothy Lefeber, Bakhtiar Bahrim, Ong Seng Lee, Heide-Marie Daniel, and Luc De Vuyst. 2013. « *Hanseniaspora opuntiae*, *Saccharomyces cerevisiae*, *Lactobacillus fermentum*, and *Acetobacter pasteurianus* Predominate during Well-Performed Malaysian Cocoa Bean Box Fermentations, Underlining the Importance of These Microbial Species for a Successful Cocoa Bean Fermentation Process ». *Food Microbiology*.

- Papalexandratou, Z., Kaasik, K., Villagra Kauffmann, L., Albert Skorstengaard, Gregoire Bouillon, Julie Leth Espensen, Lars H. Hansen, Rasmus Riemer Jakobsen, Andreas Blennow, Lukasz Krych, Josué L. Castro-Mejía, and Dennis Sandris Nielsen. 2019. « Linking Cocoa Varietals and Microbial Diversity of Nicaraguan Fine Cocoa Bean Fermentations and Their Impact on Final Cocoa Quality Appreciation ». *International Journal of Food Microbiology*.
- Pereira, GVM., da Cruz Pedrozo Miguel, MG., Lacerda Ramos, C., and Freitas Schwan, R. 2012. « Microbiological and Physicochemical Characterization of Small-Scale Cocoa Fermentations and Screening of Yeast and Bacterial Strains to Develop a Defined Starter Culture ». *Applied and Environmental Microbiology*.
- Pereira, GVM., Teixeira Magalhães, K., Gonzaga de Almeida, E., da Silva Coelho, I., and Freitas Schwan, R. 2013a. « Spontaneous Cocoa Bean Fermentation Carried out in a Novel-Design Stainless Steel Tank: Influence on the Dynamics of Microbial Populations and Physical–Chemical Properties ». *International Journal of Food Microbiology*.
- Pereira, GVM., Teixeira Magalhães-Guedes, K., and Freitas Schwan, R. 2013b. « RDNA-Based DGGE Analysis and Electron Microscopic Observation of Cocoa Beans to Monitor Microbial Diversity and Distribution during the Fermentation Process ». *Food Research International*.
- Pereira, GVM., Soccol, T., and Soccol, R. 2016. « Current State of Research on Cocoa and Coffee Fermentations ». *Current Opinion in Food Science*.
- Pereira, GVM., Alvarez, JP., C. Neto, DP., Soccol, VT., Tanobe, VOA., Rogez, H., Góes-Neto, A., and Soccol, CR. 2017. « Great Intraspecies Diversity of *Pichia kudriavzevii* in Cocoa Fermentation Highlights the Importance of Yeast Strain Selection for Flavor Modulation of Cocoa Beans ». *LWT*.
- Pettipher, GL. 1986. « Analysis of Cocoa Pulp and the Formulation of a Standardised Artificial Cocoa Pulp Medium ». *Journal of the Science of Food and Agriculture*.
- Pontillon, J. 1998. « *Cacao and chocolat: production, utilisation, caractéristiques* ». Paris: Lavoisier Tec & Doc.
- Rasschaert, G., K. Houf, H. Imberechts, L. De Zutter, and M. Heyndrickx. 2005. « Comparison of REP-, BOX-, ERIC- and (GTG)<sub>5</sub>-PCR for Molecular Discrimination of *Salmonella* Serotypes ». *Microbiology*.
- Rivas, B. 2006. « Development of a Multilocus Sequence Typing Method for Analysis of *Lactobacillus plantarum* strains ». *Microbiology*.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. 2016. « VSEARCH: A Versatile Open Source Tool for Metagenomics ». *PeerJ*.
- Romanens, E., Näf, R., Lobmaier, T., Pedan, V., Freimüller Leischtfeld, S., Meile, L., and Miescher Schwenninger, S. 2018. « A Lab-Scale Model System for Cocoa Bean Fermentation ». *Applied Microbiology and Biotechnology*.
- Ruggirello, M., Nucera, D., Cannoni, M., Peraino, A., Rosso, F., Fontana, M., Cocolin, L., and Dolci, P. 2019. « Antifungal Activity of Yeasts and Lactic Acid Bacteria Isolated from Cocoa Bean Fermentations ». *Food Research International*.
- Saeki, A., Matsushita, K., Takeno, S., Taniguchi, M., Toyama, H., Theeragool, G., Lotong, N., and Adachi, O. 1999. « Enzymes Responsible for Acetate Oxidation by Acetic Acid Bacteria ». *Bioscience, Biotechnology, and Biochemistry*.

- Sakurai, K., H. Arai, M. Ishii, and Y. Igarashi. 2011. « Transcriptome Response to Different Carbon Sources in *Acetobacter Aceti* ». *Microbiology*.
- Saltini, R., Akkerman, R. and Frosch, S. 2013. « Optimizing Chocolate Production through Traceability: A Review of the Influence of Farming Practices on Cocoa Bean Quality ». *Food Control*.
- Samagaci, L., Ouattara, H., Niamké, S., and Lemaire, M. 2016. « *Pichia kudrazevii* and *Candida nitrativorans* Are the Most Well-Adapted and Relevant Yeast Species Fermenting Cocoa in Agneby-Tiassa, a Local Ivorian Cocoa Producing Region ». *Food Research International*.
- Sato, J., Wakayama, M., and Takagi, K. 2015. « Lactate Dehydrogenase Involved in Lactate Metabolism of *Acetobacter Pasteurianus* ». *Procedia Environmental Sciences*.
- Schwan, RF. 1998. « Cocoa Fermentations Conducted with a Defined Microbial Cocktail Inoculum ». *Appl. Environ. Microbiol.*
- Schwan, RF., and Wheals, AE. 2004. « The Microbiology of Cocoa Fermentation and Its Role in Chocolate Quality ». *Critical Reviews in Food Science and Nutrition*.
- Schwan, RF., and G. H. Fleet. 2015. « *Cocoa and coffee fermentations* ». Boca Raton: CRC Press, Taylor & Francis Group.
- Sengun, IY., and Karabiyikli, S. 2011. « Importance of Acetic Acid Bacteria in Food Industry ». *Food Control*.
- Serra, JL., Gomes Moura; F., de Melo Pereira, GV., Soccol, CR., Rogez, H., and Darnet, S. 2019. « Determination of the Microbial Community in Amazonian Cocoa Bean Fermentation by Illumina-Based Metagenomic Sequencing ». *LWT*.
- Sokollek, S. J., C. Hertel, and W. P. Hammes. 1998. « Description of *Acetobacter oboediens* sp. nov. and *Acetobacter pomorum* sp. nov., Two New Species Isolated from Industrial Vinegar Fermentations ». *International Journal of Systematic Bacteriology*.
- Stackebrandt, E. 2002. « Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology ». *International Journal Of Systematic And Evolutionary Microbiology*.
- Staley, J. T. and A. Konopka. 1985. « Measurement of in Situ Activities of Nonphotosynthetic Microorganisms in Aquatic and Terrestrial Habitats ». *Annual Review of Microbiology*.
- Steiner, P. and U. Sauer. 2001. « Proteins Induced during Adaptation of *Acetobacter aceti* to High Acetate Concentrations ». *Applied and Environmental Microbiology*.
- Takemura, H., S. Horinouchi, and T. Beppu. 1991. « Novel Insertion Sequence IS1380 from *Acetobacter pasteurianus* Is Involved in Loss of Ethanol-Oxidizing Ability. » *Journal of Bacteriology*.
- Tanasupawat, S., Thawai, C., Yukphan, P., Moonmangmee, D., Itoh, T., Adachi, O., and Yamada, Y. 2004. « *Gluconobacter Thailandicus* sp. nov., an Acetic Acid Bacterium in the  $\alpha$ -Proteobacteria ». *The Journal of General and Applied Microbiology*.
- Teyssier, C., and Hamdouche Y. 2016. « Acetic Acid Bacteria ». *Fermented Foods, Part I: Biochemistry and Biotechnology*.

- Thijs, S., Op De Beeck, M. Beckers, B., Truyens, S., Stevens, V., Van Hamme, J.D., Weyens, N., and Vangronsveld, J. 2017. « Comparative evaluation of four bacteria-specific primer pairs for 16S rRNA gene surveys ». *Frontiers in Microbiology*.
- Toit, W. J. Du and I. S. Pretorius. 2002. « The Occurrence, Control and Esoteric Effect of Acetic Acid Bacteria in Winemaking ». *Ann. Microbiol.*
- Trias Mansilla, R. 2008. « Lactic Acid Bacteria as Bioprotective Agents against Foodborne Pathogens and Spoilage Microorganisms in Fresh Fruits and Vegetables». Thèse de Doctorat pour obtenir le diplôme d'état de docteur de l'Université de Gironne et Département de technologie agraire et agroalimentaire du génie chimique.
- Versalovic, J., Koeuth, T., and R. Lupski. 1991. « Distribution of Repetitive DNA Sequences in Eubacteria and Application to Fingerprinting of Bacterial Enomes ». *Nucleic Acids Research*.
- Versalovic, J., M. Schneider, F.J. de Bruijn, and J.R. Lupski. 1994. « Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. » *Methods in Molecular and Cellular Biology*.
- Visintin, S., Alessandria, V., Valette, A., Dolci, P., and Cocolin, L. 2016. « Molecular Identification and Physiological Characterization of Yeasts, Lactic Acid Bacteria and Acetic Acid Bacteria Isolated from Heap and Box Cocoa Bean Fermentations in West Africa ». *International Journal of Food Microbiology*.
- Visintin, S., Ramos, L., Batista, N., Dolci, P., Schwan, F., and Cocolin, L. 2017. « Impact of *Saccharomyces cerevisiae* and *Torulaspora delbrueckii* Starter Cultures on Cocoa Beans Fermentation ». *International Journal of Food Microbiology*.
- Walter, J., C. Hertel, G. W. Tannock, C. M. Lis, K. Munro, and W. P. Hammes. 2001. « Detection of *Lactobacillus*, *Pediococcus*, *Leuconostoc*, and *Weissella* Species in Human Feces by Using Group-Specific PCR Primers and Denaturing Gradient Gel Electrophoresis ». *Applied and Environmental Microbiology*.
- Wang, B., Shao, Y., Chen, T., Chen, W., and Chen, F. 2015. « Global Insights into Acetic Acid Resistance Mechanisms and Genetic Stability of *Acetobacter pasteurianus* Strains by Comparative Genomics ». *Scientific Reports*.
- White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. « Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics ». P. 315-22 in *PCR Protocols: A Guide to Methods and Applications*.
- Woese, C. R. and G. E. Fox. 1977. « Phylogenetic Structure of the Prokaryotic Domain: The Primary Kingdoms ». *Proceedings of the National Academy of Sciences*.
- Yamada, Y., Katsura, K., Kawasaki, H., Widyastuti, Y., Saono, S., Seki, T., Uchimura, T., and Komagata, K. 2000. « *Asaia bogorensis* gen. nov., sp. nov., an Unusual Acetic Acid Bacterium in the  $\alpha$ -Proteobacteria ». *International Journal of Systematic and Evolutionary Microbiology*.
- Yamada, Y. 2016. « Systematics of Acetic Acid Bacteria ». *Acetic Acid Bacteria. Tokyo: Springer Japan*.
- Yang, H., Yu, I., Fu, C., and Chen, F. 2019. « Bacterial acid resistance toward organic weak acid revealed by RNA-Seq transcriptomic analysis in *Acetobacter pasteurianus* ». *Frontiers in Microbiology*.

Yukphan, P.I., Malimas, T., Muramatsu, Y., Potacharoen W., Tanasupawat, S., Nakagawa, Y., Tanticharoen, M., and Yamada, Y. 2011. « *Neokomagataea* gen. nov., with Descriptions of *Neokomagataea Thailandica* sp. nov. and *Neokomagataea Tanensis* sp. nov., Osmotolerant Acetic Acid Bacteria of the  $\alpha$ - *Proteobacteria* ». *Bioscience, Biotechnology, and Biochemistry*.

Zhang, W., Du, P., Zheng, H., Yu, W., Wan, L., and Chen, C. 2014. « Whole-Genome Sequence Comparison as a Method for Improving Bacterial Species Definition ». *The Journal of General and Applied Microbiology*.



# Annexes

Annexe 1: Liste détaillée des souches de BAA isolées au cours de la thèse et les souches types de références.

Code souche	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
CI1*	Côte d'Ivoire	Wagana	déc-16	Caisse	0	1	1366	99,8	99	<i>Acetobacter pasteurianus</i>
CI10	Côte d'Ivoire	Abidjan	mai-18	Tas		4	1351	100	100	<i>Acetobacter pasteurianus</i>
CI100*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	3	1401	98	97	<i>Acetobacter pasteurianus</i>
CI101*	Côte d'Ivoire	Wagana	déc-16	Caisse	4	2	1399	99,6	97	<i>Acetobacter pasteurianus</i>
CI102*	Côte d'Ivoire	Wagana	déc-16	Caisse	4	3	1399	99,7	97	<i>Acetobacter pasteurianus</i>
CI103*	Côte d'Ivoire	Wagana	déc-16	Caisse	5	2	1396	99,5	97,5	<i>Acetobacter pasteurianus</i>
CI104*	Côte d'Ivoire	Wagana	déc-16	Caisse	7	5	1397	99,6	97,5	<i>Acetobacter pasteurianus</i>
CI11	Côte d'Ivoire	Abidjan	mai-18	Tas		4	1308	100	100	<i>Acetobacter pasteurianus</i>
CI12	Côte d'Ivoire	Abidjan	mai-18	Tas		4	1333	100	100	<i>Acetobacter pasteurianus</i>
CI13	Côte d'Ivoire	Abidjan	mai-18	Tas		4	1284	99	100	<i>Acetobacter pasteurianus</i>
CI14	Côte d'Ivoire	Wagana	déc-16	Caisse	1	5	1400	99	99	<i>Acetobacter pasteurianus</i>
<b>CI140*</b>	<b>Côte d'Ivoire</b>	<b>Wagana</b>	<b>déc-16</b>	<b>Caisse</b>	<b>5</b>	<b>1</b>	<b>1368</b>	<b>99,7</b>	<b>99,3</b>	<b><i>Acetobacter fabarum</i></b>
CI141*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	1	1367	99,9	99,9	<i>Acetobacter ghanensis</i>
CI142*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	2	1398	99,1	99,3	<i>Acetobacter ghanensis</i>
CI143*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	3	1362	99,7	100	<i>Acetobacter ghanensis</i>
CI144*	Côte d'Ivoire	Wagana	déc-16	Caisse	4	2	1364	99,9	99,9	<i>Acetobacter ghanensis</i>
CI145*	Côte d'Ivoire	Wagana	déc-16	Caisse	5	2	1354	99,9	100	<i>Acetobacter ghanensis</i>
CI146	Côte d'Ivoire	Abidjan	mai-18	Tas		2	1404	99,7	99,8	<i>Acetobacter ghanensis</i>
CI147	Côte d'Ivoire	Abidjan	mai-18	Tas		4	1389	99,6	99,1	<i>Acetobacter ghanensis</i>
CI15	Côte d'Ivoire	Abidjan	mai-18	Tas		5	1305	99,9	100	<i>Acetobacter pasteurianus</i>
CI16	Côte d'Ivoire	Abidjan	mai-18	Tas		6	1331	100	100	<i>Acetobacter pasteurianus</i>
CI17	Côte d'Ivoire	Abidjan	mai-18	Tas		6	1297	99,5	100	<i>Acetobacter pasteurianus</i>
CI18	Côte d'Ivoire	Abidjan	mai-18	Tas		6	1346	99,3	98	<i>Acetobacter pasteurianus</i>
CI19*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	5	1396	99,3	98	<i>Acetobacter pasteurianus</i>



CI2	Côte d'Ivoire	Abidjan	mai-18	Tas		1	1204	100	97	<i>Acetobacter pasteurianus</i>
CI20*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	6	1360	100	100	<i>Acetobacter pasteurianus</i>
CI21*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	0	1395	99	98	<i>Acetobacter pasteurianus</i>
CI22*	Côte d'Ivoire	Wagana	déc-16	Caisse	2	2	1363	100	100	<i>Acetobacter pasteurianus</i>
CI23*	Côte d'Ivoire	Wagana	déc-16	Caisse	2	5	1399	98	98	<i>Acetobacter pasteurianus</i>
CI24*	Côte d'Ivoire	Wagana	déc-16	Caisse	2	6	1401	98,8	99	<i>Acetobacter pasteurianus</i>
CI25	Côte d'Ivoire	Abidjan	mai-18	Tas		3	1349	100	100	<i>Acetobacter fabarum</i>
CI26*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	1	1398	98,3	98	<i>Acetobacter pasteurianus</i>
CI27*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	1	1397	97,3	98	<i>Acetobacter pasteurianus</i>
CI28*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	3	1399	99,3	98	<i>Acetobacter pasteurianus</i>
CI29*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	4	1399	98,9	98	<i>Acetobacter pasteurianus</i>
CI3	Côte d'Ivoire	Abidjan	mai-18	Tas		1	1345	99	100	<i>Acetobacter pasteurianus</i>
CI30*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	5	1400	99	99	<i>Acetobacter pasteurianus</i>
CI31*	Côte d'Ivoire	Wagana	déc-16	Caisse	4	1	1398	99,4	98	<i>Acetobacter pasteurianus</i>
CI32*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	0	1401	99,7	98	<i>Acetobacter pasteurianus</i>
CI33*	Côte d'Ivoire	Wagana	déc-16	Caisse	5	3	1398	99,7	97	<i>Acetobacter pasteurianus</i>

Code souche	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
GF148	Guyane	Combi	déc-17	Caisse		1	1404	99,9	99,6	<i>Acetobacter orientalis</i>
GF149	Guyane	Combi	déc-17	Caisse		2	1395	99	99	<i>Acetobacter orientalis</i>
MX134	Mexique	Comalcalco	oct-18	Caisse	1	1	1368	99,9	96	<i>Acetobacter malorum</i> ou <i>Cerevisiae</i>
MX135	Mexique	Comalcalco	oct-18	Caisse	2	1	1354	99,9	99,3	<i>Acetobacter malorum</i> ou <i>Cerevisiae</i>
MX136	Mexique	Comalcalco	oct-18	Caisse	2	1	1347	99,9	99,8	<i>Acetobacter malorum</i> ou <i>Cerevisiae</i>
MX137	Mexique	Comalcalco	oct-18	Caisse	2	1	1363	99,9	99,3	<i>Acetobacter malorum</i> ou <i>Cerevisiae</i>
MX138	Mexique	Comalcalco	oct-18	Caisse	2	1	1401	99,9	99	<i>Acetobacter malorum</i> ou <i>Cerevisiae</i>

Code souche	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
MX139	Mexique	Comalcalco	oct-18	Caisse	2	1	1397	99,7	99,2	<i>Acetobacter malorum ou Cerevisiae</i>
CI34*	Côte d'Ivoire	Wagana	déc-16	Caisse	6	5	1395	98,9	99	<i>Acetobacter pasteurianus</i>
CI35*	Côte d'Ivoire	Wagana	déc-16	Caisse	6	6	1398	98,9	98	<i>Acetobacter pasteurianus</i>
CI36*	Côte d'Ivoire	Wagana	déc-16	Caisse	7	5	1398	99,3	98	<i>Acetobacter pasteurianus</i>
CI37*	Côte d'Ivoire	Wagana	déc-16	Caisse	7	6	1399	99,3	98	<i>Acetobacter pasteurianus</i>
CI38*	Côte d'Ivoire	Wagana	déc-16	Caisse	7	6	1398	99	98	<i>Acetobacter pasteurianus</i>
CI39*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	1	1399	99,6	98	<i>Acetobacter pasteurianus</i>
CI4	Côte d'Ivoire	Abidjan	mai-18	Tas		1	1333	99	10	<i>Acetobacter pasteurianus</i>
CI40*	Côte d'Ivoire	Wagana	déc-16	Caisse	0	0	1398	99,5	98	<i>Acetobacter pasteurianus</i>
CI5	Côte d'Ivoire	Abidjan	mai-18	Tas		2	1346	100	100	<i>Acetobacter pasteurianus</i>
CI6	Côte d'Ivoire	Abidjan	mai-18	Tas		3	1326	99,9	100	<i>Acetobacter pasteurianus</i>
CI7	Côte d'Ivoire	Wagana	déc-16	Caisse	1	2	1401	99,7	98	<i>Acetobacter pasteurianus</i>
CI8	Côte d'Ivoire	Abidjan	mai-18	Tas		3	1295	99,7	100	<i>Acetobacter pasteurianus</i>
CI9	Côte d'Ivoire	Abidjan	mai-18	Tas		3	1260	100	100	<i>Acetobacter pasteurianus</i>
CI94*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	5	1349	100	98	<i>Acetobacter pasteurianus</i>
CI95*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	6	1344	100	98	<i>Acetobacter pasteurianus</i>
CI96*	Côte d'Ivoire	Wagana	déc-16	Caisse	2	1	1000	100	98	<i>Acetobacter pasteurianus</i>
CI97*	Côte d'Ivoire	Wagana	déc-16	Caisse	2	3	1343	99,9	98	<i>Acetobacter pasteurianus</i>
CI98*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	0	1338	100	98	<i>Acetobacter pasteurianus</i>
CI99*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	2	1400	99,6	97	<i>Acetobacter pasteurianus</i>
GF105	Guyane	Combi	déc-17	Caisse		5	1358	100	99	<i>Acetobacter pasteurianus</i>
GF106	Guyane	Combi	déc-17	Caisse		5	1358	100	97	<i>Acetobacter pasteurianus</i>
GF41	Guyane	Combi	déc-17	Caisse		2	1306	100	100	<i>Acetobacter pasteurianus</i>
GF42	Guyane	Combi	déc-17	Caisse		2	1346	99,9	100	<i>Acetobacter pasteurianus</i>
GF43	Guyane	Combi	déc-17	Caisse		2	1295	99,9	100	<i>Acetobacter pasteurianus</i>

GF44	Guyane	Combi	déc-17	Caisse		2	1310	100	100	<i>Acetobacter pasteurianus</i>
GF45	Guyane	Combi	déc-17	Caisse		3	1334	100	100	<i>Acetobacter pasteurianus</i>
Code souche	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
GF46	Guyane	Combi	déc-17	Caisse		3	1346	99,9	100	<i>Acetobacter pasteurianus</i>
GF47	Guyane	Combi	déc-17	Caisse		5	1346	100	100	<i>Acetobacter pasteurianus</i>
GF48	Guyane	Combi	déc-17	Caisse		5	1352	99,8	100	<i>Acetobacter pasteurianus</i>
GF49	Guyane	Combi	déc-17	Caisse		5	1358	100	100	<i>Acetobacter pasteurianus</i>
GF50	Guyane	Combi	déc-17	Caisse		5	1305	99,9	100	<i>Acetobacter pasteurianus</i>
MX107	Mexique	Comalcalco	févr-18	Caisse	1	4	1344	100	98	<i>Acetobacter pasteurianus</i>
MX108	Mexique	Comalcalco	févr-18	Caisse	1	4	1355	100	97,9	<i>Acetobacter pasteurianus</i>
MX109	Mexique	Comalcalco	févr-18	Caisse	1	6	1341	96	99	<i>Acetobacter pasteurianus</i>
MX110	Mexique	Comalcalco	oct-18	Caisse	1	2	1309	100	100	<i>Acetobacter pasteurianus</i>
MX111	Mexique	Comalcalco	oct-18	Caisse	2	4	1351	100	97	<i>Acetobacter pasteurianus</i>
MX52	Mexique	Comalcalco	févr-18	Caisse	2	5	1345	100	100	<i>Acetobacter pasteurianus</i>
MX53	Mexique	Comalcalco	févr-18	Caisse	2	5	1316	100	100	<i>Acetobacter pasteurianus</i>
MX54	Mexique	Comalcalco	oct-18	Caisse	1	1	1358	100	100	<i>Acetobacter pasteurianus</i>
MX55	Mexique	Comalcalco	oct-18	Caisse	1	1	1363	100	100	<i>Acetobacter pasteurianus</i>
MX56	Mexique	Comalcalco	oct-18	Caisse	1	1	1358	100	100	<i>Acetobacter pasteurianus</i>
MX57	Mexique	Comalcalco	oct-18	Caisse	1	1	1259	100	100	<i>Acetobacter pasteurianus</i>
MX58	Mexique	Comalcalco	oct-18	Caisse	1	2	1304	100	100	<i>Acetobacter pasteurianus</i>
MX59	Mexique	Comalcalco	oct-18	Caisse	1	2	1347	100	100	<i>Acetobacter pasteurianus</i>
MX60	Mexique	Comalcalco	oct-18	Caisse	1	2	1351	100	100	<i>Acetobacter pasteurianus</i>
MX61	Mexique	Comalcalco	oct-18	Caisse	1	2	1305	100	100	<i>Acetobacter pasteurianus</i>
MX62	Mexique	Comalcalco	oct-18	Caisse	1	3	1305	100	100	<i>Acetobacter pasteurianus</i>
MX63	Mexique	Comalcalco	oct-18	Caisse	1	3	1353	100	100	<i>Acetobacter pasteurianus</i>

Code souche	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
MX64	Mexique	Comalcalco	oct-18	Caisse	1	3	1355	100	100	<i>Acetobacter pasteurianus</i>
MX65	Mexique	Comalcalco	oct-18	Caisse	1	3	1352	99,8	100	<i>Acetobacter pasteurianus</i>
MX66	Mexique	Comalcalco	oct-18	Caisse	1	3	1360	100	100	<i>Acetobacter pasteurianus</i>
MX67	Mexique	Comalcalco	oct-18	Caisse	1	3	1318	100	100	<i>Acetobacter pasteurianus</i>
MX68	Mexique	Comalcalco	oct-18	Caisse	1	4	1360	100	100	<i>Acetobacter pasteurianus</i>
MX69	Mexique	Comalcalco	oct-18	Caisse	1	4	1351	100	100	<i>Acetobacter pasteurianus</i>
MX70	Mexique	Comalcalco	oct-18	Caisse	1	4	1367	100	99	<i>Acetobacter pasteurianus</i>
MX71	Mexique	Comalcalco	oct-18	Caisse	1	4	1316	99,8	100	<i>Acetobacter pasteurianus</i>
MX72	Mexique	Comalcalco	oct-18	Caisse	2	1	1346	100	100	<i>Acetobacter pasteurianus</i>
MX73	Mexique	Comalcalco	oct-18	Caisse	2	2	1346	100	100	<i>Acetobacter pasteurianus</i>
MX74	Mexique	Comalcalco	oct-18	Caisse	2	2	1358	100	100	<i>Acetobacter pasteurianus</i>
MX75	Mexique	Comalcalco	oct-18	Caisse	2	2	1310	100	100	<i>Acetobacter pasteurianus</i>
MX76	Mexique	Comalcalco	oct-18	Caisse	2	2	1314	100	100	<i>Acetobacter pasteurianus</i>
MX77	Mexique	Comalcalco	oct-18	Caisse	2	3	1353	99,8	99	<i>Acetobacter pasteurianus</i>
MX78	Mexique	Comalcalco	oct-18	Caisse	2	3	1303	99,9	100	<i>Acetobacter pasteurianus</i>
MX79	Mexique	Comalcalco	oct-18	Caisse	2	3	1351	100	100	<i>Acetobacter pasteurianus</i>
MX80	Mexique	Comalcalco	oct-18	Caisse	2	3	1353	100	100	<i>Acetobacter pasteurianus</i>
MX81	Mexique	Comalcalco	oct-18	Caisse	2	4	1353	100	100	<i>Acetobacter pasteurianus</i>
MX82	Mexique	Comalcalco	oct-18	Caisse	2	4	1353	100	100	<i>Acetobacter pasteurianus</i>
MX83	Mexique	Comalcalco	févr-18	Caisse	1	2	1360	99,9	100	<i>Acetobacter pasteurianus</i>
MX84	Mexique	Comalcalco	févr-18	Caisse	1	4	1354	100	100	<i>Acetobacter pasteurianus</i>
MX85	Mexique	Comalcalco	févr-18	Caisse	1	4	1353	100	100	<i>Acetobacter pasteurianus</i>
MX86	Mexique	Comalcalco	févr-18	Caisse	1	6	1363	99,9	100	<i>Acetobacter pasteurianus</i>
MX87	Mexique	Comalcalco	févr-18	Caisse	1	6	1358	100	100	<i>Acetobacter pasteurianus</i>
MX88	Mexique	Comalcalco	févr-18	Caisse	1	6	1358	100	100	<i>Acetobacter pasteurianus</i>

Code souche	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
MX89	Mexique	Comalcalco	févr-18	Caisse	2	4	1355	100	100	<i>Acetobacter pasteurianus</i>
MX90	Mexique	Comalcalco	févr-18	Caisse	2	4	1354	100	100	<i>Acetobacter pasteurianus</i>
MX91	Mexique	Comalcalco	févr-18	Caisse	2	4	1358	100	100	<i>Acetobacter pasteurianus</i>
MX93	Mexique	Comalcalco	févr-18	Caisse	2	5	1363	100	100	<i>Acetobacter pasteurianus</i>
CI112*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	0	1363	99	100	<i>Acetobacter tropicalis</i>
CI113*	Côte d'Ivoire	Wagana	déc-16	Caisse	3	0	1331	99,9	99,7	<i>Acetobacter tropicalis</i>
CI114*	Côte d'Ivoire	Wagana	déc-16	Caisse	4	1	1335	100	100	<i>Acetobacter tropicalis</i>
CI115*	Côte d'Ivoire	Wagana	déc-16	Caisse	5	1	1342	99,9	96,6	<i>Acetobacter tropicalis</i>
CI116*	Côte d'Ivoire	Wagana	déc-16	Caisse	5	1	1345	99,9	100	<i>Acetobacter tropicalis</i>
CI117	Côte d'Ivoire	Abidjan	mai-18	Tas		5	1354	99,9	100	<i>Acetobacter tropicalis</i>
CI118	Côte d'Ivoire	Abidjan	mai-18	Tas		6	1360	99,7	100	<i>Acetobacter tropicalis</i>
GF119	Guyane	Combi	déc-17	Caisse		2	1347	99,9	100	<i>Acetobacter tropicalis</i>
GF120	Guyane	Combi	déc-17	Caisse		2	1308	99,8	100	<i>Acetobacter tropicalis</i>
MX121	Mexique	Comalcalco	févr-18	Caisse	1	2	1348	99,9	99	<i>Acetobacter tropicalis</i>
MX122	Mexique	Comalcalco	févr-18	Caisse	1	2	1356	99,8	99,7	<i>Acetobacter tropicalis</i>
MX123	Mexique	Comalcalco	févr-18	Caisse	2	2	1340	99,8	99,7	<i>Acetobacter tropicalis</i>
MX124	Mexique	Comalcalco	févr-18	Caisse	2	5	1344	99,9	99,8	<i>Acetobacter tropicalis</i>
MX125	Mexique	Comalcalco	oct-18	Caisse	2	4	1317	99	99	<i>Acetobacter tropicalis</i>
MX126	Mexique	Comalcalco	oct-18	Caisse	2	4	1216			<i>Acetobacter tropicalis</i>
MX92	Mexique	Comalcalco	févr-18	Caisse	2	2	1294	99,9	100	<i>Acetobacter tropicalis</i>
CI127*	Côte d'Ivoire	Wagana	déc-16	Caisse	1	0	1371	99,3	99,9	<i>Gluconobacter naphelli</i>
CI133*	Côte d'Ivoire	Wagana	déc-16	Caisse	0	0	1398	99,7	96,8	<i>Gluconobacter oxydans</i>
GF128	Guyane	Combi	déc-17	Caisse		1	1116	97,85	100	<i>Gluconobacter oxydans</i>

Code souche	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
GF129	Guyane	Combi	déc-17	Caisse		2	1397	99,8	96,8	<i>Gluconobacter oxydans</i>
GF130	Guyane	Combi	déc-17	Caisse		3	1397	99,8	96,8	<i>Gluconobacter oxydans</i>
GF131	Guyane	Combi	déc-17	Caisse		3	1397	99,8	96,9	<i>Gluconobacter oxydans</i>
GF132	Guyane	Combi	déc-17	Caisse		3	1396	99,8	96,4	<i>Gluconobacter oxydans</i>
GF51	Guyane	Combi	déc-17	Caisse		2	1357	99,7	96,8	<i>Gluconobacter oxydans</i>
LMG23848T	Type									<i>Acetobacter ghanensis</i>
LMG24244T	Type									<i>Acetobacter fabarum</i>
LMG01262T	Type									<i>Acetobacter pasteurianus</i>

Annexe 2 : Liste détaillée des souches de BAL isolées au cours de la thèse.

Code	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
MX-BAL1	Mexique	Comalcalco	oct-18	Caisse	1	0	1332	99,3	100	<i>Lactobacillus plantarum</i>
MX-BAL2	Mexique	Comalcalco	oct-18	Caisse	1	0	1331	99,85	100	<i>Lactobacillus plantarum</i>
MX-BAL3	Mexique	Comalcalco	oct-18	Caisse	1	0	1073	97,6	100	<i>Lactobacillus plantarum</i>
MX-BAL4	Mexique	Comalcalco	oct-18	Caisse	1	0	1324	99,9	100	<i>Lactobacillus plantarum</i>
MX-BAL5	Mexique	Comalcalco	oct-18	Caisse	1	1	1333	97	100	<i>Lactobacillus plantarum</i>
MX-BAL6	Mexique	Comalcalco	oct-18	Caisse	1	3	1333	99	100	<i>Lactobacillus plantarum</i>
MX-BAL7	Mexique	Comalcalco	oct-18	Caisse	1	3	1331	99,4	100	<i>Lactobacillus plantarum</i>
MX-BAL8	Mexique	Comalcalco	oct-18	Caisse	1	3	1327	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL9	Mexique	Comalcalco	oct-18	Caisse	1	4	1331	99,5	100	<i>Lactobacillus plantarum</i>

Code	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
MX-BAL10	Mexique	Comalcalco	oct-18	Caisse	1	4	1330	96,9	100	<i>Lactobacillus plantarum</i>
MX-BAL11	Mexique	Comalcalco	oct-18	Caisse	1	4	1331	99,6	100	<i>Lactobacillus plantarum</i>
MX-BAL12	Mexique	Comalcalco	oct-18	Caisse	1	4	1330	99,7	100	<i>Lactobacillus plantarum</i>
MX-BAL13	Mexique	Comalcalco	oct-18	Caisse	1	4	1325	99,6	100	<i>Lactobacillus plantarum</i>
MX-BAL14	Mexique	Comalcalco	oct-18	Caisse	2	2	1328	99,7	100	<i>Lactobacillus plantarum</i>
MX-BAL15	Mexique	Comalcalco	oct-18	Caisse	2	2	1329	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL16	Mexique	Comalcalco	oct-18	Caisse	2	2	1337	99,7	100	<i>Lactobacillus plantarum</i>
MX-BAL17	Mexique	Comalcalco	oct-18	Caisse	2	2	1208	99,3	100	<i>Lactobacillus plantarum</i>
MX-BAL18	Mexique	Comalcalco	oct-18	Caisse	2	2	1334	99,3	100	<i>Lactobacillus plantarum</i>
MX-BAL19	Mexique	Comalcalco	oct-18	Caisse	2	3	1325	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL20	Mexique	Comalcalco	oct-18	Caisse	2	3	1330	99,7	100	<i>Lactobacillus plantarum</i>
MX-BAL21	Mexique	Comalcalco	oct-18	Caisse	2	4	1329	99,1	100	<i>Lactobacillus plantarum</i>
MX-BAL22	Mexique	Comalcalco	oct-18	Caisse	2	4	1333	99,5	100	<i>Lactobacillus plantarum</i>
MX-BAL23	Mexique	Comalcalco	oct-18	Caisse	2	4	1338	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL24	Mexique	Comalcalco	oct-18	Caisse	2	4	1333	99,7	100	<i>Lactobacillus plantarum</i>
MX-BAL25	Mexique	Comalcalco	oct-18	Caisse	2	4	1374	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL26	Mexique	Comalcalco	oct-18	Caisse	2	4	1368	100	100	<i>Lactobacillus plantarum</i>
MX-BAL27	Mexique	Comalcalco	févr-18	Caisse	2	2	1372	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL28	Mexique	Comalcalco	févr-18	Caisse	1	4	1349	99,7	100	<i>Lactobacillus brevis</i>
MX-BAL29	Mexique	Comalcalco	févr-18	Caisse	1	4	1343	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL30	Mexique	Comalcalco	févr-18	Caisse	1	4	1374	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL31	Mexique	Comalcalco	févr-18	Caisse	2	4	1361	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL32	Mexique	Comalcalco	févr-18	Caisse	1	6	725 *	99,17	100	<i>Lactobacillus paracasei/ rhamnosus</i>
MX-BAL33	Mexique	Comalcalco	févr-18	Caisse	1	6	1373	99,8	100	<i>Lactobacillus plantarum</i>
MX-BAL34	Mexique	Comalcalco	févr-18	Caisse	2	6	1366	99,9	100	<i>Lactobacillus plantarum</i>
MX-BAL35	Mexique	Comalcalco	févr-18	Caisse	2	6	1347	99,8	100	<i>Lactobacillus brevis</i>
MX-BAL36	Mexique	Comalcalco	févr-18	Caisse	2	6	1414	99,8	100	<i>Lactobacillus plantarum</i>

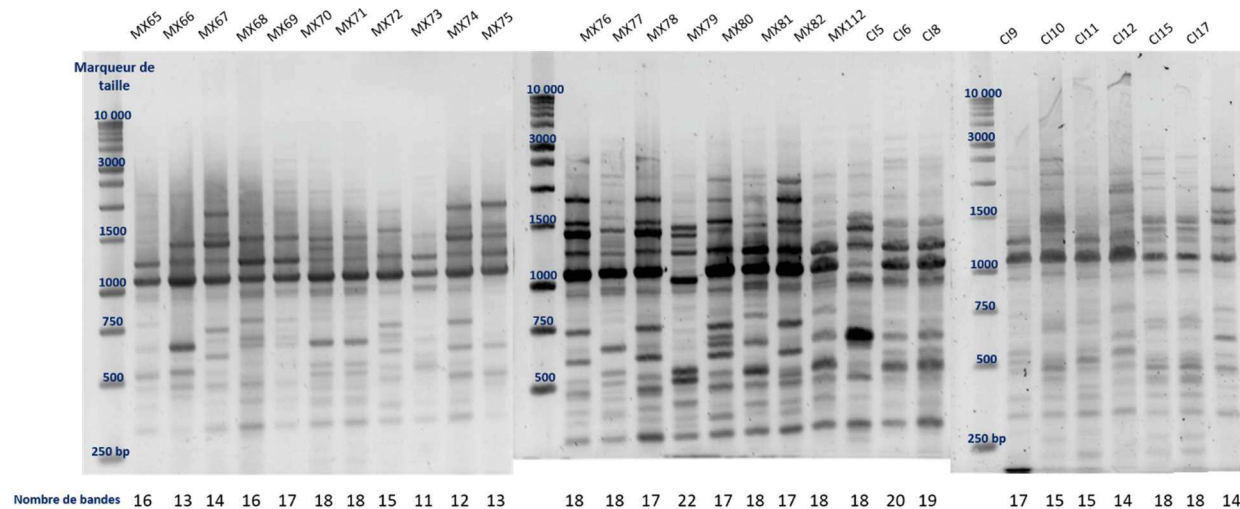
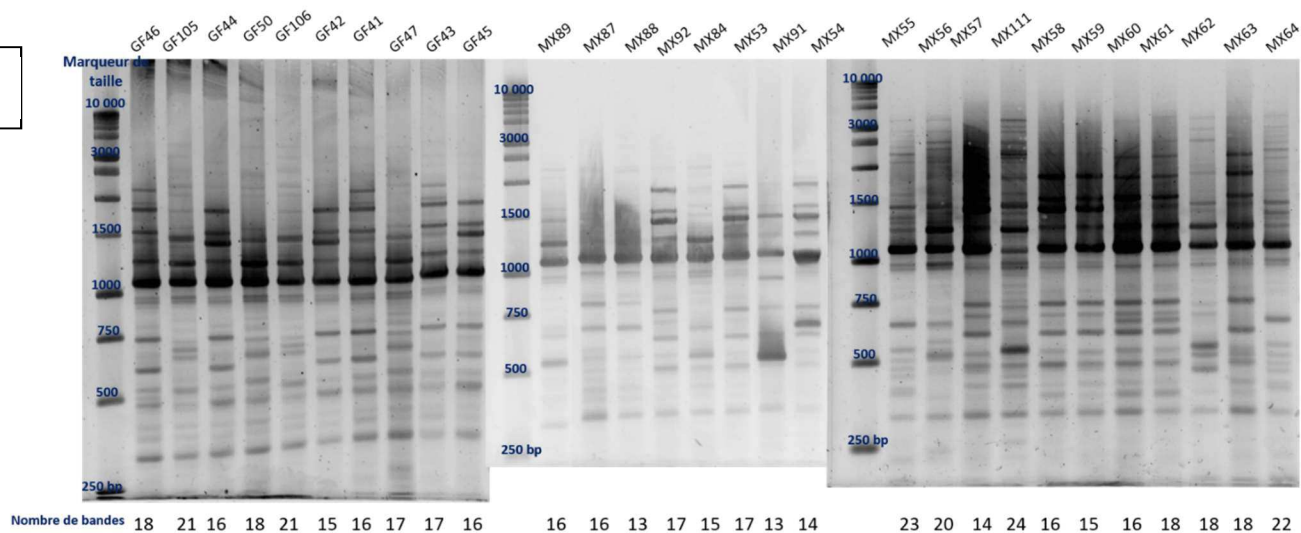
Code	Pays	Ville	Date de prélèvement	Type fermentation	Caisse	Jour de fermentation	Taille séquence	Id %	cover %	Identification
MX-BAL37	Mexique	Comalcalco	févr-18	Caisse	2	6	1314	98,6	100	<i>Lactobacillus plantarum</i>
CI-BAL38	Côte d'Ivoire	Abidjan	mai-18	Tas	1	6	1395	100	100	<i>Lactobacillus plantarum</i>
CI-BAL39	Côte d'Ivoire	Abidjan	mai-18	Tas	1	5	1343	99,8	100	<i>Lactobacillus plantarum</i>
CI-BAL40	Côte d'Ivoire	Abidjan	mai-18	Tas	1	5	1357	99,9	100	<i>Lactobacillus plantarum</i>
CI-BAL41	Côte d'Ivoire	Abidjan	mai-18	Tas	1	5	1342	99,5	100	<i>Lactobacillus plantarum</i>
CI-BAL42	Côte d'Ivoire	Abidjan	mai-18	Tas	1	2	1396	99,6	99	<i>Pediococcus acidilactici</i>
CI-BAL43	Côte d'Ivoire	Abidjan	mai-18	Tas	1	6	1383	99,9	100	<i>Lactobacillus plantarum</i>
CI-BAL44	Côte d'Ivoire	Abidjan	mai-18	Tas	1	4	1362	100	100	<i>Lactobacillus plantarum</i>
CI-BAL45	Côte d'Ivoire	Abidjan	mai-18	Tas	1	6	1366	99,9	100	<i>Lactobacillus plantarum</i>
CI-BAL46	Côte d'Ivoire	Abidjan	mai-18	Tas	1	0	1344	99,8	100	<i>Lactobacillus plantarum</i>
CI-BAL47	Côte d'Ivoire	Abidjan	mai-18	Tas	1	1	1367	99,9	100	<i>Lactobacillus plantarum</i>
CI-BAL48	Côte d'Ivoire	Abidjan	mai-18	Tas	1	1	1372	99,9	100	<i>Lactobacillus plantarum</i>
CI-BAL49	Côte d'Ivoire	Abidjan	mai-18	Tas	1	1	1404	100	100	<i>Lactobacillus plantarum</i>
CI-BAL50	Côte d'Ivoire	Abidjan	mai-18	Tas	1	5	1414	99,9	100	<i>Lactobacillus plantarum</i>
CI-BAL51	Côte d'Ivoire	Abidjan	mai-18	Tas	1	1	1404	100	100	<i>Lactobacillus plantarum</i>
CI-BAL52	Côte d'Ivoire	Abidjan	mai-18	Tas	1	5	1419	99,9	100	<i>Lactobacillus plantarum</i>
CI-BAL53	Côte d'Ivoire	Abidjan	mai-18	Tas	1	5	1352	99,9	100	<i>Lactobacillus plantarum</i>
CI-BAL54	Côte d'Ivoire	Abidjan	mai-18	Tas	1	5	1347	99	100	<i>Lactobacillus plantarum</i>
CI-BAL55	Côte d'Ivoire	Abidjan	mai-18	Tas	1	6	1403	99,6	100	<i>Lactobacillus plantarum</i>
CI-BAL56	Côte d'Ivoire	Abidjan	mai-18	Tas	1	4	1370	99	100	<i>Lactobacillus plantarum</i>
GF-BAL57	Guyane	Combi	déc-17	Caisse	1	6	1352	99,2	100	<i>Leuconostoc pseudomesenteroides</i>
GF-BAL58	Guyane	Combi	déc-17	Caisse	1	6	1341	99,9	100	<i>Fructobacillus tropaeoli</i>
GF-BAL59	Guyane	Combi	déc-17	Caisse	1	1	1344	99,9	100	<i>Leuconostoc pseudomesenteroides</i>
GF-BAL60	Guyane	Combi	déc-17	Caisse	1	1	1340	99,9	100	<i>Fructobacillus tropaeoli</i>
GF-BAL61	Guyane	Combi	déc-17	Caisse	1	3	1403	100	100	<i>Lactobacillus plantarum</i>
GF-BAL62	Guyane	Combi	déc-17	Caisse	1	3	1366	100	100	<i>Lactobacillus plantarum</i>
GF-BAL63	Guyane	Combi	déc-17	Caisse	1	3	1360	99,8	60	<i>Lactobacillus plantarum</i>



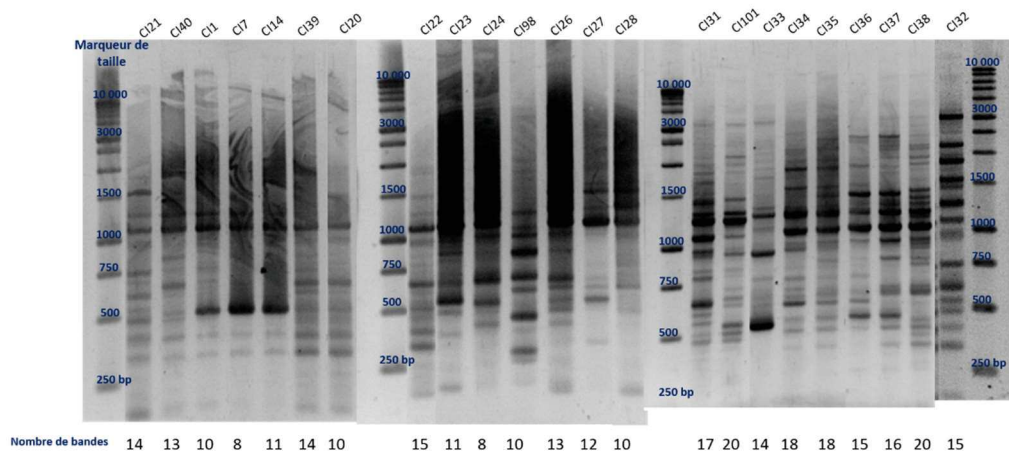
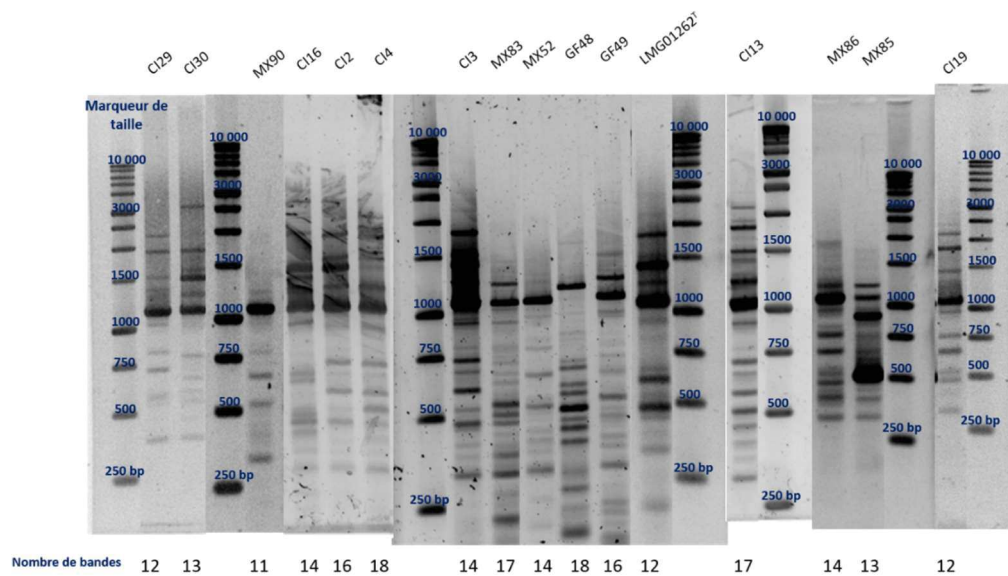
\* séquence incomplète (amorce forward)

Annexe 3 : Gels d'électrophorèse des profils génomiques obtenus pour les souches d'*A. pasteurianus* après la première amplification par PCR des à l'aide des amorces (GTG)<sub>5</sub>. M = 1kb - Les deux amplification (PCR1 et PCR2)

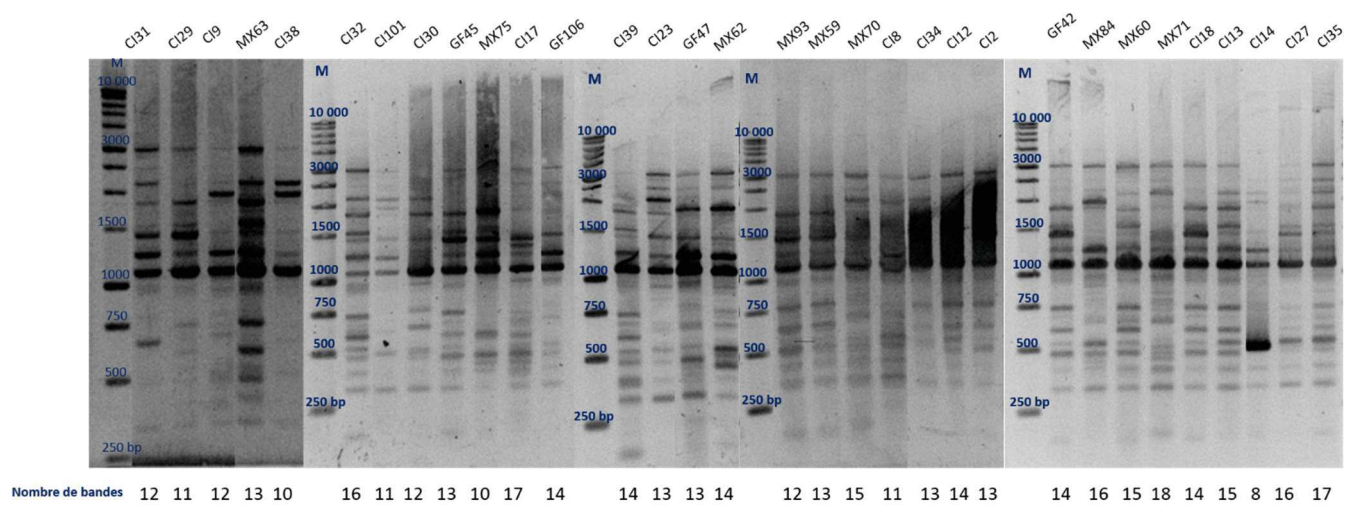
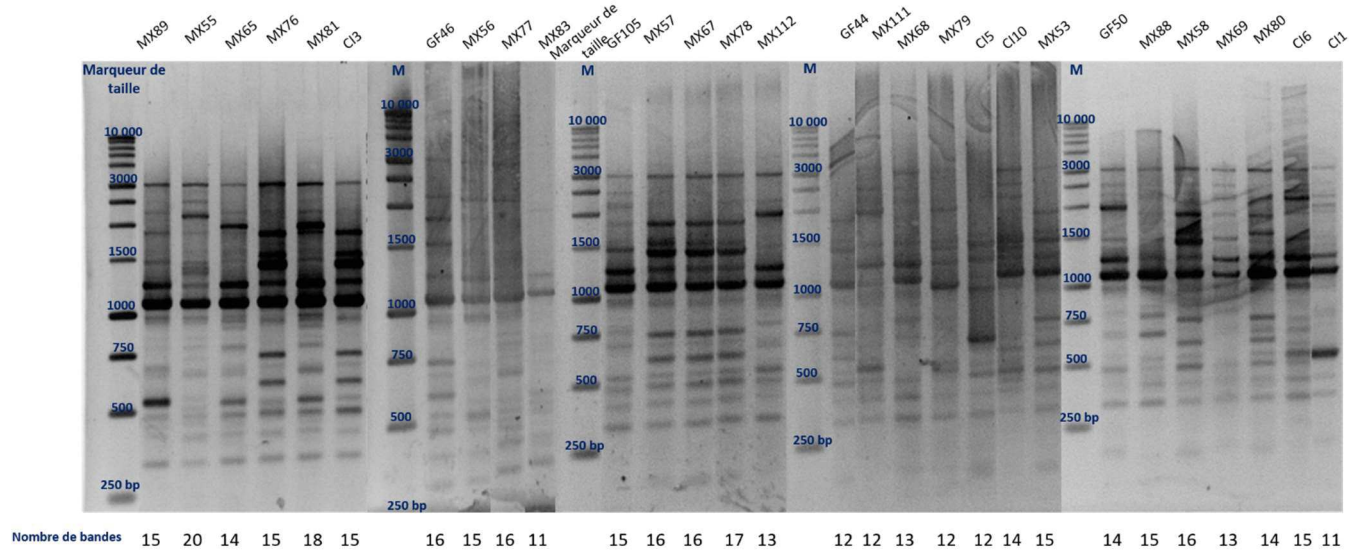
PCR1



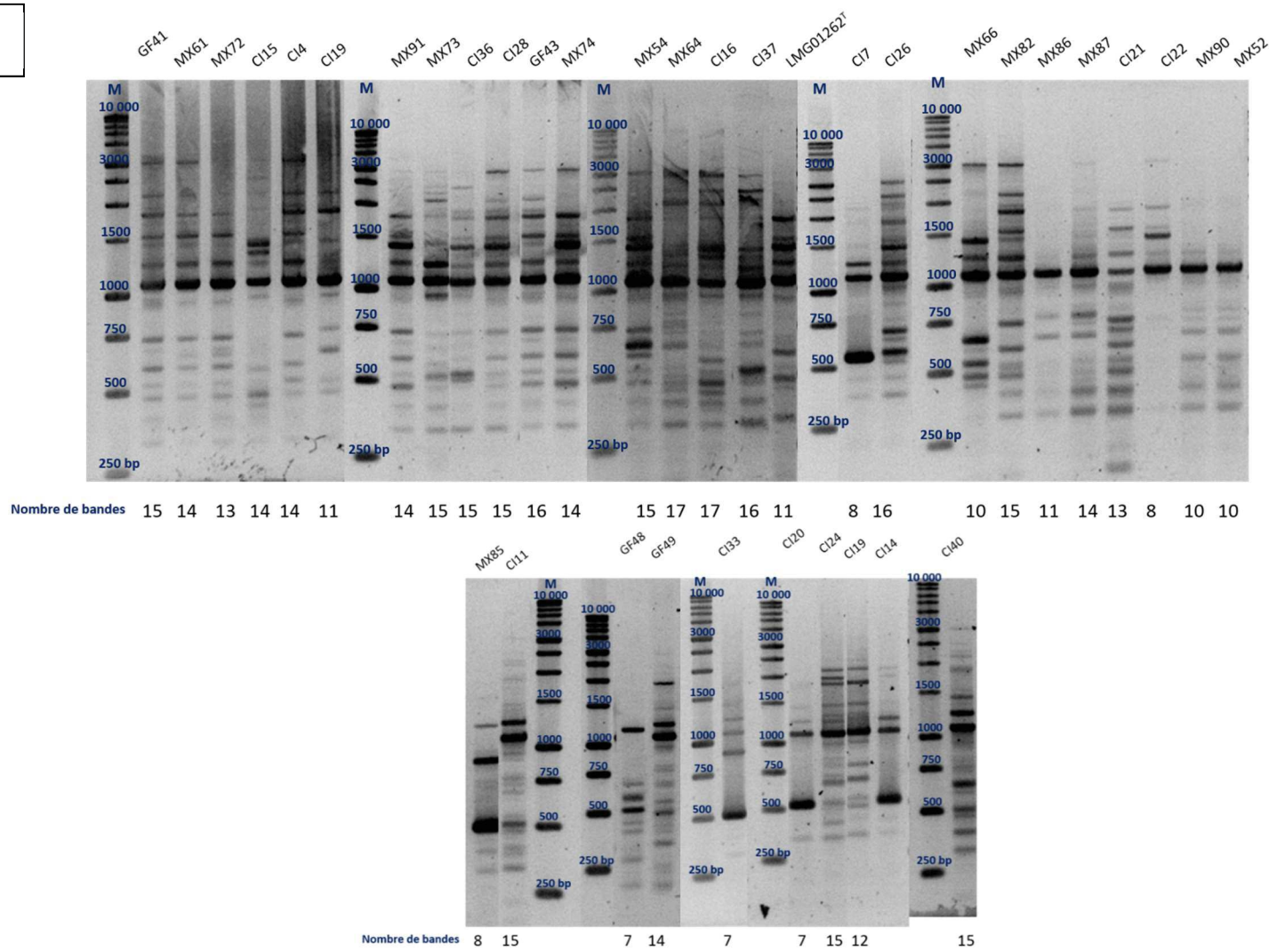
PCR1



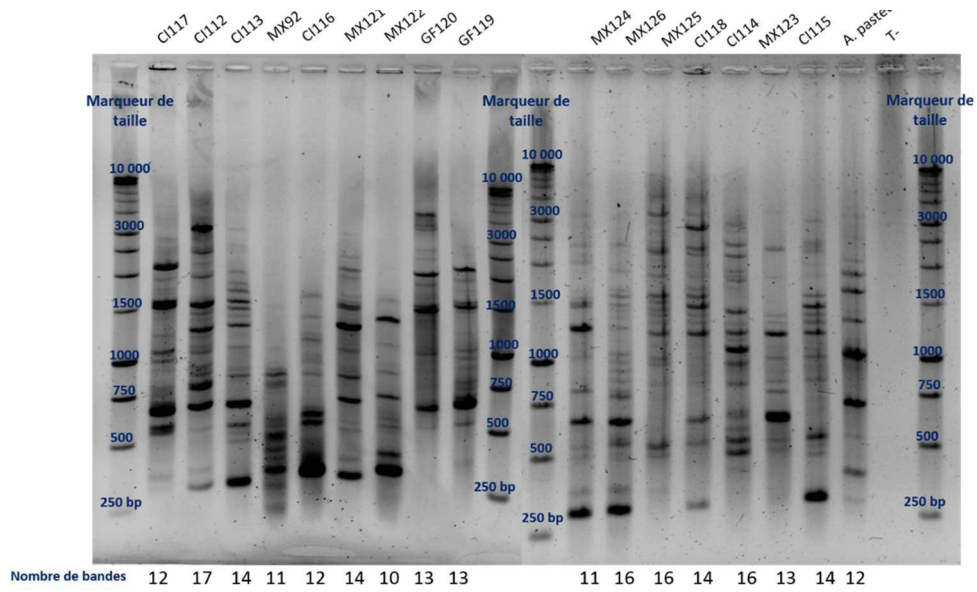
PCR2



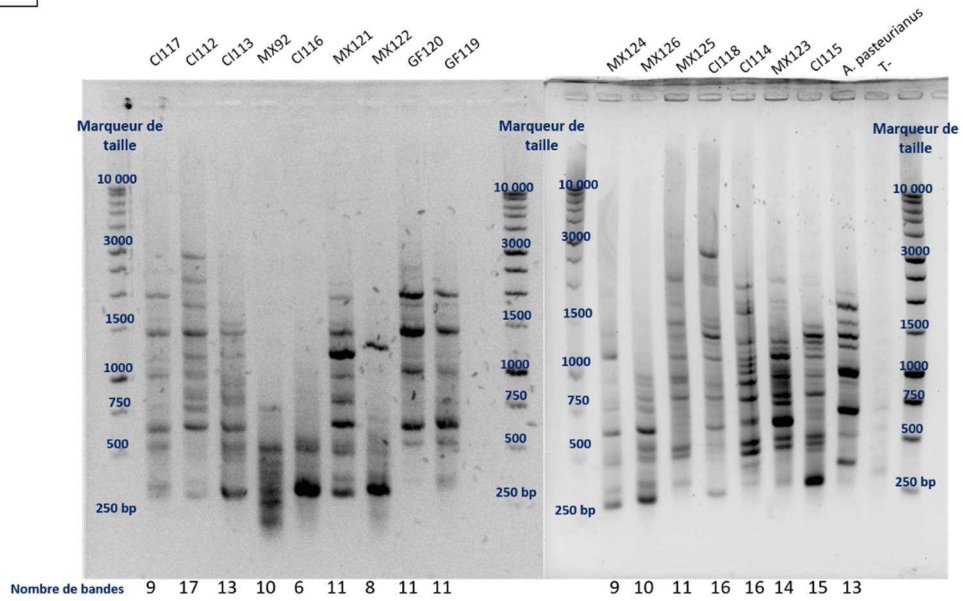
PCR2



Annexe 4 : Gels d'électrophorèse des profils génomiques obtenus pour les souches d'*A. tropicalis* après la première amplification par PCR des à l'aide des amorces (GTG)<sub>5</sub>. PCR1 et PCR2



PCR2



# Annals of Microbiology

## Genomic and metabolic intraspecific diversity among *A. pasteurianus* strains isolated from cocoa fermentation

--Manuscript Draft--

<b>Manuscript Number:</b>	ANMI-D-19-00754R1
<b>Full Title:</b>	Genomic and metabolic intraspecific diversity among <i>A. pasteurianus</i> strains isolated from cocoa fermentation
<b>Article Type:</b>	Original Articles
<b>Corresponding Author:</b>	Lucie FARRERA, M.D CIRAD Montpellier, France FRANCE
<b>Order of Authors:</b>	Lucie FARRERA, pHD Alexandre Colas de la Noue, pHD Caroline Strub, pHD Benjamin Guibert Christelle Kouame, pHD Joël Grabulos Didier Montet, pHD Corinne Teyssier, pHD
<b>Corresponding Author Secondary Information:</b>	
<b>Corresponding Author's Institution:</b>	CIRAD
<b>Corresponding Author's Secondary Institution:</b>	
<b>First Author:</b>	Lucie FARRERA, pHD
<b>First Author Secondary Information:</b>	
<b>Order of Authors Secondary Information:</b>	
<b>Funding Information:</b>	
<b>Abstract:</b>	<p><b>Purpose</b></p> <p>Acetic Acid Bacteria strains are involved in many food and beverage fermentation process. They play an important role in cocoa beans fermentation process by their acetic acid production. They initiate the development of some flavor precursors necessary for the organoleptic quality of cocoa and for the beans color. Developing starter cultures could be interesting to preserve the strains biodiversity of each country in cocoa production. This approach could avoid standardization of cocoa beans fermentation all around the producing countries.</p> <p><b>Methods</b></p> <p>One hundred and thirty Acetic Acid Bacteria were isolated from three different cocoa producing countries and identified based on their 16S rRNA gene sequence. Afterwards, intraspecific diversity of the predominant strains was assessed through analysis of their genomic polymorphism with (GTG)<sub>5</sub>-PCR fingerprinting. Finally, strains were grown in a cocoa pulp simulation medium (CPSM-AAB) in order to compare their physiological traits regarding their specific growth rate, ethanol and lactic acid specific consumption, and acetic acid specific production, and relative preferences for carbon sources.</p> <p><b>Results</b></p>

	<p>Our results showed that whatever the country, <i>Acetobacter pasteurianus</i> was the most recovered species. A great similarity was observed between strains according to their genomic polymorphisms and physiological characterization but clustering analysis showed that strains were significantly clustered according to origin's country. However, multiparametric clustering results in different groups, highlighted some differences in their basic metabolism, such as efficiency in converting carbon substrates to acetate and relative affinity for acid lactic and ethanol.</p> <p>Conclusion</p> <p>Among the 89 isolated <i>A. pasteurianus</i>, strain selected for their starter culture potential depended of their ability to oxidase ethanol and lactic acid. Ethanol and lactic acid conversion efficiency into acetic acid and relative affinity for lactic acid and ethanol could be used as criteria for acetic acid bacteria starter selection.</p>
<p><b>Response to Reviewers:</b></p>	<p>Responses to the reviewers and editor comments</p> <p>Reviewers' comments:</p> <p>Reviewer #1: In the manuscript, the authors present the characterization of acetic acid bacteria isolated from cocoa bean fermentations performed in different country, with the aim of selecting potential starters. The organization of the manuscript is good and the results are clearly presented and interesting. The introduction is well structured and focused on the most relevant aspects. I only suggest an accurate revision of the Conclusion section that seems more like a summary. The authors should draw the conclusions in terms of what are the most promising strains; how do they supposed to use them. Finally, I recommend these minor revisions:</p> <p>1. The authors should check all misprinting: for example, LINE 50: no space before the colon; LINE 88: no space after bracket etc etc. The authors should do a thorough review of English: for example, LINE 108 - focus on- LINE 112 - The analysis of...) etc etc</p> <p>The manuscript was entirely revised by a native English speaker.</p> <p>INTRODUCTION</p> <p>2. LINE 69: the authors should mention also the role of LAB in pH drop This part has been reformulated from "Action of yeasts results in a pH drop and a pulp depectinization. The yeast's pectinase activity helps to reduce the viscosity of the pulp and allows air penetration (Crafack et al. 2013). Then, LAB consume glucose, fructose and citric acid to generate lactic acid, acetic acid, ethanol and mannitol." to "Meanwhile, LAB consume glucose, fructose and citric acid to generate lactic acid, acetic acid, ethanol and mannitol. LAB metabolic activity results in a pH drop mainly due to lactic acid production. The hydrolysis of pectin by yeasts progressively reduces the viscosity of the pulp and allows air to penetrate the bean mass (Crafack et al. 2013).</p> <p>3. LINE 79: please, give more details about aroma precursors and the way acetic acid is affecting the colour. This part has been reformulated from "In addition, acetic acid initiates the production of flavor precursors necessary for the organoleptic quality of cocoa and for the beans color. Acetic acid is further eliminated by sun drying (Hansen et al. 1998)." to "Acetification plays a key role in cocoa fermentation because acetic acid penetrates the cotyledons and destabilizes the subcellular structure, with an impact on proteins, polyphenols and lipids. It promotes the production of flavour precursors, and notably amino-acids liberation through proteolysis that will further interact with sugars and polyphenols during the roasting process (John et al. 2016). The colour of the beans is also affected through the degradation of polyphenol, particularly anthocyanins, enhanced by acetic acid (De Taeye et al. 2016; Urbańska et al. 2019). It also prevents seed germination by participating in the death of the embryo, coupled with the temperature increase due to AAB exothermic metabolism."</p> <p>Materials &amp; Methods</p> <p>4. LINES 137, 139-141: what the authors mean for from days 0 to 4-6-7? Every day sampling ? Or what? Please, specify. This part has been reformulated; it was sampling each day.</p> <p>5. LINE 150: please, specify the composition of CPSM substrate. Composition of CPSM has been given "CPSM-AAB is a specific liquid medium</p>

developed by Moens et al. (2014) called Cocoa Pulp Simulation Medium for AAB (CPSM-AAB). It reproduced the conditions found after two days of fermentation necessary for the development of AAB. It is composed of 1% ethanol (Honeywell); 10 g/L lactic acid (Sigma-Aldrich); 10 g/L yeast extract (Becton Dickinson), 5 g/L soya peptone (Merck GaA), and 1 mL/L Tween 80 (PanReac) at pH 4.5.”

#### RESULTS

6. LINES 280: this sentence sound a little bit strange considering that all clusters contained strains from at least 2 countries. How do the authors explain it?

This sentence was added to the article:

“This effect was explained by strains with high percentage of similarity (>90%), which came from the same origin.”

7. LINES 284-285: I would not use the expression "acetic fermentation" which is wrong from a metabolic point of view.

The expression “acetic fermentation” has been reformulate to “acetification” (L72)  
Changed lines 284-285 « in a medium simulating the cocoa pulp conditions during fermentation process”

8. LINES 334-335: this sentence is referred to what? Is it a general evidence. Please, re-formulate this sentence.

Done, we changed to “All around the world, cocoa bean fermentations are mostly performed in a traditional way using a variety of production methods. This artisanal process leads to variable organoleptic quality of the cocoa products.”

9. LINES 337-338: intraspecific diversity should not be correlated to starter efficiency but to biodiversity maintain (LINES 339-340). Please, re-formulate this part.

Done, it was reformulated

10. LINES 351-354: The authors should underline from the beginning of the sentence that they are not talking about their experimentation.

This part has been reformulated

11. LINES 364-366: Again, the shift in commenting on your results and results of other authors is unclear. Review all discussions in this regard.

The discussion was reviewed considering the different remarks of the reviewer

12. LINE 374: The figure should not be mention in Discussion Chapter.

The mention has been deleted

13. LINES 388-389: which feature? Is not clear.

The sentence has been clarified « acetic acid overoxidation should be avoided for AAB starter »

#### CONCLUSION

14. This chapter should be reviewed. In this form it seems more like a summary. The authors should draw the conclusions in terms of, for example, what are the most promising strains; how do they supposed to use them; in combination with other strains? alone? etc etc

The conclusion was reviewed considering the different remarks of the reviewer

Reviewer #2: Comments for ANMI-D-19-00754: Role and intraspecific diversity of Acetic Acid Bacteria in cocoa beans fermentation from three different geographic origins.

Comments to the authors are presented as below.

In this MS, the authors aim to analyze the intraspecific diversity of AAB strains isolated from cocoa fermentation in three countries: Ivory Coast, Mexico and French Guiana.

1. The total number of samples collected in each country is too small, esp., only one in French Guiana. Therefore, it is not so persuade in the point of view on the diversity at different geographic origins or the diversity in the three countries. The authors should re-consider the title and aim.

The title and the aim have been changed:

- “Role and intraspecific diversity of Acetic Acid Bacteria in cocoa beans fermentation from three different geographic origins” to “Genomic and metabolic intraspecific diversity among *A. pasteurianus* strains isolated from cocoa fermentation”

- “The aim of our study is to analyze the intraspecific diversity of *A. pasteurianus* strains isolated from cocoa fermentation in three countries: Ivory Coast, Mexico and French Guiana” to “ The aim of our study is to analyze the intraspecific diversity of AAB



strains isolated from cocoa fermentation”

2. The data obtained from the identification of 130 AAB isolated by 16S rRNA gene sequencing should be analyzed and reported more details, including Genbank accession No.; length of analyzed sequences; similarity values with closest validated AAB species phylogenetic tree(s). The blasting on NCBI only may not give precise identification to species.

Details for the 89 *A. pasteurianus* strains and the 3 non AAB strains analyzed in this study are given in supplemental material: table S1

Sequences will be online on Genbank when the article will be published.

3. Please show the (GTG)<sub>5</sub>-PCR banding patterns of strains and references together with the dendrogram. Papalexandratou et al. (2009) found some (GTG)<sub>5</sub> PCR banding patterns from well identified *A. pasteurianus* strains, those data should be mentioned in your discussion.

Migration was performed 11 by 11, and it was for that reason that we had to normalized migration bands thanks to a software (Cliqs) as mentioned in M&M, so it was difficult for us to include banding patterns in our dendrogram. Furthermore, the protocol of migration cited in the paper of Papalexandratou was different from our study.

4. Please give detailed references for each part in Biochemical characterization, esp., statistical analysis. And, please give reference for the sentence in line 292-293.

References have been added to the manuscript “(Steensels et al. 2014, Brase 2017).” For the statistical analysis.

Minor comments:

1. Tables 3 and 4: the isolates should be presented by groups of origin countries.

We did it by cluster to differentiate and better describe the cluster in order to indicate the percentage of strains from the different origin countries

2. Figure S1: Positions in all of chart graphs should be the same for each isolate and strain. The isolates should be presented by groups of origin countries or clusters, also. And, please give title for vertical axis of the first chart graph.

Position are now the same in each graph and separated by origin countries

3. Please give title for vertical axis in chart graphs that is still lacking.

Vertical axis has been given for all chart graphs

4. Please use normal style but not capitalized for all "Acetic Acid Bacteria", since "acetic acid bacteria" is a common noun but not a proper noun.

"Acetic Acid Bacteria" has been changed to "acetic acid bacteria"

5. Please make explanation for "UOD" and correct mistyping, "UDO" (Line 312, 325, 316).

Explanation has been done and mistyping removed.

6. Please improve the English sentences and mistyping, i.e., line 23, 36, 345-346, 554,...

Mistyping have been removed

7. Please check again the references in lines 477, 547-549, and 562-564.

References have been checked

## Responses to the reviewers and editor comments

Dear editor,

We have taken into account all the remarks made by the reviewers and have improved our article accordingly.. We have improved the English with the help of a native speaker. Please find our responses for these comments in red hereafter. All the corrections are highlighted in yellow in the manuscript.

Thank you very much for the delay and for reconsidering our article for publication.

Yours sincerely,

### Reviewers' comments:

Reviewer #1: In the manuscript, the authors present the characterization of acetic acid bacteria isolated from cocoa bean fermentations performed in different country, with the aim of selecting potential starters. The organization of the manuscript is good and the results are clearly presented and interesting. The introduction is well structured and focused on the most relevant aspects. I only suggest an accurate revision of the Conclusion section that seems more like a summary. The authors should draw the conclusions in terms of what are the most promising strains; how do they supposed to use them. Finally, I recommend these minor revisions:

*1. The authors should check all misprinting: for example, LINE 50: no space before the colon; LINE 88: no space after bracket etc etc. The authors should do a thorough review of English: for example, LINE 108 - focus on- LINE 112 - The analysis of...) etc etc*

*The manuscript was entirely revised by a native English speaker.*

### INTRODUCTION

*2. LINE 69: the authors should mention also the role of LAB in pH drop*

*This part has been reformulated from "Action of yeasts results in a pH drop and a pulp depectinization. The yeast's pectinase activity helps to reduce the viscosity of the pulp and allows air penetration (Crafack et al. 2013). Then, LAB consume glucose, fructose and citric acid to generate lactic acid, acetic acid, ethanol and mannitol." to "Meanwhile, LAB consume glucose, fructose and citric acid to generate lactic acid, acetic acid, ethanol and mannitol. LAB metabolic activity results in a pH drop mainly due to lactic acid production. The hydrolysis of pectin by yeasts progressively reduces the viscosity of the pulp and allows air to penetrate the bean mass (Crafack et al. 2013).*

*3. LINE 79: please, give more details about aroma precursors and the way acetic acid is affecting the colour.*

This part has been reformulated from “In addition, acetic acid initiates the production of flavor precursors necessary for the organoleptic quality of cocoa and for the beans color. Acetic acid is further eliminated by sun drying (Hansen et al. 1998).” to “Acetification plays a key role in cocoa fermentation because acetic acid penetrates the cotyledons and destabilizes the subcellular structure, with an impact on proteins, polyphenols and lipids. It promotes the production of flavour precursors, and notably amino-acids liberation through proteolysis that will further interact with sugars and polyphenols during the roasting process (John et al. 2016). The colour of the beans is also affected through the degradation of polyphenol, particularly anthocyanins, enhanced by acetic acid (De Taeye et al. 2016; Urbańska et al. 2019). It also prevents seed germination by participating in the death of the embryo, coupled with the temperature increase due to AAB exothermic metabolism.”

#### Materials & Methods

4. *LINES 137, 139-141: what the authors mean for from days 0 to 4-6-7? Every day sampling ? Or what? Please, specify.*

This part has been reformulated; it was sampling each day.

5. *LINE 150: please, specify the composition of CPSM substrate.*

Composition of CPSM has been given “CPSM-AAB is a specific liquid medium developed by Moens et al. (2014) called Cocoa Pulp Simulation Medium for AAB (CPSM-AAB). It reproduced the conditions found after two days of fermentation necessary for the development of AAB. It is composed of 1% ethanol (Honeywell); 10 g/L lactic acid (Sigma-Aldrich); 10 g/L yeast extract (Becton Dickinson), 5 g/L soya peptone (Merck GaA), and 1 mL/L Tween 80 (PanReac) at pH 4.5.”

#### RESULTS

6. *LINES 280: this sentence sound a little bit strange considering that all clusters contained strains from at least 2 countries. How do the authors explain it?*

This sentence was added to the article:

“This effect was explained by strains with high percentage of similarity (>90%), which came from the same origin.”

7. *LINES 284-285: I would not use the expression "acetic fermentation" which is wrong from a metabolic point of view.*

The expression “acetic fermentation” has been reformulate to “acetification” (L72)

Changed lines 284-285 « in a medium simulating the cocoa pulp conditions during fermentation process”

8. *LINES 334-335: this sentence is referred to what? Is it a general evidence. Please, re-formulate this sentence.*

Done, we changed to “All around the world, cocoa bean fermentations are mostly performed in a traditional way using a variety of production methods. This artisanal process leads to variable organoleptic quality of the cocoa products.”

9. *LINES 337-338: intraspecific diversity should not be correlated to starter efficiency but to biodiversity maintain (LINES 339-340). Please, re-formulate this part.*

Done, it was reformulated

10. LINES 351-354: *The authors should underline from the beginning of the sentence that they are not talking about their experimentation.*

This part has been reformulated

11. LINES 364-366: *Again, the shift in commenting on your results and results of other authors is unclear. Review all discussions in this regard.*

The discussion was reviewed considering the different remarks of the reviewer

12. LINE 374: *The figure should not be mention in Discussion Chapter.*

The mention has been deleted

13. LINES 388-389: *which feature? Is not clear.*

The sentence has been clarified « acetic acid overoxidation should be avoided for AAB starter »

CONCLUSION

14. This chapter should be reviewed. In this form it seems more like a summary. The authors should draw the conclusions in terms of, for example, what are the most promising strains; how do they supposed to use them; in combination with other strains? alone? etc etc

The conclusion was reviewed considering the different remarks of the reviewer

Reviewer #2: Comments for ANMI-D-19-00754: Role and intraspecific diversity of Acetic Acid Bacteria in cocoa beans fermentation from three different geographic origins.

Comments to the authors are presented as below.

In this MS, the authors aim to analyze the intraspecific diversity of AAB strains isolated from cocoa fermentation in three countries: Ivory Coast, Mexico and French Guiana.

1. *The total number of samples collected in each country is too small, esp., only one in French Guiana. Therefore, it is not so persuade in the point of view on the diversity at different geographic origins or the diversity in the three countries. The authors should re-consider the title and aim.*

The title and the aim have been changed:

- "Role and intraspecific diversity of Acetic Acid Bacteria in cocoa beans fermentation from three different geographic origins" to "Genomic and metabolic intraspecific diversity among *A. pasteurianus* strains isolated from cocoa fermentation"

- "The aim of our study is to analyze the intraspecific diversity of *A. pasteurianus* strains isolated from cocoa fermentation in three countries: Ivory Coast, Mexico and French Guiana" to "The aim of our study is to analyze the intraspecific diversity of AAB strains isolated from cocoa fermentation"

2. *The data obtained from the identification of 130 AAB isolated by 16S rRNA gene sequencing should be analyzed and reported more details, including Genbank accession No.; length of analyzed sequences; similarity values with closest validated AAB species phylogenetic tree(s). The blasting on NCBI only may not give precise identification to species.*

Details for the 89 *A. pasteurianus* strains and the 3 non AAB strains analyzed in this study are given in supplemental material: table S1

Sequences will be online on Genbank when the article will be published.

3. Please show the (GTG)<sub>5</sub>-PCR banding patterns of strains and references together with the dendrogram. Papalexandratou et al. (2009) found some (GTG)<sub>5</sub> PCR banding patterns from well identified *A. pasteurianus* strains, those data should be mentioned in your discussion.

Migration was performed 11 by 11, and it was for that reason that we had to normalized migration bands thanks to a software (Cliqs) as mentioned in M&M, so it was difficult for us to include banding patterns in our dendrogram. Furthermore, the protocol of migration cited in the paper of Papalexandratou was different from our study.

4. Please give detailed references for each part in Biochemical characterization, esp., statistical analysis. And, please give reference for the sentence in line 292-293.

References have been added to the manuscript "(Steensels et al. 2014, Brase 2017)." For the statistical analysis.

Minor comments:

1. Tables 3 and 4: the isolates should be presented by groups of origin countries.

We did it by cluster to differentiate and better describe the cluster in order to indicate the percentage of strains from the different origin countries

2. Figure S1: Positions in all of chart graphs should be the same for each isolate and strain. The isolates should be presented by groups of origin countries or clusters, also. And, please give title for vertical axis of the first chart graph.

Position are now the same in each graph and separated by origin countries

3. Please give title for vertical axis in chart graphs that is still lacking.

Vertical axis has been given for all chart graphs

4. Please use normal style but not capitalized for all "Acetic Acid Bacteria", since "acetic acid bacteria" is a common noun but not a proper noun.

"Acetic Acid Bacteria" has been changed to "acetic acid bacteria"

5. Please make explanation for "UOD" and correct mistyping, "UDO" (Line 312, 325, 316).

Explanation has been done and mistyping removed.

6. Please improve the English sentences and mistyping, i.e., line 23, 36, 345-346, 554,...

Mistyping have been removed

7. Please check again the references in lines 477, 547-549, and 562-564.

References have been checked

[Click here to view linked References](#)

# 1 **Genomic and metabolic intraspecific diversity among *A. pasteurianus* strains isolated from cocoa** 2 **fermentation**

3  
4  
5  
6  
7 **Lucie Farrera<sup>ab\*</sup>, Alexandre Colas de la Noue<sup>b</sup>, Caroline Strub<sup>b</sup>, Benjamin Guibert<sup>a</sup>, Christelle**  
8 **Kouame<sup>a</sup>, Joël Grabulos<sup>a</sup>, Didier Montet<sup>a</sup> and Corinne Teyssier<sup>b</sup>**

9  
10  
11  
12  
13  
14 <sup>a</sup> UMR 95 QualiSud, Centre de Coopération Internationale en Recherche Agronomique pour le Développement  
15 (CIRAD), Université de Montpellier, Montpellier SupAgro, Université d'Avignon et des Pays de Vaucluse,  
16  
17 Université de la Réunion, 34398 Montpellier Cedex 5, France,

18  
19  
20  
21 <sup>b</sup> UMR 95 QualiSud, Université de Montpellier, CIRAD, Montpellier SupAgro, Université d'Avignon et des Pays  
22 de Vaucluse, Université de la Réunion, 34095 Montpellier Cedex 5, France,

23  
24  
25  
26  
27 Corresponding author: [lucie.farrera@gmail.com](mailto:lucie.farrera@gmail.com)\*

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  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

## 15 **ABSTRACT**

### 16 **Purpose**

17 Acetic acid bacteria are involved in many food and beverage fermentation process. They play an important role in  
18 cocoa beans fermentation by their acetic acid production. They initiate the development of some of the flavour  
19 precursors necessary for the organoleptic quality of cocoa and for bean colour. Developing starter cultures with  
20 local strains would enable the preservation of the microbial biodiversity of each country in cocoa-producing areas.  
21 This approach could avoid the standardization of cocoa bean fermentation in the producing countries.

### 22 **Methods**

23 One hundred and thirty acetic acid bacteria were isolated from three different cocoa producing countries and  
24 identified based on their 16S rRNA gene sequence. The intraspecific diversity of the predominant strains was then  
25 assessed through analysis of their genomic polymorphism with (GTG)<sub>5</sub>-PCR fingerprinting. Finally, strains were  
26 grown in a cocoa pulp simulation medium (CPSM-AAB) to compare their physiological traits regarding their  
27 specific growth rate, ethanol and lactic acid consumption, acetic acid production, and relative preferences of carbon  
28 sources.

### 29 **Results**

30 Our results showed that *Acetobacter pasteurianus* was the most recovered species in all origins with 89 isolates  
1 31 out of 130 cultures. A great similarity was observed between strains according to their genomic polymorphisms  
2 32 and physiological characterization, but clustering analysis showed that strains were significantly clustered  
3 33 according to each country. However, multiparametric clustering results in different groups highlighted some  
4 34 differences in their basic metabolism, such as efficiency in converting carbon substrates to acetate and relative  
5 35 affinity to lactic acid and ethanol.  
6  
7  
8

## 9 36 **Conclusion**

10  
11 37 *A. pasteurianus* strains showed different behaviors regarding their ability to oxidize ethanol and lactic acid into  
12 38 acetic acid and their relative preference for each substrate. The impact of these behaviors on cocoa quality should  
13 39 be investigated and considered as a criterion for the selection of acetic acid bacteria starters.  
14  
15  
16

17 40 **Keywords:** Cocoa beans fermentation, acetic acid bacteria, (GTG)<sub>5</sub> rep PCR, lactic acid, acetic acid, ethanol.  
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

## 43 Introduction

44  
45 *Acetobacter* is a genus of acetic acid bacteria (AAB) that belong to the family of *Acetobacteraceae* in the  
46 *Alphaproteobacteria* class. They are Gram negative anaerobe, ellipsoidal or rod-shape, oxidase negative, and  
47 catalase positive (Sengun and Karabiyikli, 2011; Teyssier and Hamdouche, 2016). *Acetobacter* species are  
48 involved in many food and beverage fermentation processes with sugar and alcohol rich environments including  
49 vinegar, beer, water kefir, kombucha, coffee, etc. (De Roos and De Vuyst. 2018). They also play an important  
50 role in the fermentation of cocoa beans. Cocoa beans are collected from fruits called pods of the cocoa tree  
51 *Theobroma cacao* L, which are transformed by a succession of post-harvest treatments e.g. opening, fermentation,  
52 and drying (Barel 2013). Nowadays cocoa is produced in 50 tropical countries in a well-defined area called the  
53 cocoa belt. The main producing countries are Ivory Coast, Ghana, Nigeria and Cameroon in Africa, Indonesia and  
54 Malaysia in Asia, and Brazil, Ecuador, Mexico and Peru in Latin America (Ozturk and Young 2017).

55 The spontaneous fermentation of cocoa beans lasts between 4 to 7 days based on the succession of microorganisms,  
56 which are mainly yeasts, lactic acid bacteria (LAB) and AAB (Schwan and Wheals 2004; Nielsen et al. 2007;  
57 Camu et al. 2007, 2008; Daniel et al. 2009; Garcia-Armisen et al. 2010; Lefeber et al. 2012; Meersman et al. 2013;  
58 De Vuyst and Weckx, 2016). The fermentation process takes place in the mucilaginous pulp surrounding the beans,  
59 forming an impenetrable mass by air. The pulp is white, viscous and has a low pH ranging between 3 and 4. It is  
60 mainly composed of 80-90% water, 10-15% sugar (glucose, fructose, and sucrose), 2-3% pentosans, 2% citric acid  
61 and 1% pectin. Its composition varies all along the fermentation process following microbial activities (Pettipher  
62 1986; Schwan and Wheals 2004). The choice of mature, healthy pods in plantations, and the smooth progress of a  
63 fermentation are crucial for the establishment of the microorganisms involved in the fermentation, which  
64 influences the final cocoa organoleptic quality (Camu et al. 2007). According to the local farming practices,  
65 fermentation process can be performed in a heap (on leaves) or in boxes. Bean turning is often performed in order  
66 to oxygenate pulp mass and promote aerobic microorganisms.

67 The anaerobic fermentation is performed by yeasts and LAB. Yeasts metabolize glucose, fructose and citric acid  
68 to produce ethanol and carbon dioxide (Ardhana 2003; Schwan and Wheals 2004; Nielsen et al.2007; Camu et al  
69 .2007, 2008; L. Papalexandratou et al. 2011a, 2011b, 2011c,2013; De Vuyst and Weckx 2016). Meanwhile, LAB  
70 consume glucose, fructose and citric acid to generate lactic acid, acetic acid, ethanol and mannitol. LAB metabolic  
71 activity results in a pH drop mainly due to lactic acid production. The hydrolysis of pectin by yeasts progressively



72 reduces the viscosity of the pulp and allows air to penetrate the bean mass (Crafack et al. 2013). This air intake  
73 allowed by pulp liquefaction, combined with lactic acid and ethanol, progressively leads to favorable conditions  
74 for the growth of AAB. In this environment, AAB quickly metabolizes ethanol to produce acetic acid, but also  
75 lactic acid that is considered as an undesirable fermentation by-product which can lead to off flavours (Holm et al.  
76 1993; Camu et al. 2007; Lefeber et al. 2012). Acetification plays a key role in cocoa fermentation because acetic  
77 acid penetrates the cotyledons and destabilizes the subcellular structure, with an impact on proteins, polyphenols  
78 and lipids. It promotes the production of flavour precursors, and notably amino-acids liberation through proteolysis  
79 that will further interact with sugars and polyphenols during the roasting process (John et al. 2016). The colour of  
80 the beans is also affected through the degradation of polyphenol, particularly anthocyanins, enhanced by acetic  
81 acid (De Taeye et al. 2016; Urbańska et al. 2019). It also prevents seed germination by participating in the death  
82 of the embryo, coupled with the temperature increase due to AAB exothermic metabolism. Acetic acid is further  
83 eliminated by sun drying (Hansen et al. 1998).

84 Interspecific diversity of AAB involved in cocoa fermentation has been well documented using culture dependent  
85 and independent methods. All studies revealed that *Acetobacter* is the most prevalent genus, followed by two other  
86 genera, *Gluconobacter*, *Gluconacetobacter* (Schwan and Wheals 2004; Camu et al. 2007). The major species  
87 found are *Acetobacter pasteurianus*, *Acetobacter tropicalis*, *Acetobacter senegalensis* and *Acetobacter syzygii*  
88 (Camu et al. 2008; Papalexandratou et al. 2011a, 2011b; Pereira et al. 2012; Crafack et al. 2013; Pereira et al.  
89 2013a, 2013b; Bortolini et al. 2016). Other species are reported in different countries such as *Acetobacter malorum*  
90 and *Acetobacter fabarum* in Brazil (Papalexandratou et al. 2011b; Pereira et al. 2013a, 2013b), and *Acetobacter*  
91 *lovaniensis* in Ghana and Ivory Coast (Camu et al. 2007; Hamdouche et al. 2015). However, *A. pasteurianus*  
92 dominates fermentations among all producing countries (Ardhana 2003; Camu et al. 2007, 2008; Papalexandratou  
93 et al. 2011a, 2011b, 2011c, 2013; Hamdouche et al. 2015; Bortolini et al. 2016). Although *A. pasteurianus* has  
94 been frequently isolated and appears to have a key role in the fermentation process, to date no study has reported  
95 a potential intraspecific diversity according to their country of origin.

96 Some studies have reported the interspecific diversity of AAB using (GTG)<sub>n</sub> fingerprinting molecular method  
97 based on the analysis of genomic polymorphisms (Camu et al. 2007, De Vuyst et al. 2008). This method allows  
98 the identification of isolates at the species level and could be used for intraspecific diversity analysis. A  
99 physiological characterization of AAB strains isolated in Ivory Coast has been reported by Visintin et al. (2016).  
100 Optimal growth temperature was between 25 and 30°C, and most of the strains could grow at temperatures up to  
101 45°C, which is the temperature usually obtained during the last days of cocoa fermentation. High concentration of

102 glucose or fructose (30%) reduce the growth of AAB due to osmotic stress, and resistance to acetic acid and lactic  
103 acid concentration (5%) is strain dependent. Among AAB species, most *A. pasteurianus* strains show physiological  
104 traits particularly well suited to the harsh conditions found in cocoa fermentation (Visintin et al. 2016). This species  
105 is highly tolerant to temperature, ethanol and acetic acid in low pH environments (Illegghems et al. 2013). Lefeber  
106 et al. (2010) studied the metabolism of three *A. pasteurianus* strains representative of different sub clusters  
107 discriminated by (GTG)<sub>5</sub>-PCR banding patterns in a previous study (Camu et al. 2007). They compared their  
108 growth kinetics, substrate consumption, and metabolites production and highlighted substantial differences  
109 regarding their metabolism. However, this study only focused on strains isolated in the same country.

110 In cocoa fermentation, *Acetobacter* and more specifically *A. pasteurianus* strains use lactic acid and ethanol (Camu  
111 et al. 2007, 2008) and display a separate metabolism during co-consumption of these two substrates. While ethanol  
112 oxidation leads mostly to acetic acid production, lactic acid is mainly degraded through the tricarboxylic acid  
113 (TCA) cycle via pyruvate, and provide a carbon source for biomass precursors, notably for amino acids synthesis.  
114 A substantial part of lactic acid is also oxidized to acetoin, and a smaller part to acetic acid production (Adler et  
115 al. 2014); The relative consumption of these substrates and their use by AAB strains of different origins can provide  
116 interesting criteria for the selection of efficient starter cultures.

117 As cocoa bean fermentation is a spontaneous process, starters composed of strains from the microbial consortium  
118 of cocoa (yeasts, LAB and AAB) have been developed to control and improve cocoa fermentation (Lefeber et al.  
119 2010; 2011; Crafacck et al. 2014; Moens et al. 2014; Magalhães da Veiga Moreira et al. 2017). Inoculation of a  
120 microbial cocktail isolated from Ghana and composed of *Saccharomyces cerevisiae* H5S5K23, *Lactobacillus*  
121 *fermentum* 222 and *A. pasteurianus* 386B, enhanced the cocoa bean fermentation process and improved the final  
122 chocolate quality (Lefeber et al. 2012). A similar result was obtained in Brazil, by inoculation of *S. cerevisiae*  
123 UFLA CCMA 200, *Lactobacillus plantarum* CCMA 0238 and *A. pasteurianus* CCMA 0241 strains (Magalhães  
124 da Veiga Moreira et al. 2017). The analysis of the intraspecific diversity of strains isolated from different countries  
125 could be of interest for the development of indigenous starters. It would allow the maintenance of the biodiversity  
126 of each producing country by avoiding the use of microorganisms from foreign countries. Moreover, it might help  
127 to avoid standardization of cocoa beans fermentation all around the globe.

128 **The aim of our study was to analyze the intraspecific diversity of *A. pasteurianus* strains isolated from cocoa**  
129 **fermentation** in three different countries. Genomic polymorphism analysis was performed using (GTG)<sub>5</sub> rep-PCR  
130 fingerprinting and the behavior of AAB strains was studied by analyzing the target metabolites in a medium that

131 simulate cocoa pulp, considering their specific growth rate, ethanol and lactic acid consumption, and acetic acid  
132 production.

## 133 **Material and methods**

### 134 **Cocoa beans sampling**

135 Cocoa beans fermentation sampling was performed in different countries: Mexico (MX), Ivory Coast (CI) and  
136 French Guiana (GF) (Table 1). Samplings were performed making sure to respect local practices from each  
137 producing country. In Mexico, two cocoa bean samplings were realized during the main crop (February 2018) and  
138 during the mid-crop (October 2018) respectively, in Comalcalco and Tabasco. For the main crop, sampling was  
139 realized at 2, 4 and 5 days of fermentation, and for the mid-crop all fermentation days were sampled **from days 0**  
140 **to 4**. Two campaigns of sampling were performed in Ivory Coast: the first in 2016 in Wagana (Fermentation Center  
141 of Cemoi) where sampling was performed **daily from day 0 to day 6**. The second sampling was done in Abidjan  
142 in 2018 where cocoa beans were sampled **daily from 0 to 7**. In addition, one cocoa bean fermentation realized in  
143 French Guiana was sampled **daily from 0 to 7**.

### 144 **Acetic acid bacteria strains isolation**

145 Samples were withdrawn from the fermenting cocoa bean mass for microbiological analysis. A specific medium  
146 for AAB, Kneifel medium was used (OIV, 2010), containing ethanol (20 mL/L), an antifungal (Cycloheximide,  
147 Sigma-Aldrich, 100 mg/L) and an antibiotic (Penicillin G, Sigma-Aldrich, 12.5 mg/L) that prevent the growth of  
148 yeast and LAB. This medium also contained a pH indicator (Bromocresol green, Sigma-Aldrich) which turned  
149 from blue to yellow /green when the medium acidified. *Acetobacter pasteurianus* LMG 01262<sup>T</sup> was used as  
150 reference strain. It was purchased from the BCCM/LMG Bacteria collection (Ghent, Belgium). Recommendations  
151 of the BCCM/LMG were followed for its culture and conservation (<http://bccm.belspo.be/>). All strains were kept  
152 frozen at -80°C in CPSM-AAB supplemented with 33% (v/v) glycerol. **CPSM-AAB is a specific liquid medium**  
153 **developed by Moens et al. (2014) called Cocoa Pulp Simulation Medium for AAB (CPSM-AAB). It reproduced**  
154 **the conditions found after two days of fermentation necessary for the development of AAB. It is composed of 1%**  
155 **ethanol (Honeywell); 10 g/L lactic acid (Sigma-Aldrich); 10 g/L yeast extract (Becton Dickinson), 5 g/L soya**  
156 **peptone (Merck GaA), and 1 mL/L Tween 80 (PanReac) at pH 4.5.**

### 157 **Bacterial species identification**

158 Gram stain was performed on colonies with a yellow / green halo on Kneifel medium. Then Gram-negative  
159 ellipsoidal bacilli were identified by 16S rRNA gene sequencing. A rapid DNA extraction was performed as  
160 described previously (Teyssier et al. 2003). Then the 16S rRNA gene was amplified by PCR using the primer pair  
161 27F (5'-GTGCTGCAGAGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-  
162 CACGGATCCTACGGGTACCTTGTTACGACTT-3'). The size of the expected amplicon was about 1500 bp.  
163 PCR was performed in 50 µL reaction volume containing GoTaq® reaction buffer (1X), 2 µM of each dNTP, 0.2  
164 µM of each primer, 1.25 of GoTaq® G2 DNA polymerase (Promega) and 5-40 ng of DNA. Amplification was  
165 carried out under the following conditions: an initial denaturation step at 94°C for 3 min followed by 35 cycles of  
166 at 94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec and a final elongation step at 72°C for 5 min (Tulini et al.  
167 2013). Control of amplification was performed on 0.8% agarose gels in 1 X TAE [Tris acetate-EDTA, Euromedex]  
168 at 100 V for 30 min. Then the PCR products were sequenced according to the Sanger method (Genseq,  
169 Montpellier). Sense or antisense strand DNA sequences were analyzed and corrected with Bioedit Sequence  
170 Alignment Editor (version 7.0.5.3 - Applied Biosystems USA). The software Geneious (version 2019.0.4-  
171 Invitrogen Corporation, NZ) allowed us to obtain a complete sequence. The sequence was compared with Genbank  
172 using Blastn algorithm (Altschul 1997) on the NCBI website (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

### 173 **Genomic diversity analysis by rep-PCR**

174 The genomic polymorphism of the AAB strains was analyzed by (GTG)<sub>5</sub> amplification of GTG repeated sequences  
175 (5' GTGGTGGTGGTGGTG 3') as previously described by Versalovic et al. (1991, 1995) and Mercier et al.  
176 (1996). PCR mixtures contained AmpliTaq Gold™ 360 Master Mix reaction buffer (1X, ThermoFisher®), 2 µM  
177 of primer, 3.5% DMSO (Sigma-Aldrich) and 1 µL DNA (10 ng/µL) in a final volume of 50 µL. Amplification  
178 was carried out under the following conditions: an initial denaturation step at 95°C for 3 min followed by 30 cycles  
179 at 90°C for 30 sec, 42°C for 30 sec, 65°C for 4 min and a final elongation step at 65°C for 8 min (Gonzalez et  
180 al. 2004). Migration of PCR amplicons was performed on 1.5% agarose gels containing 1 x TAE at 100 volts for  
181 2 h, and stained hereafter by DNA's intercalant (GelRed, Biotium) for 1 h. PCR products were loaded with size  
182 markers at both extremities of the gel in order to facilitate normalization of migration distances. Gels were  
183 visualized and photographed on a 312 nm UV transilluminator with a digital camera controlled by the software  
184 Gel Smart 7.3 (Clara Vision, France). Amplifications were realized in duplicates and PCR products were deposited  
185 randomly in the gels to eliminate a potential gel variation effect. Normalization of migration distance (Rf) between  
186 electrophoresis gels was achieved using the CLIQS software (version 1.2.0.044, Totallab, UK). The (GTG)<sub>5</sub> PCR

187 patterns were visually compared, and the statistical analyses were performed with Rstudio (Version 1.1.453 – ©  
1 2009-2018 RStudio, Inc, USA). To generate (GTG)<sub>5</sub> PCR patterns for each strain, only bands which were shared  
2 188 between the duplicates amplifications were kept, and used for statistical analyses. Distance matrices were  
3 189 calculated according to the Sørensen-Dice index (band presence or absence), a dendrogram was generated and a  
4 190 permutational multivariate analysis of variance (PERMANOVA) was performed.  
5  
6  
7  
8  
9

## 10 191 **Biochemical characterization**

### 11 192 **Determination of growth kinetics**

12  
13  
14  
15 193 Strains were transferred from - 80°C stock cultures to Kneifel agar medium and incubated aerobically at 30°C for  
16 194 48 h. Subsequently, the strains were cultivated overnight at 32°C to obtain final pre-cultures in CPSM-AAB.  
17 195 Absorbance were measured at 600 nm after the strains had been placed in sterile 96-well plates in a final volume  
18 196 of 200 µL and an initial optical density (OD) of 0.01. Growth kinetics was monitored by absorbance measurement  
19 197 at 600 nm with a microplate reader (TECAN). The plates were stirred at 150 rpm for 600 sec and, OD at 600 nm  
20 198 (OD<sub>600</sub>) was monitoring automatically every 10 min 30 sec for 20 h. Specific growth rate was calculated by  
21 199 extracting the slope of  $\ln(\text{OD}_{600})=f(\text{time})$  and taking the maximal value.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32

### 33 201 **Quantification of metabolites consumption and production**

34 202 After an overnight pre-culture at 32°C in CPSM for AAB (pH 4.5), OD was measured, and cultures were developed  
35 203 in CPSM-AAB (pH 4.5) at 32°C without agitation at an initial OD of 0.02. Sampling was performed at 18 h and  
36 204 40 h and only samples having an OD<sub>600</sub> < 3 were selected, because the main substrate ethanol was not exhausted,  
37 205 in order to determine the relative consumption of lactic acid, mannitol and ethanol without bias.  
38  
39  
40  
41  
42  
43  
44  
45

46 206 After a vigorous shaking, 2 mL of cell suspension was removed from the cell culture tubes and centrifuged at  
47 207 10000 rpm for 10 min. Supernatants were filtered through regenerated cellulose membrane with pore diameter of  
48 208 0.45 µm then kept frozen at -20°C in glass vials until HPLC analysis. OD<sub>600</sub> were measured for each strain in  
49 209 order to evaluate biomass production. Residual substrates (ethanol, lactic acid and mannitol) and metabolites  
50 210 (acetic acid) were quantified by high performance liquid chromatography (HPLC). HPLC analysis was performed  
51 211 using an Aminex HPX-87H column (Biorad, Hemel Hempstead, UK, 300 x 7.8 mm) and detection by both  
52 212 refractive index (for ethanol and mannitol) and UV absorption at 210 nm (lactic and acetic acid). Elution was  
53 213 performed at 30°C with 5 mM sulphuric acid at a flow rate of 0.6 mL/min. Biomass yield was calculated as  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

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

214 followed:  $Y_{X/S}$  where X is the biomass (OD) and S substrates (ethanol and lactic acid) consumption in mM (value  
215 by HPLC). The conversion ratio of ethanol into acetic acid ( $Y_{AA/EtOH}$ ) was calculated by dividing the moles of  
216 acetic acid produced by the moles of ethanol consumed. A conversion of 1 mole of acetic acid by 1 mole of ethanol  
217 ( $Y_{AA/EtOH} = 1$ ) was considered as the maximal stoichiometric ratio. Biomass yield, determined by acetic acid  
218 produced, was defined as followed:  $Y_{X/AA, S}$  where X is the biomass (OD600) and AA is the acetic acid production  
219 in mM. The molar ratio ( $R_{LA/EtOH}$ ) of lactic acid and ethanol consumption was also calculated in order to determine  
220 the relative affinity for ethanol and lactic acid as carbon sources for their growth and/or to produce acetic acid.

## 221 **Statistical analysis**

222 All statistical analyses were performed using XLSTAT software version 2019.1.3 (Addinsoft, France) and Rstudio  
223 (Version 1.1.453 – © 2009-2018 RStudio, Inc, USA). Box plots were developed in Excel and dendrogram,  
224 heatmap in Rstudio. Preprocessing of the HPLC data consisted of normalization on the biomass product (OD) to  
225 correct inter-strain differences in consumption in lab-scale fermentation experiments. Then conversion to Z-scores  
226 to correct noise was performed, calculated as followed:  $Z\text{-score} = (X - \mu) / \sigma$ , where X is the value measurement,  $\mu$   
227 is the mean value of all strains, and  $\sigma$  is the standard deviation of values of all strains. Cluster analysis was  
228 performed using Euclidean distance for building the matrix and UPGMA algorithm (Steensels et al. 2014, Brase  
229 2017). Statistical analysis tests were Shapiro-Wilk (normality test) and Kruskal- Wallis (one-way analysis of  
230 variance).

## 231 **Results**

### 233 **Identification of the acetic acid bacteria strains by 16S rRNA gene sequencing**

234 One hundred thirty Gram-negative ellipsoidal bacilli were isolated on Kneifel medium. All strains were identified  
235 as acetic acid bacteria by 16S rRNA gene sequencing with a percentage of identity higher than 98% (Table S1).

236 These AAB strains were affiliated to *Acetobacter* (n= 122) and to *Gluconobacter* (n=8) genera (Table 2). Among  
237 *Acetobacter* strains, *A. pasteurianus* was the most recovered species with 73,4% of the isolates, followed by *A.*  
238 *tropicalis*, *Acetobacter ghanensis*, *A. fabarum* and *Acetobacter orientalis*, which represented respectively 13%,  
239 5.6%, 1.6%, and 1.6% of the total *Acetobacter* isolates. Six *Acetobacter* strains could not be affiliated at the species  
240 level, because *Acetobacter cerevisiae* or *A. malorum* were very close genetically and 16S rRNA gene sequence

241 was not sufficient to differentiate them (Valera et al. 2013). Among *Gluconobacter* strains, *Gluconobacter oxydans*  
242 was the most encountered species with 87.5% and *Gluconobacter nephelii* with 12.5% of the total of  
243 *Gluconobacter* strains. Distribution of acetic acid bacteria strains according to the origin's countries shows that *A.*  
244 *pasteurianus* and *A. tropicalis* strains were distributed over the 3 countries, while *A. fabarum* and *A. ghanensis*  
245 strains were isolated only in Ivory Coast, *A. orientalis* in French Guiana and *A. cerevisiae/malorum* strains in  
246 Mexico. *Gluconobacter oxydans* strains were isolated in Ivory Coast and French Guiana and *Gluconobacter*  
247 *nephelii* only in Ivory Coast (Table 2).

### 248 **Genomic polymorphism analysis of acetic acid bacteria by (GTG)<sub>5</sub> PCR**

249 According to the results of strain identification by 16S rRNA gene sequencing, *A. pasteurianus* was the most  
250 common species among AAB isolates. Thus, we further analyzed the genomic polymorphism of 89 *A.*  
251 *pasteurianus* strains using one single oligonucleotide primer (GTG)<sub>5</sub> with 3 other AAB strains (*A. tropicalis*, *A.*  
252 *fabarum*, *G. oxydans*). Strains used for this analysis were numbered from 1 to 92, with name code CI, MX and GF  
253 for Ivory Coast, Mexico and French Guiana respectively and Genbank accession number from MN909060 (CI1)  
254 to MN909151 (MX92) (Table S1). *A. pasteurianus* type strain LMG 01262<sup>T</sup> was also included for the analysis.  
255 The name code for the 3 AAB from other species was: *A. tropicalis* MX92, *A. fabarum* CI25 and *G. oxydans*  
256 GF51.

257 After amplification and gel electrophoresis, banding patterns contained between 7 to 23 DNA fragments ranging  
258 from 350 to 3000 bp. Visually, all banding patterns of *A. pasteurianus* strains shared several bands at 350, 380,  
259 480, 980 and 1100 bp (data not shown). The results of the numerical and statistical analysis of the generated  
260 (GTG)<sub>5</sub> PCR banding patterns are shown in the dendrogram (**Fig. 1**).

261 All *A. pasteurianus* strains were clustered into 7 groups (I to VII, Table 3) having a minimum of 55% of similarity  
262 and all clusters contained strains from at least 2 countries. The majority of strains were clustered in 5 clusters with  
263 a variable number of isolates for each group: 6, 39, 4, 12 and 12 isolates for cluster II, III, IV, V and VI  
264 respectively, with a similarity between clusters of 60% (**Fig. 3**). In *A. pasteurianus* clusters, most of the strains  
265 seems to be genetically closed, with a percentage of similarity around 80% indicating a low intraspecific genomic  
266 polymorphism. Furthermore, three strains pairs obtained 100% similarity: MX75 and MX76 isolated in Mexico in  
267 the same fermentation time; MX90 and MX52 isolated in Mexico at two different fermentation times (4 and 5  
268 days respectively) and MX78 and MX67 isolated in Mexico at the same 3 days of fermentation but in 2 different

269 batches. Strains CI12 and CI2 were isolated at two fermentation times presented a similarity at 95%. Strains CI40  
1 270 and CI21, isolated at the same time of fermentation but in two different batches had also a similarity of 95%.  
2  
3 271 Regarding these strains with high similarities (95%), we can assume that they were the same strains isolated twice.  
4  
5 272 On the other hand, others strain pairs, isolated in different countries, were clustered with a high percentage of  
6  
7 273 similarity 95%, 95% and 90% for CI27/GF42, MX61/GF41 and MX84/CI34 respectively. In addition, strain CI28  
8  
9 274 was clustered with the strains pair MX61/GF41 at 90% of similarity. All these results indicated that strains could  
10  
11 275 be found in two or three different countries.  
12  
13 276 Strains grouped in cluster VII shown 55% of similarity with the other cluster. It was composed of 2 strains, one  
14  
15 277 isolated in Mexico (MX70) and one in Ivory Coast (CI8) and strains in cluster I, shown 58% of similarity with the  
16  
17 278 other strains. Cluster I was composed by 8 isolates from Mexico, 6 from Ivory Coast and 1 from French Guiana.  
18  
19 279 As expected, the 3 non *A. pasteurianus* strains; CI25, MX92 and GF51 were clustered far from *A. pasteurianus*  
20  
21 280 strains and showed less than 35% of similarity. *A. fabarum* being the closest relative and *G. oxydans* the most  
22  
23 281 distant to *A. pasteurianus* strains.  
24  
25 282 In order to evaluate the influence of the strains' origin countries or fermentation day of strain isolation on the  
26  
27 283 clustering, statistical tests were performed by analyzing variance-using distance matrices created with Sørensen-  
28  
29 284 Dice index. It appeared that origin country had a strong effect on the clustering (p-value=0.001\*\*) while  
30  
31 285 fermentation day of strain isolation had not (p-value > 0.05). This effect was explained by strains with high  
32  
33 286 percentage of similarity (>90%), which came from the same origin.  
34  
35  
36

## 287 **Comparative analysis of *A. pasteurianus* strains' metabolism in CPSM-AAB**

38  
39  
40

41 288 In order to investigate the intraspecific diversity of the 89 strains of *A. pasteurianus* isolated from three regions  
42  
43 289 of the world, their basic metabolism was studied in a medium simulating the cocoa pulp conditions during the  
44  
45 290 fermentation process developed by Moens et al. (2014). Different parameters were calculated from substrate  
46  
47 291 consumption and acetic acid production demonstrated by HPLC: biomass yield ( $Y_{X/S}$ ) to determine biomass  
48  
49 292 productivity, production yield of ethanol into acetic acid ( $Y_{AA/EtOH}$ ) and biomass yield produced by acetic acid  
50  
51 293 produced ( $Y_{X/AA}$ ). Lactic acid on ethanol ratio ( $R_{LA/EtOH}$ ) was also calculated to determine relative affinity for  
52  
53 294 ethanol and lactic acid as carbon source for their growth and/or to produce acetic acid. These different parameters  
54  
55 295 were used together with specific growth ( $\mu$ ) rate to compare the strains. An overview is showed in the dendrogram  
56  
57 296 (Fig. 2).  
58  
59  
60  
61  
62  
63  
64  
65



## 297 Dendrogram and heatmap analysis

1  
2  
3 298 Strains were compared by using their Z-scores. According to the histogram of the Z-scores data (**Fig. 2**), most of  
4  
5 299 the Z-score data were grouped between -1 and 1 indicating a strong similarity between strains. In order to evaluate  
6  
7 300 the influence of geographical origin and fermentation time on the strains clustering, statistical tests were performed  
8  
9 301 by variance analysis of Euclidean distance matrices. It appeared that the origin country had a significant effect on  
10  
11 302 the clustering (p-value=0.028\*) while the fermentation time had not (p-value>0.05).

12  
13  
14 303 *A. pasteurianus* strains were clustered into 2 large clusters with a strong similarity between strains (Euclidean  
15  
16 304 distance less than 2 on a scale of 8) (**Fig. 2**). The repartition of the strains according to their origin in each cluster  
17  
18 305 is detailed in Table 4. Cluster I was composed of 45 strains with 22 isolates from Ivory Coast (56.5 % of Ivorian  
19  
20 306 strains), 17 isolates from Mexico (42.5% of Mexican strains), 4 strains from French Guiana (40% of Guyanese  
21  
22 307 strains) *A. pasteurianus* LMG01262<sup>T</sup> and *G. oxydans* strain (GF51). Moderates  $Y_{X/S}$ ,  $Y_{X/AA}$ , and conversion  
23  
24 308 efficiency of ethanol into acetic acid were observed for the strains belonging to this cluster. Cluster II was  
25  
26 309 composed of 45 strains, with 22 from Mexico (55% of Mexican strains), 17 isolates from Ivory Coast (43.5% of  
27  
28 310 Ivorian strains), and 6 from French Guiana (60% of Guyanese strains). Strains in this cluster were characterized  
29  
30 311 by higher  $Y_{AA/EtOH}$ ,  $R_{LA/EtOH}$ ,  $Y_{X/S}$ , and  $Y_{X/AA}$  to the other strains. Strain MX61, was not clustered with other *A.*  
31  
32 312 *pasteurianus* and presented the highest  $Y_{AA/EtOH}$  with low  $Y_{X/S}$ , and  $Y_{X/AA}$  value compared to other strains. *A.*  
33  
34 313 *fabarum* CI25 and *A. tropicalis* MX92 were not clustered with *A. pasteurianus* strains. *A. fabarum* had a lower  
35  
36 314 conversion efficiency of ethanol in acetic acid than *A. tropicalis*. The latter is characterized by a high  $R_{LA/EtOH}$  and  
37  
38 315  $Y_{X/AA}$  compared to *A. pasteurianus* strains.

## 42 316 Comparison of molecular consumption and production between clusters

43  
44  
45 317 Most of the strains did not or negligibly consume the mannitol present in the medium in our experimental  
46  
47 318 conditions (0 to 10 mM/ Optical Density Units, ODU, data not shown). Thus, it was not considered in the  
48  
49 319 comparative analysis of clusters. In order to identify parameters which had a significant effect within each cluster  
50  
51 320 a two-by-two analysis was performed. All selected parameters except for specific growth rate participated in the  
52  
53 321 clustering (p-value < 0.0001\*\*\*). Strains from cluster I exhibited lower values for  $Y_{X/S}$  (median= 9.38 ODU/mol)  
54  
55 322 and  $Y_{X/AA}$  (median= 17.5 ODU/mol) than strains in cluster II with  $Y_{X/S}$  (median= 19.35 ODU/mol) and  $Y_{X/AA}$   
56  
57 323 (median= 31 ODU/mol) (**Fig. 3**). Both clusters had a very high conversion efficiency of ethanol into acetic acid  
58  
59 324 ( $Y_{AA/EtOH}$ ), cluster I=0.83 and cluster II=0.94-median value, clusterII had a higher  $Y_{AA/EtOH}$ , explained by strains

325 which had a value higher than 1 while in cluster I, strains had a value less than 0.5 (Fig. 3). Both clusters had a  
326 preference for ethanol than lactic acid, indeed, their  $R_{AL/EtOH}$  were less than 1. However, strains in cluster II had  
327 higher  $R_{LA/EtOH}$ , with a median value of 0.5 while cluster I had a median  $R_{LA/EtOH}$  of 0.31. The analysis of ethanol  
328 and lactic acid specific consumption of the 89 *A. pasteurianus* strains confirmed that they generally consumed  
329 more ethanol than lactic acid, ranging between  $25.58 \pm 3.49$  and  $132.15 \pm 8.34$  mM/ODU and  $7.5 \pm 0.79$  and  $56.71 \pm 5.07$   
330 mM/ODU respectively (mean value taken for each strain) (Fig. S1). Regarding the specific growth rate, the median  
331 value for each cluster was around  $0.6 \text{ h}^{-1}$ , and statistical tests showed a non-significant difference between clusters  
332 with a p-value of 0.091. Finally, 2 subclusters in cluster II, cluster II-A and cluster II-B would appear to have a  
333 higher preference for lactic acid with high  $Y_{AA/EtOH}$  cluster II-A= 1.16 and cluster II-B=0.9, median value for each  
334 sub cluster) (Fig. 3). Ratios  $R_{LA/EtOH}$  were higher than strains in cluster I (cluster II-A=0.61 and cluster II-B=0.5)  
335 with strains which had a  $R_{LA/EtOH}$  equal or close to 1, as MX90, CI19, MX81, these 3 strains had also a  $Y_{AA/EtOH}$   
336 close to or higher than 1. These sub clusters had also a great biomass productivity resulting in higher yield of acetic  
337 acid production. Cluster II-A was composed of 7 Mexican strains and 2 Ivorian strains and cluster II-B of 3  
338 Mexican strains, 2 Guyanese strains and 2 Ivorian strains.

## 339 Discussion

340 All around the world, cocoa bean fermentations are mostly performed in a traditional way using a variety of  
341 production methods. This artisanal process leads to variable organoleptic quality of the cocoa products. The  
342 development of a starter culture with mixed microorganisms has been proposed for successful fermentations for  
343 final cocoa product with better quality and flavor. However, only a few published studies compared the  
344 intraspecific diversity of *A. pasteurianus* strains to develop starters with biodiversity maintain (Camu et al. 2007;  
345 Lefeber et al. 2010). To preserve the bacterial biodiversity relative to a specific cocoa production area, it seemed  
346 interesting to further investigate the intraspecific diversity of strains from different producing countries for the  
347 elaboration of starter cultures.  
348 Our results showed that *A. pasteurianus* was the most recovered species among 130 AAB isolates, representing  
349 69% of the strains. This observation confirms those findings where the mains AAB species involved in the  
350 spontaneous cocoa beans fermentations appeared to belong to *A. pasteurianus*, perfectly suited for the cocoa  
351 environment regarding its ability to oxidize most substrate (ethanol, lactic acid and mannitol) and its temperature  
352 and pH tolerance (De Vuyst and Weckx, 2016, Figueroa-Hernández et al. 2019). A first attempt to elucidate  
353 intraspecific diversity was performed using strains from one origin country, in Ghana (Camu et al. 2007; Lefeber  
354 et al. 2010). (GTG)<sub>5</sub> fingerprinting and analysis of their kinetics of growth, substrate consumption, and metabolite

355 production were performed. However, our study was a first attempt to elucidate intraspecific diversity of a large  
1 356 panel of *A. pasteurianus* strains according to their geographical origins with genomic and metabolic intraspecific  
2 357 diversity analysis. Genomic polymorphism analysis with (GTG)<sub>5</sub>-rep-PCR showed a great similarity between  
3 358 isolates from three countries and according to the statistical tests they were clustered according to their origin  
4 359 country.

9 360 Many studies have been performed aiming at elucidating genomics and genetics polymorphisms of *A. pasteurianus*  
10 361 strains involved in vinegar production. The chromosome sequence of 11 *A. pasteurianus* strains from vinegar  
11 362 fermentation and cocoa beans fermentation have been compared and showed a high degree of homology (Wang  
12 363 et al. 2015). These high homologies between chromosome of *A. pasteurianus* strains could explain the great  
13 364 similarities between (GTG)<sub>5</sub> pattern profiles obtained for isolates in our study. However, Illeghems et al (2013)  
14 365 found that 68% of the predicted proteins obtained by genome sequencing were shared between *A. pasteurianus*  
15 366 386B isolated in cocoa beans fermentation and 5 others *A. pasteurianus*. Moreover, 27 genes were specific to the  
16 367 386B strain, probably linked to its unique performance in cocoa fermentation process. These findings support the  
17 368 fact that high intraspecific metabolic diversity could be found among *A. pasteurianus* species.

22 369 The main carbon sources for *A. pasteurianus* strains are ethanol and lactic acid. In our study, the analysis of the  
23 370 consumption of ethanol, lactic acid and mannitol by several *A. pasteurianus* strains isolated in three different  
24 371 countries has been performed. This phenotypic characterization remains of basic relevance to screen strains in  
25 372 selection strategies providing data to cluster strains according to convenient metabolic profiles (Wu et al. 2010).  
26 373 Kinetics experiments regarding metabolites consumption performed in our study showed that mannitol was not  
27 374 consumed until ethanol was completely exhausted and lactate remained in low quantity (<2g/L) (data not shown).  
28 375 This condition was never achieved in the experiments considered for this analysis (all OD<sub>600</sub><3 and ethanol  
29 376 remained in the medium). It has already been shown that *A. pasteurianus* usually oxidizes sugar very slowly until  
30 377 ethanol exhaustion (Sakurai et al. 2011; Moens et al. 2014). Our experiments were performed on 89 strains and  
31 378 confirmed this behavior regarding mannitol consumption. All *A. pasteurianus* strains studied here seemed to  
32 379 exhibit a stable phenotype regarding this feature. Most analyzed strains in our study showed a higher affinity for  
33 380 ethanol than lactic acid, with R<sub>LA/EtOH</sub> lower than 1, confirming results of previous studies (Lefeber et al. 2010,  
34 381 Moens et al. 2014). However, 3 strains of sub clusters cluster II-A and cluster II-B, MX90, CI19 and MX81, had  
35 382 a high yield of ethanol to acetic acid transformation and appeared to have a higher affinity for lactic acid than other  
36 383 strains. These strains could be interesting as starter culture, since lactic acid is an undesirable metabolite in cocoa  
37 384 fermentation due to its unpleasant flavor (Holm et al. 1993). Furthermore, these strains had a Y<sub>AA/EtOH</sub> higher than

385 1, suggesting that ethanol consumption alone cannot explain acetic acid production. This finding highlights the  
1  
2 386 potential significant involvement of lactic acid in acetic acid production by some strains of *A. pasteurianus* and *A.*  
3  
4 387 *tropicalis*. The presence of the enzymes mandatory for the transformation of lactate to acetate through pyruvate  
5  
6 388 and acetaldehyde by *A. pasteurianus* has been described (Sato et al. 2015). Lactate has been shown to be mostly  
7  
8 389 degraded through the tricarboxylic acid cycle (TCA) and via its oxidation to acetoin, but poorly to acetate (Adler,  
9  
10 390 2014). However the balance between these two pathways has been investigated on a single *A. pasteurianus* strain.  
11  
12 391 Our results tends to show that this specificity may not be generalized to all *A. pasteurianus* strains. Strains in  
13  
14 392 cluster I generally had  $Y_{AA/EtOH}$ , lower than 1, with some strains being beyond 0.5. This finding suggests an  
15  
16 393 overoxidation of acetic acid produced from ethanol. Acetic acid overoxidation has been reported to be provoked  
17  
18 394 by an increased activity of the tricarboxylic acid (TCA) cycle enzymes and acetyl-CoA synthetase for energy  
19  
20 395 supply. However, among acetic acid bacteria, the mechanism of switching between incomplete oxidation and  
21  
22 396 assimilatory oxidation, and the control of energy and carbon metabolism remains poorly understood (Saeki et  
23  
24 397 al.1999; Sakurai et al.2011; Adler et al. 2014, Yang et al. 2019). As acetic acid is a desirable product for the  
25  
26 398 completeness of cocoa fermentation, acetic acid overoxidation should be avoided for AAB starter. For the three  
27  
28 399 other AAB, *G. oxydans* was clustered with *A. pasteurianus* strains with a moderate  $Y_{AA/EtOH}$  and lower  $Y_{X/AA}$ ,  
29  
30 400 explained by slower ethanol oxidation than *A. pasteurianus* strains (De Ley and Swings, 1984, De Muyneck et al.  
31  
32 401 2008). *A.tropicalis* had an  $Y_{X/AA}$  higher than *A. pasteurianus* strains; lactic acid and ethanol oxidation have been  
33  
34 402 reported to be effective in this strain (Pereira et al. 2012). *A. fabarum* strain had a lower conversion efficiency of  
35  
36 403 ethanol to acetic acid than other strains as demonstrated in a previous study (Moens et al. 2014).  
37  
38 404 The differences obtained by biochemical analysis could be explained by genetic instability of *A. pasteurianus*  
39  
40 405 species. Genetic instability is known to occur at very high frequencies due to spontaneous mutations that cause  
41  
42 406 loss or acquisition of various physiological and phenotypic properties (Azuma et al. 2009; Wang et al. 2015).  
43  
44 407 Genetic mechanisms which cause these instabilities are not fully understood but according to Beppu (1994) and  
45  
46 408 Matsutani et al. (2013) phenotypic modifications in ethanol oxidation can occur by transposase insertion of a  
47  
48 409 sequence element. Phenotypic variability of *A. pasteurianus* has been demonstrated, on strains found in solid-state  
49  
50 410 fermentation in vinegar technologies (Wu et al. 2010). The 27 specific genes identified for *A. pasteurianus* 386B  
51  
52 411 as compared to other *A. pasteurianus* strains found by Illegheems and co-workers (2013) also promote the idea that  
53  
54 412 a metabolic diversity can be encountered in this species.  
55  
56 413 Our data indicated that phenotypic information clustering was not related to genomic clustering of *A. pasteurianus*  
57  
58 414 strains. Analyses showed high genomic and phenotypic similarities between strains, however, multiparametric  
59  
60  
61  
62  
63  
64  
65

415 clustering resulted in different groups. Some strains which were 100% identical according to (GTG)<sub>5</sub> rep-PCR  
1  
2 416 were not clustered together in the dendrogram based on basic metabolism. This finding suggests that (GTG)<sub>5</sub>  
3  
4 417 fingerprinting might not be sufficient for the discrimination of intraspecific variability among strains of  
5  
6 418 *Acetobacter pasteurianus*, notably when selection regarding their technological potential has to be performed.  
7  
8 419

## 420 **Conclusion**

11  
12 421 Intraspecific analysis of strains isolated in three different countries around the world has been studied according  
13  
14 422 to their genomic polymorphism, and their consumption and production of main metabolites under cocoa-like  
15  
16 423 conditions. The great similarity between strains indicated a low intraspecific diversity.

17  
18 424 Genomic polymorphism was not correlated with the analysis of metabolic polymorphism, revealing that (GTG)<sub>5</sub>-  
19  
20 425 rep-PCR might not be sufficient for the differentiation of strains metabolic features. To gain further information  
21  
22 426 on AAB intraspecific diversity, a complementary method might be considered such as whole genome sequencing  
23  
24 427 (WGS) (Illegheems et al. 2013). That method could allow us to highlight subtle differences between similar strains  
25  
26 428 obtained with (GTG)<sub>5</sub> fingerprinting method and might help to establish correlations with the intraspecific  
27  
28 429 biochemical diversity.

30 430 High ethanol conversion efficiency into acetic acid and a high affinity for lactic acid could be used as selection  
31  
32 431 criteria for AAB starter selection. Strains from sub clusters cluster II-A and cluster II-B seems to be the most  
33  
34 432 promising candidates for an efficient conversion of carbon sources into acetic acid, with a higher affinity for lactic  
35  
36 433 acid than other strains. Furthermore, the production of acetoin from lactic acid might help in understanding the  
37  
38 434 diversity observed among *A. pasteurianus* species regarding their carbon metabolism. It will be interesting to  
39  
40 435 compare their metabolic pathways using transcriptomics methods in order to study the expression of the genes  
41  
42 436 involved in the conversion of lactate and ethanol into acetate (Sato et al. 2015) and coupled through the analysis  
43  
44 437 of metabolic fluxes using labelled <sup>13</sup>C isotope experiments (Adler et al. 2014).

46 438 Furthermore, strains characterized by distant metabolic profiles associated with yeast and LAB might be tested in  
47  
48 439 cocoa fermentation at the lab-scale in order to better understand the influence of *A. pasteurianus* metabolic  
49  
50 440 diversity on the final quality of cocoa product.  
51

## 52 441

## 53 442

## 54 443 **Acknowledgments**

55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

444 Data used in this work were partly produced through the GenSeq technical facilities of the « Institut des Sciences  
1  
2 445 de l'Evolution de Montpellier » with the support of LabEx CeMEB, an ANR "Investissements d'avenir" program  
3  
4 446 (ANR-10-LABX-04-01). The authors are grateful to Frenchoc Premium Project for support samples. We would  
5  
6 447 like to thank Pr Guehi Tagro from Nangui Abrogoua University, Abidjan –Ivory Coast, Dr Suárez Mirna Leonor  
7  
8 448 and Hernández Medina Samuel from the technological institute of Veracruz - Mexico for providing the cocoa  
9  
10 449 beans samples

11  
12 450

### 14 451 **Disclosure of potential conflicts of interest:**

16  
17 452 The authors declare that they have no conflict of interest.

### 19 453 **Funding:**

21  
22 454 Sampling in Mexico was supported by DUFRENOY - Crédit Agricole d'Ile de France Mécenat grant.

### 24 455 **Research involving Human Participants and/or Animal:**

26  
27 456 This article does not contain any studies with human participant or animal performed by any of the authors.

### 29 457 **Informed consent:**

31  
32 458 N/A

### 34 459 **Figure caption:**

36  
37 460 **Fig. 1** Dendrogram generated after statistical analysis of the similarity of the (GTG)<sub>5</sub> PCR banding patterns of the  
38 461 89 *Acetobacter pasteurianus*, 3 AAB; *Acetobacter fabarum*, *Acetobacter tropicalis* and *Gluconobacter oxydans*  
39  
40 462 strains isolated from spontaneous cocoa fermentations in three areas around the world and one *A. pasteurianus*  
41 463 type strain (LMG0162T). The Sørensen-Dice index was used to analyze similarity and a UGPMA clustering  
42  
43 464 algorithm was applied to cluster the data. Tree branches were colored according strain origin , Mexico (blue),  
44 465 Ivory Coast (black), French Guiana (red), type strain (brown), the three other AAB species in green

46  
47 466 **Fig. 2** Biochemical relatedness of the 89 *Acetobacter pasteurianus* strains, *A. pasteurianus* type strain and the 3  
48 467 AAB. Measurements were converted in Z-scores, pairwise similarities were calculated by Euclidean distance and  
49  
50 468 an UGPMA clustering algorithm was applied to cluster the data. With  $Y_{EtOH/AA}$ = conversion efficiency of ethanol  
51 469 into acetic acid,  $R_{LA/EtOH}$ = correspond to the lactic acid consumption divided by the ethanol consumption,  $Y_{X/S}$ =  
52  
53 470 biomass yield,  $Y_{X/AA}$ = Biomass yield produced by acetic acid produced and specific growth rate ( $\mu$ ). Where X was  
54 471 the biomass, S = substrates consumption, LA= lactic acid, AA = acetic acid. EtOH= ethanol. Tree branches colors  
55  
56 472 codes are identical to those on **Fig. 1**

57  
58 473 **Fig. 3** Comparison in boxplot of physiological traits of *Acetobacter pasteurianus* strains from different  
59  
60 474 geographical origins and one type strain according to the two large clusters. With  $Y_{EtOH/AA}$ = conversion efficiency

475 of ethanol into acetic acid,  $R_{LA/EtOH}$  correspond to the lactic acid consumption divided by the ethanol  
1 476 consumption,  $Y_{X/S}$  = biomass yield,  $Y_{X/AA}$  = Biomass yield produced by acetic acid produced and specific growth  
2 477 rate in  $h^{-1}$ . Where X was the biomass, S = substrates consumption, LA= lactic acid, AA = acetic acid. EtOH=  
3 478 ethanol  
4  
5  
6

## 7 479 **References**

- 8  
9 480 Adler P, Frey LJ, Berger A et al (2014) The key to acetate: metabolic fluxes of acetic acid bacteria under cocoa  
10 481 pulp fermentation-simulating conditions. Appl Environ Microbiol 80:4702–4716.  
11 482 <https://doi.org/10.1128/AEM.01048-14>  
12  
13  
14 483 Altschul S (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.  
15 484 Nucleic Acids Res 25:3389–3402. <https://academic.oup.com/nar/article/25/17/3389/1061651>  
16  
17  
18 485 Ardhana M (2003) The microbial ecology of cocoa bean fermentations in Indonesia. Int J Food Microbiol 86:87–  
19 486 99. [https://doi.org/10.1016/S0168-1605\(03\)00081-3](https://doi.org/10.1016/S0168-1605(03)00081-3)  
20  
21  
22 487 Azuma Y, Hosoyama A, Matsutani M et al (2009) Whole-genome analyses reveal genetic instability of  
23 488 *Acetobacter pasteurianus*. Nucleic Acids Res 37:5768–5783. <https://doi.org/10.1093/nar/gkp612>  
24  
25  
26  
27 489 Barel M (2013) Qualité du cacao: L’impact du traitement post-récolte. Quae éditions 21-63. ISBN 978-2-7592-  
28 490 2052-6 2-7592-2052-4  
29  
30  
31 491 Beppu T (1994) Genetic organization of *Acetobacter* for acetic acid fermentation. Antonie van Leeuwenhoek  
32 492 64:121–135. <https://doi.org/10.1007/BF00873022>  
33  
34  
35 493 Bortolini C, Patrone V, Puglisi E, Morelli L (2016) Detailed analyses of the bacterial populations in processed  
36 494 cocoa beans of different geographic origin, subject to varied fermentation conditions. Int J Food  
37 495 Microbiol 236:98–106. <https://doi.org/10.1016/j.ijfoodmicro.2016.07.004>  
38  
39  
40  
41 496 Brase CH (2017) Understanding basic statistics. Cengage Learning, Inc. 8th edition. ISBN 978-1337558075  
42  
43 497 Camu N, De Winter T, Verbrugghe K et al (2007) Dynamics and biodiversity of populations of lactic acid bacteria  
44 498 and acetic acid bacteria involved in spontaneous heap fermentation of cocoa beans in Ghana. Appl  
45 499 Environ Microbiol 73:1809–1824. <https://doi.org/10.1128/AEM.02189-06>  
46  
47  
48  
49 500 Camu N, Gonzalez A, De Winter T et al (2008) Influence of turning and environmental contamination on the  
50 501 dynamics of populations of lactic acid and acetic acid bacteria involved in spontaneous cocoa bean heap  
51 502 fermentation in Ghana. Appl Environ Microbiol 74:86–98. <https://doi.org/10.1128/AEM.01512-07>  
52  
53  
54 503 Cleenwerck I, De Vos P, De Vuyst L (2010) Phylogeny and differentiation of species of the genus  
55 504 *Gluconacetobacter* and related taxa based on multilocus sequence analyses of housekeeping genes and  
56 505 reclassification of *Acetobacter xylinus* subsp. *sucrofermentans* as *Gluconacetobacter sucrofermentans*  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 506 (Toyosaki et al.1996) sp. nov., comb. nov. Int J Syst Evol Microbiol 60:2277–2283.  
1 507 <https://doi.org/10.1099/ijs.0.018465-0>  
2  
3  
4 508 Crafacck M, Keul H, Eskildsen CE et al (2014) Impact of starter cultures and fermentation techniques on the volatile  
5 509 aroma and sensory profile of chocolate. Food Res Int 63:306–316.  
6 510 <https://doi.org/10.1016/j.foodres.2014.04.032>  
7  
8  
9 511 Crafacck M, Mikkelsen MB, Saerens S et al (2013) Influencing cocoa flavour using *Pichia kluyveri* and  
10 512 *Kluyveromyces marxianus* in a defined mixed starter culture for cocoa fermentation. Int J Food Microbiol  
11 513 167:103–116. <https://doi.org/10.1016/j.ijfoodmicro.2013.06.024>  
12  
13  
14  
15 514 Daniel H-M, Vrancken G, Takrama JF et al (2009) Yeast diversity of Ghanaian cocoa bean heap fermentations.  
16 515 FEMS Yeast Res 9:774–783. <https://doi.org/10.1111/j.1567-1364.2009.00520.x>  
17  
18  
19 516 De Ley J. and Swings J (1984) Genus *Gluconobacter*. In: Bergey's Manual of Systematic Bacteriology. Krieg, N.,  
20 517 and Holt, J., Eds., Williams and Wilkins, London, pp. 275–278  
21  
22  
23 518 De Muynck C, Pereira CSS, Naessens M et al (2007) The Genus *Gluconobacter oxydans* : Comprehensive  
24 519 overview of biochemistry and biotechnological applications. Crit Rev Biotechnol 27:147–171.  
25 520 <https://doi.org/10.1080/07388550701503584>  
26  
27  
28  
29 521 De Roos J, De Vuyst L (2018) Acetic acid bacteria in fermented foods and beverages. Curr Opin Biotechnol  
30 522 49:115–119. <https://doi.org/10.1016/j.copbio.2017.08.007>  
31  
32  
33 523 De Vuyst L, Camu N, De Winter T et al (2008) Validation of the (GTG)<sub>5</sub>-rep-PCR fingerprinting technique for  
34 524 rapid classification and identification of acetic acid bacteria, with a focus on isolates from Ghanaian  
35 525 fermented cocoa beans. Int J Food Microbiol 125:79–90.  
36 526 <https://doi.org/10.1016/j.ijfoodmicro.2007.02.030>  
37  
38  
39  
40 527 De Taeye C, Eyamo Evina VJ, Cullet G, Niemenak N, Collin S (2016) Fate of Anthocyanins through Cocoa  
41 528 Fermentation. Emergence of New Polyphenolic Dimers. J Agric Food Chem 64: 8876-8885  
42 529 <https://pubs.acs.org/doi/10.1021/acs.jafc.6b03892>  
43  
44  
45  
46 530 De Vuyst L, Weckx S (2016) The cocoa bean fermentation process: from ecosystem analysis to starter culture  
47 531 development. J Appl Microbiol 121:5–17. <https://doi.org/10.1111/jam.13045>  
48  
49  
50 532 Figueroa-Hernández C, Mota-Gutierrez J, Ferrocino I et al (2019) The challenges and perspectives of the selection  
51 533 of starter cultures for fermented cocoa beans. Int J Food Microbiol 301:41–50.  
52 534 <https://doi.org/10.1016/j.ijfoodmicro.2019.05.002>  
53  
54  
55 535 Garcia-Armisen T, Papalexandratou Z, Hendryckx H et al (2010) Diversity of the total bacterial community  
56 536 associated with Ghanaian and Brazilian cocoa bean fermentation samples as revealed by a 16 S rRNA  
57 537 gene clone library. Appl Microbiol Biotechnol 87:2281–2292. [https://doi.org/10.1007/s00253-010-2698-](https://doi.org/10.1007/s00253-010-2698-9)  
58  
59 538 [9](https://doi.org/10.1007/s00253-010-2698-9)  
60  
61  
62  
63  
64  
65

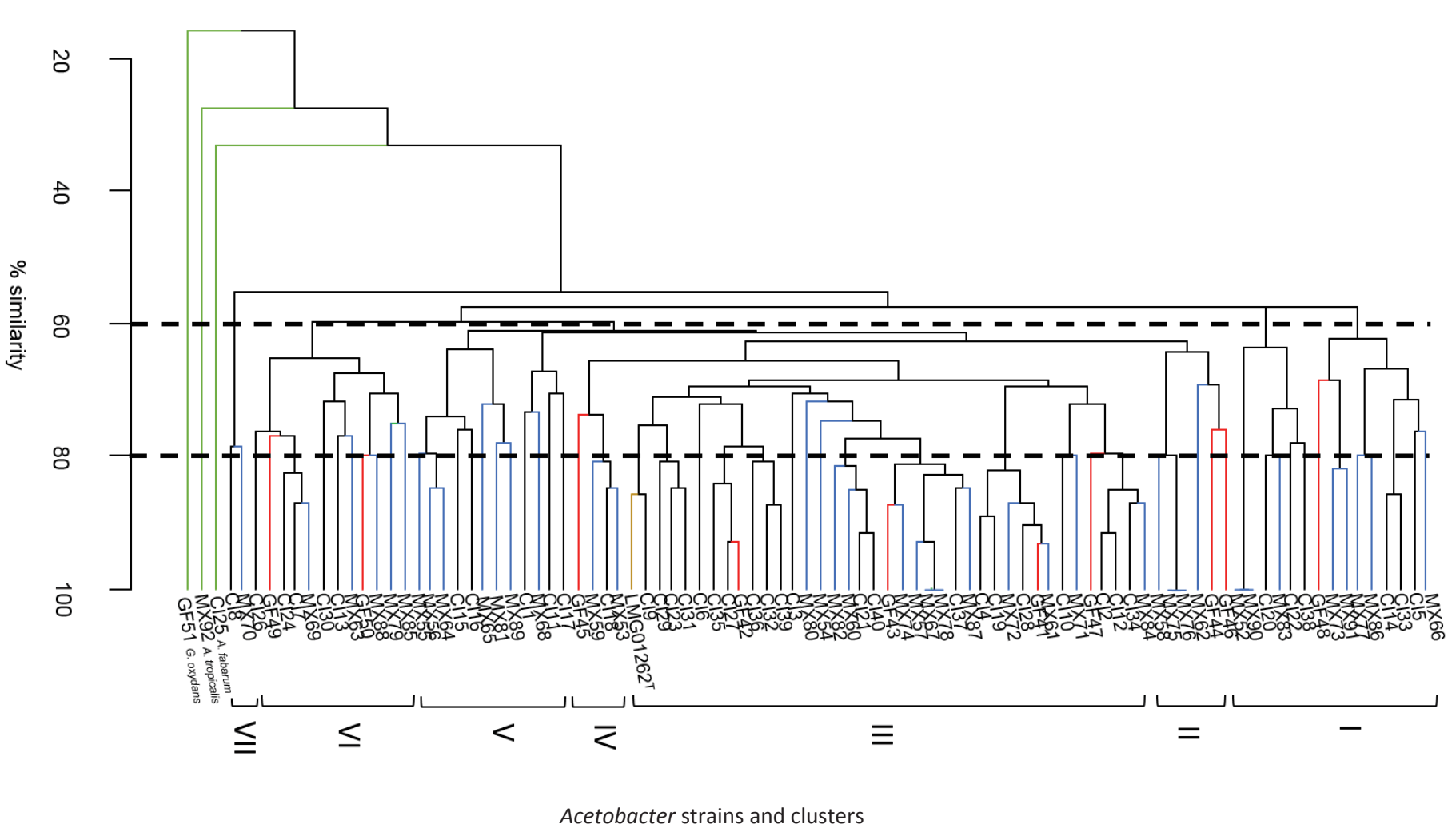


- 539 Gonzalez A, Hierro N, Poblet M et al (2004) Application of molecular methods for the differentiation of acetic  
1 540 acid bacteria in a red wine fermentation. J Appl Microbiol 96:853–860. <https://doi.org/10.1111/j.1365->  
2 [2672.2004.02220.x](https://doi.org/10.1111/j.1365-2672.2004.02220.x)  
3 541  
4  
5 542 Hamdouche Y, Guehi T, Durand N et al (2015) Dynamics of microbial ecology during cocoa fermentation and  
6 drying: Towards the identification of molecular markers. Food Control 48:117–122.  
7 543 <https://doi.org/10.1016/j.foodcont.2014.05.031>  
8 544  
9  
10  
11 545 Hansen CE, Olmo M del, Burri C (1998) Enzyme activities in cocoa beans during fermentation. J Sci Food Agric  
12 546 77:273–281. [https://doi.org/10.1002/\(SICI\)1097-0010\(199806\)77:2<273::AID-JSFA40>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1097-0010(199806)77:2<273::AID-JSFA40>3.0.CO;2-M)  
13  
14  
15 547 Holm CS, Aston JW, Douglas K (1993) The effects of the organic acids in cocoa on the flavour of chocolate. J Sci  
16 548 Food Agric 61:65–71. <https://doi.org/10.1002/jsfa.2740610111>  
17  
18  
19 549 Illegghems K, De Vuyst L, Weckx S (2013) Complete genome sequence and comparative analysis of *Acetobacter*  
20 550 *pasteurianus* 386B, a strain well-adapted to the cocoa bean fermentation ecosystem. BMC genomics  
21 551 14:526. <https://doi.org/10.1186/1471-2164-14-526>  
22  
23  
24  
25 552 John W Kumari N, Böttcher N, Koffi K. J et al (2016) Aseptic artificial fermentation of cocoa beans can be  
26 553 fashioned to replicate the peptide profile of commercial cocoa bean fermentations. Food res Int 89:764-  
27 554 772 <https://doi.org/10.1016/j.foodres.2016.10.011>  
28  
29  
30  
31 555 Lefeber T, Janssens M, Camu N, De Vuyst L (2010) Kinetic analysis of strains of lactic acid bacteria and acetic  
32 556 acid bacteria in cocoa pulp simulation media toward development of a starter culture for cocoa bean  
33 557 fermentation. Appl Environ Microbiol 76:7708–7716. <https://doi.org/10.1128/AEM.01206-10>  
34  
35  
36 558 Lefeber T, Janssens M, Moens F et al (2011) Interesting starter culture strains for controlled cocoa bean  
37 559 fermentation revealed by simulated cocoa pulp fermentations of cocoa-specific lactic acid bacteria. Appl  
38 560 Environ Microbiol 77:6694–6698. <https://doi.org/10.1128/AEM.00594-11>  
39  
40  
41  
42 561 Lefeber T, Papalexandratou Z, Gobert W et al (2012) On-farm implementation of a starter culture for improved  
43 562 cocoa bean fermentation and its influence on the flavour of chocolates produced thereof. Food Microbiol  
44 563 30:379–392. <https://doi.org/10.1016/j.fm.2011.12.021>  
45  
46  
47 564 Li L, Wieme A, Spitaels F et al (2014) *Acetobacter sicerae* sp. nov., isolated from cider and kefir, and identification  
48 565 of species of the genus *Acetobacter* by dnaK, groEL and rpoB sequence analysis. Int J Syst Evol  
49 566 Microbiol 64:2407–2415. <https://doi.org/10.1099/ijs.0.058354-0>  
50  
51  
52  
53 567 Magalhães da Veiga Moreira I, De Figueiredo Vilela L, Da Cruz Pedroso M et al (2017) Impact of a microbial  
54 568 cocktail used as a starter culture on cocoa fermentation and chocolate flavor. Molecules 22:766.  
55 <https://doi.org/10.3390/molecules22050766>  
56 569  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 570 Matsutani M, Nishikura M, Saichana N et al (2013) Adaptive mutation of *Acetobacter pasteurianus* SKU1108  
1 571 enhances acetic acid fermentation ability at high temperature. J Biotechnol 165:109–119.  
2  
3 572 <https://doi.org/10.1016/j.jbiotec.2013.03.006>  
4
- 5 573 Meersman E, Steensels J, Mathawan M et al (2013) Detailed analysis of the microbial population in Malaysian  
6 574 spontaneous cocoa pulp fermentations reveals a core and variable microbiota. PLoS ONE 8:e81559.  
7  
8 575 <https://doi.org/10.1371/journal.pone.0081559>  
9
- 10  
11 576 Moens F, Lefeber T, De Vuyst L (2014) Oxidation of metabolites highlights the microbial interactions and role of  
12 577 *Acetobacter pasteurianus* during cocoa bean fermentation. Appl Environ Microbiol 80:1848–1857.  
13  
14 578 <https://doi.org/10.1128/AEM.03344-13>  
15
- 16 579 Nielsen DS, Teniola OD, Ban-Koffi L et al (2007) The microbiology of Ghanaian cocoa fermentations analysed  
17 580 using culture-dependent and culture-independent methods. Int J Food Microbiol 114:168–186.  
18  
19  
20  
21 581 OIV- Organisation internationale de la vigne et du vin. 2010. Recueil des méthodes internationales d'analyses,  
22 582 Analyse microbiologique des vins et des moûts : détection, différenciation et dénombrement des  
23 583 microorganismes (OIV-Oeno 206-2010). 2nd ed. OIV.
- 24  
25  
26 584 Ozturk G, Young GM (2017) Food Evolution: The Impact of society and science on the fermentation of cocoa  
27 585 beans. J Compr Rev Food Sci Food Saf 16:431–455. <https://doi.org/10.1111/1541-4337.12264>  
28  
29
- 30 586 Papalexandratou Z, Camu N, Falony G, De Vuyst L (2011a) Comparison of the bacterial species diversity of  
31 587 spontaneous cocoa bean fermentations carried out at selected farms in Ivory Coast and Brazil. Food  
32 588 Microbiol 28:964–973. <https://doi.org/10.1016/j.fm.2011.01.010>  
33  
34  
35
- 36 589 Papalexandratou Z, Vrancken G, De Bruyne K et al (2011b) Spontaneous organic cocoa bean box fermentations  
37 590 in Brazil are characterized by a restricted species diversity of lactic acid bacteria and acetic acid bacteria.  
38 591 Food Microbiol 28:1326–1338. <https://doi.org/10.1016/j.fm.2011.06.003>  
39  
40
- 41 592 Papalexandratou Z, Falony G, Romanens E et al (2011c) Species diversity, community dynamics, and metabolite  
42 593 kinetics of the microbiota associated with traditional Ecuadorian spontaneous cocoa bean fermentations.  
43 594 Appl Environ Microbiol 77:7698–7714. <https://doi.org/10.1128/AEM.05523-11>  
44  
45  
46
- 47 595 Papalexandratou Z, Lefeber T, Bahrim B et al (2013) *Hanseniaspora opuntiae*, *Saccharomyces cerevisiae*,  
48 596 *Lactobacillus fermentum*, and *Acetobacter pasteurianus* predominate during well-performed Malaysian  
49 597 cocoa bean box fermentations, underlining the importance of these microbial species for a successful  
50 598 cocoa bean fermentation process. Food Microbiol 35:73–85. <https://doi.org/10.1016/j.fm.2013.02.015>  
51  
52  
53
- 54 599 Pereira GV de M, Miguel MG da CP, Ramos CL, Schwan RF (2012) Microbiological and physicochemical  
55 600 characterization of small-scale cocoa fermentations and screening of yeast and bacterial strains to develop  
56 601 a defined starter culture. Appl Environ Microbiol 78:5395–5405. <https://doi.org/10.1128/AEM.01144-12>  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 602 Pereira GV, Magalhães KT, De Almeida EG et al (2013a) Spontaneous cocoa bean fermentation carried out in a  
1 603 novel-design stainless steel tank: Influence on the dynamics of microbial populations and physical–  
2 604 chemical properties. Int J Food Microbiol 161:121–133.  
3 605 <https://doi.org/10.1016/j.ijfoodmicro.2012.11.018>  
4  
5  
6  
7 606 Pereira GV, Magalhães KT, Schwan RF (2013b) rDNA-based DGGE analysis and electron microscopic  
8 607 observation of cocoa beans to monitor microbial diversity and distribution during the fermentation  
9  
10 608 process. Food Res Int 53:482–486. <https://doi.org/10.1016/j.foodres.2013.05.030>  
11  
12  
13 609 Pettipher GL (1986) Analysis of cocoa pulp and the formulation of a standardised artificial cocoa pulp medium. J  
14 610 Sci Food Agric 37:297–309. <https://doi.org/10.1002/jsfa.2740370315>  
15  
16  
17 611 Saeki A, Matsushita K, Takeno S et al (1999) Enzymes responsible for acetate oxidation by acetic acid bacteria.  
18 612 Biosci Biotechnol Biochem 63:2102–2109. <https://doi.org/10.1271/bbb.63.2102>  
19  
20  
21 613 Sakurai K, Arai H, Ishii M, Igarashi Y (2011) Transcriptome response to different carbon sources in *Acetobacter*  
22 614 *acetii*. Microbiol 157:899–910. <https://doi.org/10.1099/mic.0.045906-0>  
23  
24  
25 615 Sato J, Wakayama M, Takagi K (2015) Lactate dehydrogenase involved in lactic acid metabolism of *Acetobacter*  
26 616 *Pasteurianus*. Proc Environ Sci 28:67–71. <https://doi.org/10.1016/j.proenv.2015.07.010>  
27  
28  
29 617 Schwan RF, Wheals AE (2004) The Microbiology of cocoa fermentation and its role in chocolate quality. Crit Rev  
30 618 Food Sci Nutr 44:205–221. <https://doi.org/10.1080/10408690490464104>  
31  
32  
33 619 Sengun IY, Karabiyikli S (2011) Importance of acetic acid bacteria in food industry. Food Control 22:647–656.  
34 620 <https://doi.org/10.1016/j.foodcont.2010.11.008>  
35  
36  
37 621 Steensels J, Meersman E, Snoek T, et al (2014) Large-Scale Selection and Breeding To Generate Industrial Yeasts  
38 622 with Superior Aroma Production. Appl Environ Microbiol 22:6965–6975.  
39 623 <http://dx.doi.org/10.1128/AEM.02235-14>.  
40  
41  
42  
43 624 Teyssier C, Hamdouche Y (2016) Acetic acid bacteria. Fermented Foods, Part I: Biochemistry and Biotechnology  
44 625 112-126. CRC Press. ISBN 978-1-4987-4081-4  
45  
46  
47 626 Teyssier C, Marchandin H, Simeon De Buochberg M et al (2003) Atypical 16S rRNA gene copies in  
48 627 *Ochrobactrum intermedium* strains reveal a large genomic rearrangement by recombination between rrn  
49 628 Copies. J Bacteriol 185:2901–2909. <https://doi.org/10.1128/JB.185.9.2901-2909.2003>  
50  
51  
52  
53 629 Tulini FL, Winkelströter LK, De Martinis ECP (2013) Identification and evaluation of the probiotic potential of  
54 630 *Lactobacillus paraplantarum* FT259, a bacteriocinogenic strain isolated from Brazilian semi-hard  
55 631 artisanal cheese. Anaerobe 22:57–63. <https://doi.org/10.1016/j.anaerobe.2013.06.006>  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

- 632 Urbańska B, Derewiaka D, Lenart A, Kowalska J (2019) Changes in the composition and content of polyphenols  
1 633 in chocolate resulting from pre-treatment method of cocoa beans and technological process Technol Eur  
2  
3 634 Food Res 245: 2101. <https://doi.org/10.1007/s00217-019-03333-w>  
4
- 5 635 Valera MJ, Torija MJ, Mas A, Mateo E (2013) *Acetobacter malorum* and *Acetobacter cerevisiae* identification  
6  
7 636 and quantification by Real-Time PCR with TaqMan-MGB probes. Food Microbiol 36:30–39.  
8  
9 637 <https://doi.org/10.1016/j.fm.2013.03.008>  
10
- 11 638 Visintin S, Alessandria V, Valente A et al (2016) Molecular identification and physiological characterization of  
12  
13 639 yeasts, lactic acid bacteria and acetic acid bacteria isolated from heap and box cocoa bean fermentations  
14  
15 640 in West Africa. Int J Food Microbiol 216:69–78. <https://doi.org/10.1016/j.ijfoodmicro.2015.09.004>  
16
- 17 641 Wang B, Shao Y, Chen T et al (2015) Global insights into acetic acid resistance mechanisms and genetic stability  
18  
19 642 of *Acetobacter pasteurianus* strains by comparative genomics. Sci Rep 5:18330  
20  
21 643 <https://doi.org/10.1038/srep18330>  
22
- 23 644 Wu J, Gullo M, Chen F, Giudici P (2010) Diversity of *Acetobacter pasteurianus* strains isolated from solid-state  
24  
25 645 fermentation of cereal vinegars. Curr Microbiol 60:280–286. <https://doi.org/10.1007/s00284-009-9538-0>  
26
- 27 646 Yang H, Yu Y, Fu C, Chen F (2019) Bacterial acid resistance toward organic weak acid revealed by RNA-Seq  
28  
29 647 transcriptomic analysis in *Acetobacter pasteurianus*. Front Microbiol 10:1616  
30  
31 648 <https://doi.org/10.3389/fmicb.2019.01616>  
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



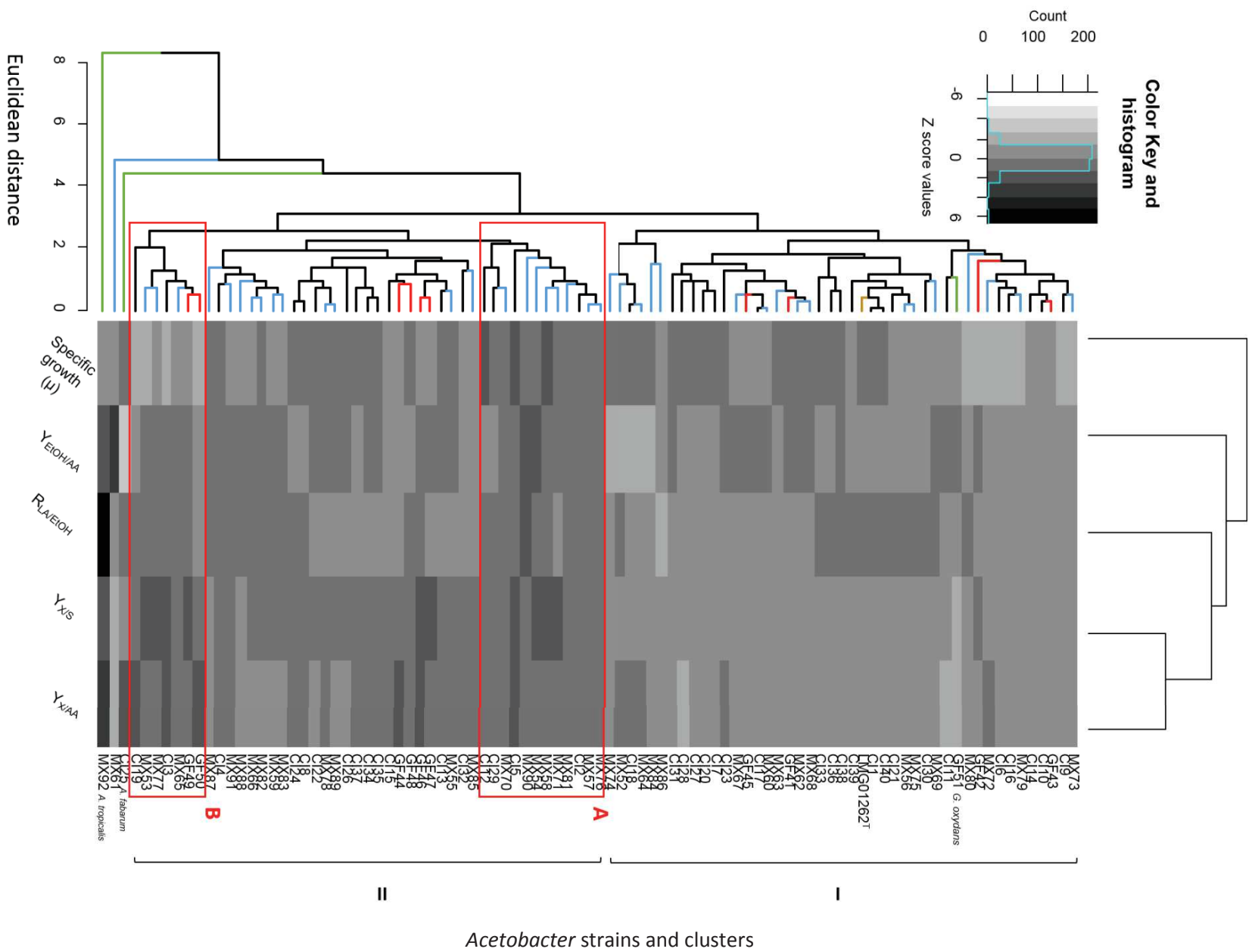
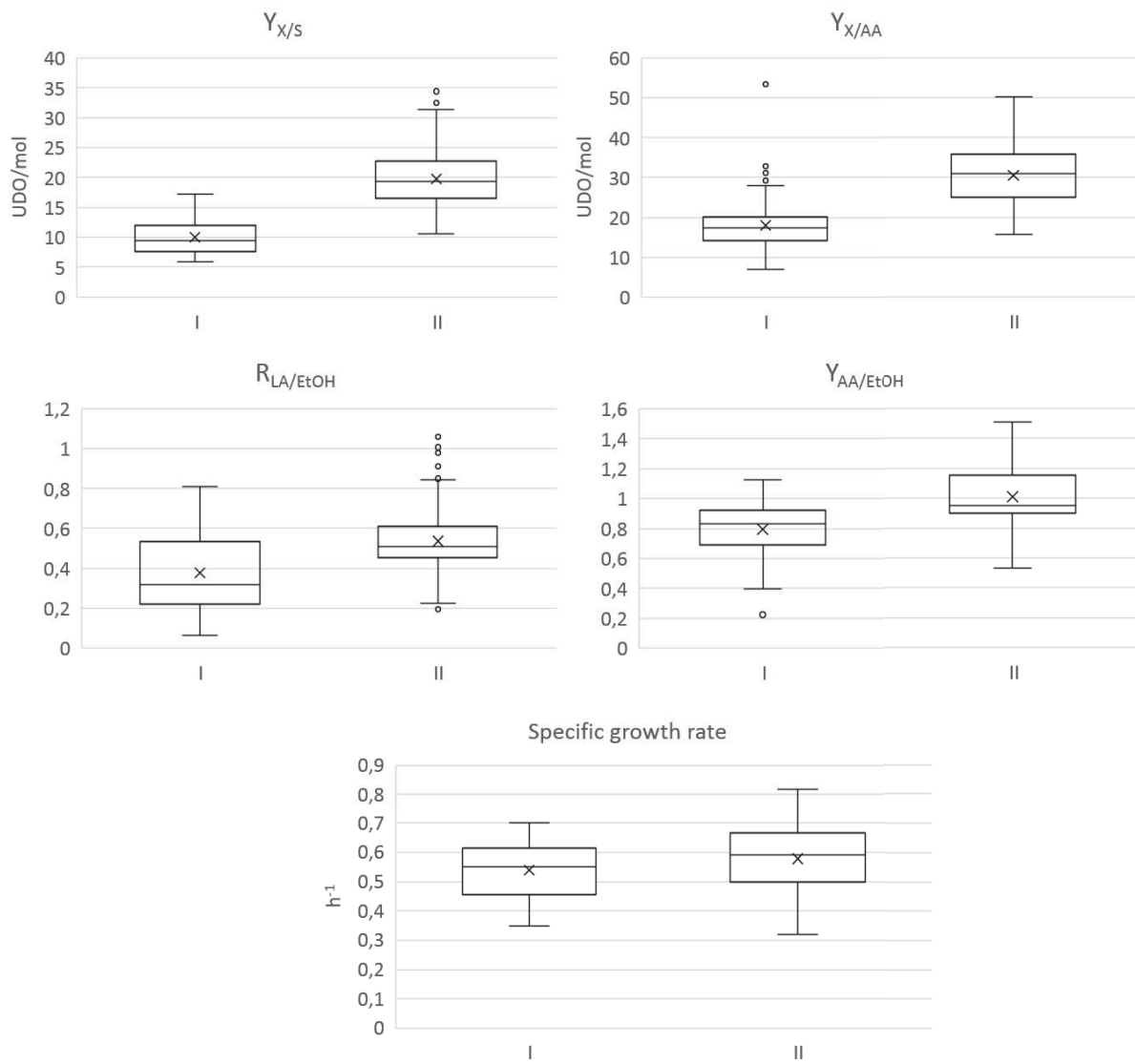


Fig. 2



**Fig. 3**

Table 1: Data concerning cocoa beans sampling

<b>Origin</b>	<b>Sampling date</b>	<b>Geographical area</b>	<b>Fermentation type</b>	<b>Duration of fermentation</b>	<b>Number of fermentation box/heap</b>
Ivory Coast	2016	Wagana	Box	6	6
Ivory Coast	2018	Abidjan	Heap	7	1
French Guiana	2017	Combi	Box	7	1
Mexico	02 / 2018	Comalcalco	Box	5	2
Mexico	10 / 2018	Comalcalco	Box	4	2



Table 2: Molecular identification of the 130 AAB strains by 16S rRNA gene sequencing and repartition of the isolates according to the three countries, Ivory Coast, French Guiana and Mexico.

Bacterial species identification	Total strains	Ivory Coast		French Guiana	Mexico	
		2016	2018		Main	Mid crop
<i>A. pasteurianus</i>	<b>89</b>	24	15	10	11	29
<i>A. tropicalis</i>	<b>16</b>	5	2	2	5	2
<i>A. fabarum</i>	<b>2</b>	1	1			
<i>A. ghanensis</i>	<b>7</b>	5	2			
<i>A. orientalis</i>	<b>2</b>			2		
<i>A. malorum / A. cerevisiae</i>	<b>6</b>					6
<i>G. oxydans</i>	<b>7</b>	1		6		
<i>G. nephelii</i>	<b>1</b>	1				
<b>Sub-total</b>		<b>37</b>	<b>20</b>		<b>16</b>	<b>37</b>
<b>Total strains</b>	<b>130</b>		<b>57</b>	<b>20</b>		<b>53</b>

Table 3: Clustering of *A. pasteurianus* strains according to genomic polymorphism analysis by (GTG)<sub>5</sub> PCR

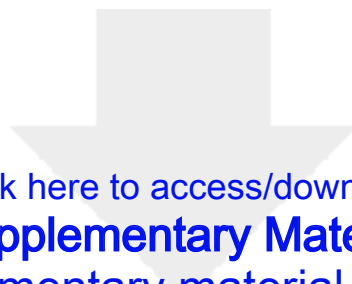
(GTG) <sub>5</sub> grouping	Isolates number	% clustering	Isolates	Origin country
I	15	58%	MX66; CI5; CI33; CI14; MX86; MX77; MX91; MX73; GF48; CI38; CI22; MX83; CI20; MX90; MX52	Mexico (20%*), Ivory Coast (15%*) & French Guiana (10%*)
II	6	62.5%	GF46; GF44; MX62; MX76; MX75; MX58	Mexico (10%*) & French Guiana (20%*)
III	39	62.5%	MX84; CI34; CI12; CI2; GF47; MX71; CI10; MX61; GF41; CI28; MX72; CI19; CI4; MX87; CI37; MX78; MX67; MX57; MX74; GF43; CI40; CI21; MX60; MX82; MX54; MX80; CI3; CI39; CI32; CI36; GF42; CI27; CI35; CI6; CI31; CI23; CI29; CI9; LMG01262 <sup>†</sup>	Ivory Coast (54%*), Mexico (32.5%*), French Guiana (40%*)
IV	4	62.5%	MX53; CI18; MX59, GF45	Mexico (5%*), Ivory Coast (2.5%*) & French Guiana (10%*)
V	12	62.5%	CI17; CI11; MX68; CI1; MX89; MX81; MX65; CI16; CI15; MX64; MX56; MX55	Mexico (17.5%*), Ivory Coast (13%*)
VI	12	60%	MX85; MX79; MX88; GF50; MX63; CI13; CI30; MX69; CI7; CI24; GF49; CI26	Mexico (12.5%*), Ivory Coast (13%*) & French Guiana (20%*)
VII	2	55%	MX70; CI8	Mexico (2.5%*) & Ivory Coast (2.5%*)

\* Percentage is the number of strains of the country in the cluster divided by total number of strains from this country.

Table 4: Clustering of the *A. pasteurianus* strains according to physiological traits

Metabolism group	Isolates number	Isolates	Origin country
<b>I</b>	45	LMG01262 <sup>T</sup> ; CI1; CI10; CI11; CI14; CI16; CI17; CI18; CI20; CI21; CI23; CI27; CI28; CI30; CI31; CI33; CI36; CI38; CI39; CI40; CI6; CI7; CI9; GF41; GF42; GF43; GF45; MX52; MX56; MX60; MX62; MX63; MX64; MX67; MX68; MX69; MX72; MX73; MX74; MX75; MX79; MX80; MX84; MX86; GF51	Ivory Coast (22/39, 56.5%*) & Mexico (17/40, 42.5%*), and French Guiana (4/10, 40%*)
<b>II</b>	45	CI12; CI13; CI15; CI19; CI2; CI22; CI24; CI26; CI29; CI3; CI32; CI34; CI35; CI37; CI4; CI5; CI8; GF44; GF46; GF47; GF48; GF49; GF50; MX53; MX54; MX55; MX57; MX58; MX59; MX65; MX66; MX70; MX71; MX76; MX77; MX78; MX81; MX82; MX83; MX85; MX87; MX88; MX89; MX90; MX91	Mexico (22/40, 55%*), Ivory Coast (18/39, 43.5%*), French Guiana (6/10, 60%*)

\* Percentage is the number of strains of the country in the cluster divided by total number of strains from this country.



[Click here to access/download](#)

**Supplementary Material**

Farrera-Supplementary material - reviewed.docx



## Résumé

La fermentation du cacao est une fermentation spontanée qui dure de 4 à 8 jours et repose principalement sur la succession de trois groupes de microorganismes : les levures, les bactéries lactiques et les bactéries acétiques qui réalisent les fermentations alcoolique, lactique et acétique respectivement. Les fèves saines sont stériles jusqu'à l'ouverture de la cabosse. L'inoculation des fèves est réalisée le plus souvent naturellement à l'aide de l'environnement autour de l'écabossage et de la mise en fermentation. Les procédés de traitements post-récolte sont différents d'un pays à l'autre et influencent le déroulement de la fermentation. Cependant, trois espèces de bactéries lactiques et acétiques (*Lactobacillus plantarum*, *Lactobacillus fermentum* et *Acetobacter pasteurianus*) dominent les fermentations dans tous les pays. En revanche, leur diversité intra spécifique n'a été que rarement étudiée. Dans cette étude nous avons utilisé des méthodes moléculaires indépendantes de la culture (PCR-DGGE et metabarcoding) pour étudier les communautés bactériennes associées à la fermentation des fèves de cacao selon trois pays : le Mexique, la Côte d'Ivoire et la Guyane. La méthode de metabarcoding a également été utilisée pour identifier la contribution des surfaces liées à l'environnement pré et post récolte des cabosses de cacao lors d'une fermentation réalisée au Mexique. La dominance des genres *Lactobacillus* et *Acetobacter* au cours de la fermentation dans chaque pays a été confirmée. De plus, la présence de genres spécifiques à chaque pays a été mise en évidence lors du premier jour de la fermentation. L'ensemble des surfaces lié à l'environnement de la fermentation semble participer à l'inoculation des genres dominants. Elles agissent en tant que réservoirs bactériens. Une collection de souches de bactéries lactiques et acétiques a été constituée. *L. plantarum* et *A. pasteurianus* ont été les deux espèces isolées en majorité. La diversité intra spécifique des souches d'*A. pasteurianus* a été étudiée. Pour cela, leurs polymorphismes génomiques ont été analysés à l'aide d'une amplification par PCR sur des séquences répétées et leurs caractéristiques biochimiques ont été comparées dans un milieu simulant les conditions de la pulpe de cacao au 2<sup>ème</sup> jour de fermentation. Notre étude a contribué à montrer que les mêmes souches d'*A. pasteurianus* peuvent être présentes dans les 3 pays différents. Certaines souches se distinguent pour leur plus grande affinité pour l'acide lactique que les autres, ce qui est intéressant pour améliorer la qualité organoleptique du cacao final. Les résultats sur la diversité intra spécifique nous permettent de proposer des candidats potentiels pour l'élaboration de starters de culture pour la fermentation des fèves de cacao.

**Mots clés :** fermentation du cacao, bactéries acétiques, bactéries lactiques, écologie, metabarcoding, PCR-DGGE, acide lactique, acide acétique

## Abstract

Cocoa fermentation is a spontaneous fermentation that lasts 4 to 8 days. It is mainly based on the succession of three groups of microorganisms: yeasts, lactic acid bacteria and acetic bacteria that carry out respectively the alcoholic, lactic and acetic fermentation. The beans are sterile until the opening of the pod. The inoculation of the beans is usually naturally done using the environment around the pod opening and the fermentation process. Post-harvest treatment processes differ from one country to another and influence the fermentation progress. However, three species of lactic and acetic bacteria (*L. plantarum*, *L. fermentum* and *A. pasteurianus*) dominate the fermentations in all countries. On the other hand, their intraspecific diversity was rarely studied. In this study, we used the metabarcoding method to study the interspecific diversity of bacterial communities associated with the fermentation of cocoa beans in 3 countries: Mexico, Ivory Coast and Guyana. In addition, this method was used to identify the contribution of the surfaces related to the pre- and post-harvest environment of cocoa pods during the fermentation, which was carried out in Mexico. The dominance of the genera *Lactobacillus* and *Acetobacter* during fermentation in each country has been confirmed. In addition, the presence of country-specific genera was founded on the first day of fermentation. All the surfaces linked to the fermentation environment participate to the inoculation of the dominant genera. They act as bacterial tanks. A collection of lactic and acetic bacteria strains was produced. *L. plantarum* and *A. pasteurianus* were the most isolated species. Intra-specific diversity of *A. pasteurianus* strain was studied. For this, their genomic polymorphisms were analyzed using PCR amplification on repeated sequences and their biochemical characteristics were compared in a specific medium, simulating the conditions of the cocoa pulp at the 2nd day of fermentation. Our study showed that the strains of *A. pasteurianus* could be present in the three different countries. Some strains were distinguished for their greater affinity for lactic acid than the others, which is interesting in order to improve the organoleptic quality of the final cocoa. The results on intra-specific diversity allow us to propose potential candidates to produce culture starters for the fermentation of cocoa beans.

**Key words:** cocoa fermentation, acetic acid bacteria, lactic acid bacteria, ecology, metabarcoding, PCR-DGGE, lactic acid, acetic acid