



HAL
open science

Biologie des populations du complexe d'espèces *Ralstonia solanacearum* à Madagascar et dans le sud-ouest de l'océan indien

Hasina Ny Aina Rasoamanana

► **To cite this version:**

Hasina Ny Aina Rasoamanana. Biologie des populations du complexe d'espèces *Ralstonia solanacearum* à Madagascar et dans le sud-ouest de l'océan indien. Biologie végétale. Université de la Réunion, 2022. Français. NNT : 2022LARE0017 . tel-03773531

HAL Id: tel-03773531

<https://theses.hal.science/tel-03773531>

Submitted on 9 Sep 2022

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

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

"Cette thèse a reçu le soutien financier de la Région Réunion et de l'Union Européenne (Fonds Européen de Développement Régional – FEDER)"



UNIVERSITÉ DE LA RÉUNION

UMR Peuplements Végétaux et Bio-agresseurs en Milieu Tropical

Centre National de Recherche Appliquée au Développement Rural *FOFIFA*

École doctorale Sciences Technologies Santé (STS, ED 542)

THÈSE

Pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ DE LA REUNION

Discipline du doctorat : biologie des populations et écologie

"Biologie des populations du complexe d'espèces *Ralstonia solanacearum* à Madagascar et dans le Sud-ouest de l'océan indien"

par Rasoamanana Hasina Ny Aina

Date de soutenance : 08 juillet 2022

Pascale BESSE, Professeure, HDR, Université de La Réunion
Séverine Zirah, MCF, HDR, MNHN Paris
Christophe Lemaire, MCF, HDR, Université d'Angers
Santatra Ravelomanantsoa, Chercheuse, FOFIFA
Stéphane Poussier, Professeur, HDR, Université de La Réunion

Présidente du jury
Rapporteur
Rapporteur
Invitée
Directeur de thèse



« Insensé est celui qui ne fait pas mieux que ses parents »

Proverbe malgache

merci

Mes sincères remerciements vont vers :

Le Directeur de l'UMR PVBMT, Pr. Bernard REYNAUD, pour m'avoir accueilli au sein de son laboratoire et permis de réaliser ma thèse dans les meilleures conditions de travail.

Mon Directeur de thèse, Pr. Stéphane Poussier, et mes encadreur, Dr. Santatra Ravelomanantsoa, Dr. Isabelle Robène, Dr. Fabien Guerin, Dr. Yann Pécricx, pour m'avoir donné l'opportunité de réaliser cette thèse. Merci de m'avoir formé, dirigé, et soutenu. Je n'aurai pu rêver un meilleur encadrement, j'ai toujours ressenti du respect et de la bienveillance au cours de notre collaboration et je vous en suis très reconnaissante.

Les membres de mon comité de thèse, Dr. Alice Guidot et Dr. Christian Vernière, pour leurs conseils avisés et leurs encouragements.

Les personnels du Pôle de Protection des Plantes, 3P, pour leurs sympathies et collaborations tout au long de mon séjour au 3P.

Ma famille pour avoir partagé les joies et les peines de ma thèse. Mille mercis à mes parents qui m'ont toujours poussé à faire de mon mieux et à mes deux frères pour leurs soutiens indéfectibles.

À tous ceux qui ont participés de près ou de loin à la réalisation de ce travail.

Table des matières

<i>Introduction générale</i>	1
Synthèse bibliographique.....	3
I. Epidémiologie moléculaire.....	3
1. Diversité et structure génétique des populations bactériennes	4
2. Intérêt de la méthode de génotypage MultiLocus Variable number of tandem repeats (VNTR) Analysis (MLVA) en épidémiologie moléculaire.....	8
3. Facteurs impliqués dans l'émergence de lignée génétique phytopathogène	11
II. Bactériocines	14
1. Généralités sur les bactériocines	14
2. Classification des bactériocines.....	15
2.1. Bactériocines produites par les bactéries de type gram positif	15
2.2. Bactériocines produites par les bactéries de type gram négatif	17
3. Rôle biologique des bactériocines	20
4. Application des bactériocines	21
4.1. Dans le domaine de l'industrie agroalimentaire.....	21
4.2. Dans le domaine de la santé humaine et animale.....	23
4.3. Dans le domaine de la phytopathologie.....	24
III. Modèle d'étude : le complexe d'espèces <i>Ralstonia solanacearum</i>	26
1. Distribution géographique, spectre d'hôtes et importance socioéconomique	26
2. Caractéristiques morphologiques et génomiques	26
3. Taxonomie et diversité phylogénétique	27
4. Cycle infectieux et symptomatologie.....	28
5. Stratégies de lutte et de gestion.....	31
6. Distribution de la diversité génétique du ceRs	32
6.1. À l'échelle mondiale	32
6.3. À Madagascar.....	33
7. Etudes antérieures sur l'activité bactériocine du ceRs	35
IV. Objectifs et questions de recherche	36
<i>Chapitre 1</i>	37
Article 1	39
Contrasting genetic diversity and structure among Malagasy <i>Ralstonia pseudosolanacearum</i> phylotype I populations inferred from an optimized Multilocus Variable Number of Tandem Repeat Analysis scheme	39
Introduction	40
Materials and Methods	42
1. Bacterial isolates and phylogenetic assignment.....	42

2. VNTR identification and selection for MLVA	44
3. Primer design and assay optimization	45
4. MLVA genotyping	46
5. Data analysis.....	46
Results	48
1. Development of the RS1-MLVA14 scheme	48
2. Genotypic resolution of the RS1-MLVA14 scheme	53
3. RS1-MLVA14 revealed genetic diversity among Malagasy phylotype I populations and was discriminative at the field scale	53
4. RS1-MLVA14 revealed genetic structure among Malagasy phylotype I populations and showed the singularity of the most prevalent genetic cluster	56
Discussion	59
1. RS1-MLVA14 scheme, a genotyping tool adapted to study the population genetics of phylotype I, both at field and global scales.....	60
2. RS1-MLVA14 unveiled contrasting genetic diversity and epidemiology among Malagasy phylotype I populations.....	61
3. Comparative epidemiology of Malagasy phylotype I, IIB-1 and III populations	63
Article 2	77
The epidemiology of the Malagasy <i>Ralstonia pseudosolanacearum</i> phylotype I genetic lineages might be impacted by bacteriocin production.....	77
Introduction	78
Material and methods	80
1. Bacterial collections and phylogenetic assignment	80
2. MLVA typing	83
3. Genetic structure and diversity analysis.....	84
4. Inhibition activity assays.....	85
5. Characterization of inhibitory activity in culture supernatants	85
6. Data analysis.....	86
Results	86
1. Phylogenetic assignments of the Malagasy isolates	86
2. Genetic diversity of the Malagasy phylotype I isolates	87
3. Population structure analysis of the Malagasy phylotype I isolates	91
4. Genetic links between the Malagasy and worldwide phylotype I isolates	93
5. Growth inhibition activity.....	94
Discussion	96
1. High prevalence of the phylotype I sequevar 18 in Madagascar	96
2. Broad spectrum inhibition activity of the sequevars 18 and 31.....	98
Chapitre 2.....	105

Introduction	106
Matériels et méthodes.....	108
1. Souches bactériennes étudiées.....	108
2. Protocole d'expérimentation de typage bactériocine	113
2.1. Production de bactériocines.....	113
2.2. Préparation des tapis bactériens et dépôts des spots de bactériocine.....	113
3. Contrôle de l'activité des bactériocines.....	116
4. Analyse des données de typage bactériocine	117
5. Protocole d'extraction d'ADN génomique à haut poids moléculaire	118
6. Séquençage, assemblage et annotation des génomes	119
7. Recherche <i>in silico</i> des clusters de gènes bactériocine	119
8. Analyse pangénomique des souches du ceRs	120
Résultats	121
1. Activités bactériocines des souches du ceRs.....	121
2. Qualité et quantité des ADN génomiques extraits	124
3. Qualité de l'assemblage des génomes.....	125
4. Caractérisation génomique des bactériocines chez les souches du ceRs.....	128
5. Vers la recherche d'autres gènes candidats pouvant être impliqués dans le succès épidémiologique du I-18 à Madagascar et du I-31 dans les petites îles du SOOI	130
Discussion	138
1. Mise au point d'un protocole d'extraction d'ADN génomique à haut poids moléculaire chez le ceRs.....	138
2. Qu'est-ce qui pourrait expliquer le succès épidémiologique des lignées génétiques I-18 à Madagascar et I-31 dans les petites îles du SOOI?.....	138
Conclusion générale et perspectives.....	141
1. Pertinence du schéma RS1-MLVA14 pour l'étude épidémiologique du phylotype I à Madagascar	142
2. Situation épidémiologique du ceRs à Madagascar et liens génétiques avec les souches internationales	143
3. Activité antibactérienne entre les souches du ceRs	145
4. Potentiel des bactériocines dans le biocontrôle du ceRs	147
Résumé.....	149
Abstract.....	150
Bibliographie.....	151
ANNEXES	166
<i>Annexes chapitre 1</i>	167
<i>Annexes chapitre 2</i>	196

Liste des figures et tableaux:

Introduction :

Figure 1: Facteurs influençant le développement d'une épidémie végétale.....	3
Figure 2: Mécanismes de transfert horizontal de gène chez la bactérie	5
Figure 3: Profil génétique de deux souches (A et B) après analyse MLVA.....	9
Figure 4: Modèle du Slipped Strand Mismatching (SSM).....	10
Figure 5: Principaux modèles d'évolution des VNTR. A) Modèle IAM et B) Modèle SSM.....	11
Figure 6: Micrographie électronique des tailocines. A) R2 pyocine; B) monocine de type F issue du <i>Listeria monocytogenes</i>	20
Figure 7: Répartition géographique du ceRs	26
Figure 8: A) Image au microscope électronique du ceRs ; B) Colonies virulentes du ceRs sur un milieu contenant du TTC ; C) Colonies avirulentes du ceRs sur un milieu contenant du TTC.	27
Figure 9: Classification du ceRs	28
Figure 10: Cycle infectieux du ceRs	29
Figure 11: Symptômes du flétrissement bactérien causé par le ceRs sur différentes plantes hôtes.	30
Figure 12: Distribution mondiale des quatre phylotypes du ceRs	33
Figure 13: Minimum Spanning Trees (MSTs)	34
Tableau 1: Indices de diversité et différenciation génétique des populations	7
Tableau 2: Classifications des bactériocines produites par les bactéries de type gram positif	15

Chapitre 1 :

Article 1:

Fig 1. Collection area of RSSC phylotype I in Madagascar	43
Fig 2. Distribution and frequency of phylotype I haplotypes collected in the 16 fields of the Atsinanana, Analamanga and Itasy regions.....	55
Fig 3. Multidimensional scaling (MDS) of the Malagasy phylotype I (total = 31 haplotypes, 291 isolates)	57
Fig 4: Minimum spanning tree (MST) of the Malagasy phylotype I according to the region of collect.....	57
Fig 5. Minimum spanning tree of the Malagasy and worldwide phylotype I.....	58

Table 1: Features of the 14 selected VNTR loci composing the RS1-MLVA14 scheme	51
--	----

Article 2 :

Fig_ 1: Collection area of RSSC phylotype I in Madagascar.....	82
Fig_ 2: Multidimensional scaling (MDS) of the 97 Malagasy haplotypes	91
Fig_ 3. Minimum spanning tree (MST) of the Malagasy phylotype I according to the province of collect.....	92
Fig_ 4. Minimum spanning tree (MST) of the Malagasy and worldwide phylotype I.....	93
Fig_ 5. Inhibitory activity of 18 producing isolates (horizontal) on 27 indicator isolates (vertical).	95

Table_ 1: Genetic diversity estimated from RS1-MLVA14 data of C1 and C2 RSSC collection	90
---	----

Chapitre 2 :

Figure_ 1: Répartition géographique des sequevars I-18 et I-31 à Madagascar, dans les petites îles du SOOI et en Afrique de l'Est	107
Figure_ 2: Schéma récapitulatif du protocole d'expérimentation de typage bactériocine	115
Figure_ 3: Identification des halos d'inhibition	117
Figure_ 4: Activité inhibitrice de 37 souches productrices de bactériocine (horizontal) sur 99 souches indicatrices (vertical)	124
Figure_ 5: Gel digital de migration des ADN génomiques extraits.....	125
Figure_ 6: Alignement des séquences génomiques nouvellement assemblées à la séquence de référence du ceRs GMI1000	128
Figure_ 7: Diversité des bactériocines chez le ceRs	129
Figure_ 8: Diagramme de Venn représentant la distribution des 8014 familles de gènes du pangéome entre les différentes lignées génétiques.....	132

Tableau_ 1: Information sur les souches bactériennes étudiées	109
Tableau_ 2: Qualité et quantité d' ADN génomiques extraits	124
Tableau_ 3: Résultats de l'analyse Busco présentant les pourcentages d'intégrité des génomes, les pourcentages de genes dupliqués, fragmentés et manquants parmi les 124 gènes recherchés	125
Tableau_ 4: Résultats de l'analyse Quast	126
Tableau_ 5: Comparaison deux à deux du nombre de SNPs entre les 20 génomes du ceRs.....	131
Tableau_ 6: Comparaison des effecteurs de type III chez les 20 génomes du ceRs.....	133

Liste des abréviations :

- **Organismes et institutions :**

3P	Pôle de Protection des Plantes
CIRAD	Centre de coopération Internationale en Recherche Agronomique pour le Développement
FIFAMANOR	Centre National de Développement Rural et de Recherche Appliquée
FOFIFA	Centre National de Recherche Appliquée au Développement Rural
UMR PVBMT	Unité Mixte de Recherche Peuplements Végétaux et Bioagresseurs en Milieu Tropical

- **Autres abréviations :**

A	Allelic Richness
ABC	ATP Binding Cassette
ADN	Acide Désoxyribonucléique
ARN	Acide ribonucléique
CC	Clonal Complex
ceRs	Complexe d'espèces <i>Ralstonia solanacearum</i>
CNV	Copy Number Variation
DLV	Double Locus Variant
EPS	Exopolysaccharides
FB	Flétrissement Bactérien
F _{ST} , R _{ST}	Indice de différenciation
G	Number of haplotypes
GSM	Generalized Stepwise Mutation model
HGDI	Indice de Discrimination Hunter-Gaston
H _{nb}	Indice de diversité génétique non biaisé de Nei
IAM	Infinite Allele Model
kb	Kilobase
kDa	<i>kilodalton</i>
LAMP	Loop-Mediated Isothermal Amplification
LAPs	Linear Azole containing Peptides
Mb	Megabase
MCC	Malagasy Clonal Complex
MLSA	MultiLocus Sequence Analysis
MLST	MultiLocus Sequence Typing (typage de séquences multilocus)
MLVA	Multiple Loci VNTR Analysis (analyse de plusieurs loci VNTR)
MST	Minimum Spanning Tree
MT	Multilocus-Type (MLVA)
n	Nombre de souches
Na	Nombre moyen d'allèle par locus
NPB	Non Pathogenic on Banana
OMS	Organisation Mondiale de la Santé
P	Taux de polymorphisme
PacBio	Pacific Bioscience

pb	Paire de Base
PCR	Amplification par polymérisation en chaîne
RBP	Receptor Binding Proteins
RFLP	Restriction Fragment Length Polymorphism
RSSC	<i>Ralstonia Solanacearum</i> Species Complex
SLV	Single Locus Variant (un locus variant)
SMM	Stepwise Mutation Model
SMRT	Single Molecule Real-Time
SNP	Single Nucleotide Polymorphism
SOOI	South West Indian Ocean
ST	Séquence-type (MLST)
TLV	Triple Locus Variant
TPM	Two Phase Model
TR	Tandem Repeat (répétition en tandem)
VNTR	Variable Number of Tandem Repeat (répétition en tandem polymorphe)

Introduction générale

La pratique de l'agriculture a beaucoup changé depuis son avènement il y a 2000 à 12000 ans avant notre ère [1–3]. Initialement, les premières cultures étaient probablement un mélange d'espèces issues de populations sauvages présentant une grande diversité génétique [1]. Ensuite, au cours du néolithique, les agriculteurs ont commencé à sélectionner les phénotypes les plus intéressants, à pratiquer la monoculture d'espèces et probablement à répandre leurs pratiques dans différentes régions du monde [1–3]. Depuis, l'augmentation de la densité des cultures agricoles et l'uniformisation accrue des espèces et variétés cultivées ont favorisé l'émergence de nouveaux phytopathogènes et modifié significativement les populations d'agents phytopathogènes déjà existantes [1]. En effet, ce type d'agroécosystème facilite la dispersion des agents phytopathogènes, augmente la taille de leurs populations, favorise une plus grande diversité génétique et plus de mutations disponibles pour la sélection. Par ailleurs, l'utilisation massive et irraisonnée de pesticides favorise la sélection des génotypes résistants pouvant se multiplier rapidement et se disperser dans les parcelles cultivées [1].

En parallèle, le développement des moyens de transport (terrestre, maritime, aérien) a favorisé le commerce national et international des produits agricoles et le déplacement des populations humaines. Ainsi, différentes espèces de plantes agricoles, parfois accompagnées de leurs agents pathogènes, ont été introduites dans différentes régions du monde, souvent bien loin de leurs aires d'origine [1]. Ces cas d'introductions d'agents phytopathogènes peuvent avoir de lourdes conséquences dans leur nouvel environnement, car les plantes-hôtes qui n'ont pas coévolué avec ceux-ci peuvent avoir un niveau de résistance plus faible. Citons en exemple, l'introduction au XIX^{ème} siècle du *Phytophthora infestans* en Irlande qui a engendré une grave épidémie de mildiou sur la pomme de terre et entraîné une grande famine [4]; l'introduction au début des années 1900 du *Cryphonectria parasitica*, l'agent pathogène responsable du chancre du châtaignier aux Etats-Unis, qui a détruit l'équivalent de 9 millions d'hectare d'arbres [5] ; et l'introduction en 2013 du *Xylella fastidiosa* en Italie qui décime les oliviers dans le pays menaçant ainsi la filière [6].

Afin de développer des stratégies de gestion d'une épidémie, la connaissance de la biologie des populations de l'agent pathogène, notamment la génétique des populations (diversité et structure génétique) et ses traits biologiques (aptitude à la compétition,

pouvoir pathogène, etc.) s'avèrent nécessaires. Pour cela, de nombreuses méthodes de génotypage ont été développées permettant de différencier les souches et les populations au sein d'une espèce bactérienne donnée [7].

Le flétrissement bactérien causé par les bactéries du complexe d'espèces *Ralstonia solanacearum* (ceRs), est considéré comme étant la deuxième phyto bactériose la plus dommageable au monde [8]. Ces bactéries attaquent plus de 200 espèces végétales [9] dont nombreuses sont d'intérêt économique telles que la pomme de terre, la tomate, la banane, l'arachide, le haricot, le géranium et l'eucalyptus. De plus, le ceRs est mondialement réparti sévissant en milieu tropical, subtropical et tempéré [10]. À Madagascar, l'étude de la biologie des populations du ceRs a été initiée par Ravelomanantsoa lors de sa thèse de 2013 à 2016 [11]. Ses travaux réalisés dans les principaux bassins de production de pomme de terre et les zones de production maraîchères des hautes terres centrales malgaches ont montré des épidémiologies contrastées entre les souches du phylotype IIB-1 introduites et les souches endémiques de phylotype III [12]. Par ailleurs, dans l'optique de mener une étude sur la distribution du ceRs au sein des espèces agricoles dans des zones agroécologiques très diversifiées, une étude plus exhaustive couvrant une plus large prospection dans les principales zones de productions maraîchères en moyenne et basse altitude s'est avérée nécessaire. Ce qui a amené à ce projet de thèse réalisé au sein de l'équipe « Génomique et épidémiologie des agents pathogènes émergents » de l'Unité Mixte de Recherche Peuplement des Végétaux et Bioagresseurs en Milieu Tropical (UMR PVBMT, La Réunion) en collaboration avec le Centre National de Recherche Appliquée au Développement Rural (FOFIFA/CENRADERU, Madagascar). Les principaux objectifs de cette thèse sont la mise en lumière et la compréhension de la situation épidémiologique du ceRs à Madagascar ainsi que dans les petites îles du Sud-Ouest de l'Océan Indien (SOOI) en analysant la diversité et la structure génétique des populations responsables du flétrissement bactérien et en identifiant des traits biologiques pouvant expliquer cette situation épidémiologique. *In fine*, cette thèse contribuera à la mise en place de stratégie de prévention et de gestion des risques associés au ceRs, notamment la dissémination de clones hautement virulents aux niveaux national (Madagascar) et régional (Madagascar - petites îles du SOOI). De plus, elle contribuera à l'identification de stratégie de biocontrôle de cet agent phytopathogène.

Synthèse bibliographique

I. Epidémiologie moléculaire

Constamment exposée aux agents phytopathogènes tels que les champignons, virus et bactéries, la plante peut développer une maladie infectieuse dont la propagation rapide sur une zone géographique donnée entraînera une épidémie [13]. Non maîtrisée, celle-ci peut avoir des conséquences lourdes pour l'agriculteur. Selon la FAO, les ravageurs et maladies des plantes causent chaque année jusqu'à 40% de pertes en cultures vivrières, privant des millions de gens de nourriture et portant un grave préjudice à l'agriculture qui est la principale source de revenus pour les communautés rurales pauvres (<http://www.fao.org/plant-health-2020/about/fr/>).

Les facteurs influençant le développement d'une épidémie végétale sont classiquement représentés sous la forme d'un tétraèdre [14] dont les sommets correspondent à l'agent pathogène (caractérisé par exemple par son pouvoir pathogène - virulence, agressivité, sa résistance aux pesticides, son aptitude à la compétition, etc.), à l'hôte (sensibilité, résistance naturelle ou acquise, importance agronomique, répartition géographique), aux conditions environnementales (température, humidité, force et direction du vent, type de sol), et à l'homme (pratiques agricoles, échange et commerce de matériel végétal infecté favorisant la dispersion de l'agent pathogène) (Figure 1).

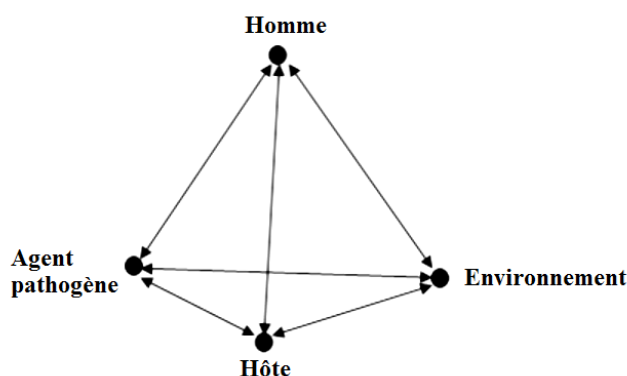


Figure 1: Facteurs influençant le développement d'une épidémie végétale

D'après [14]

Afin de mieux comprendre et maîtriser une épidémie, l'épidémiologiste se doit de caractériser les souches phytopathogènes incriminées pour avoir des informations sur leurs sources d'inoculum, leurs voies de dissémination et la dynamique d'évolution de

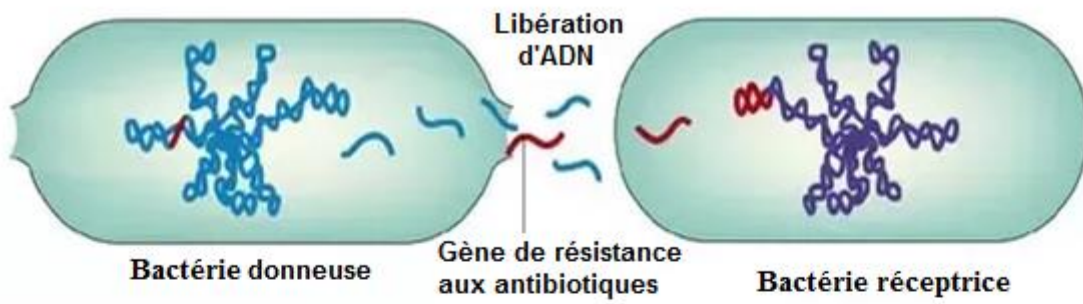
leurs populations. La différenciation des souches bactériennes a longtemps été basée sur des caractères phénotypiques conventionnels [15]. Cependant, le développement des techniques de biologie moléculaire basées sur l'analyse d'acides nucléiques a permis d'identifier des marqueurs moléculaires stables, discriminants, reproductibles et applicables à un grand nombre de souches [15]. L'utilisation de ces différents outils moléculaires à des fins épidémiologiques est à l'origine de la dénomination « épidémiologie moléculaire ».

1. Diversité et structure génétique des populations bactériennes

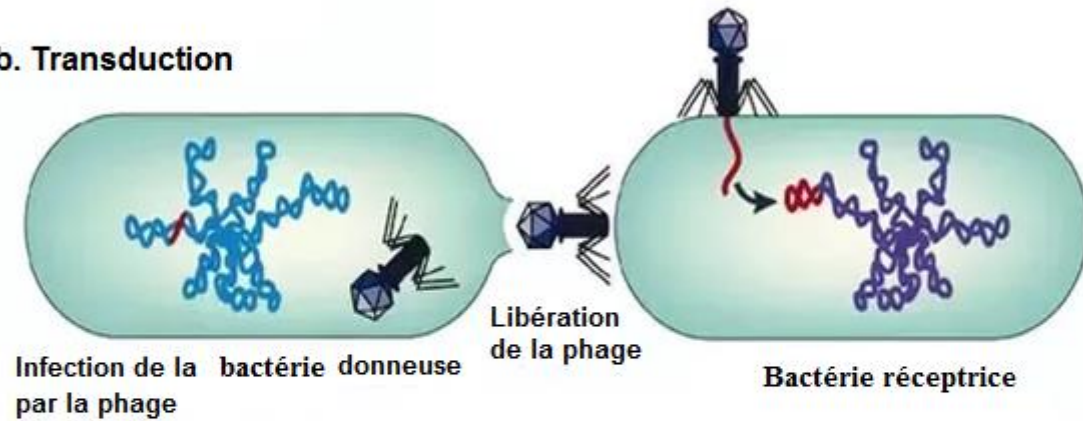
Une population peut être définie comme un ensemble d'individus dont on considère qu'il fonctionne comme une unité, du point de vue évolutif et écologique [16]. Des variations peuvent apparaître au niveau du génome des individus qui la composent amenant de la diversité génétique au sein de la population. Chez les bactéries, deux principaux mécanismes interviennent afin d'apporter de la variabilité génétique : la mutation et le transfert horizontal de gènes.

La mutation conduit à des changements dans la séquence d'ADN des gènes individuels et crée ainsi de nouveaux allèles dans la population [17]. Elle peut être bénéfique, délétère ou silencieuse. La mutation peut générer différents types de polymorphisme : *i)* le polymorphisme mononucléotidique (SNP) dû à la substitution d'une seule paire de bases à un locus donné [18]; *ii)* un polymorphisme de nombre de copies (CNV) dû à des événements de duplication ou de délétion entraînant la variabilité du nombre de copies d'une même séquence [19]; *iii)* un polymorphisme de répétitions (VNTR) dû à la délétion ou à l'insertion d'un motif répété suite à un glissement de la Taq polymérase lors de la réplification ou suite à une recombinaison [20,21]. Le transfert horizontal de gène est quant à lui un processus par lequel une bactérie va transférer du matériel génétique à une autre bactérie sans qu'elle soit sa propre descendance. Ce mode de transfert d'information génétique, indépendant de la division cellulaire, est une cause particulièrement efficace de dissémination de gènes conférant des avantages d'adaptation [22].

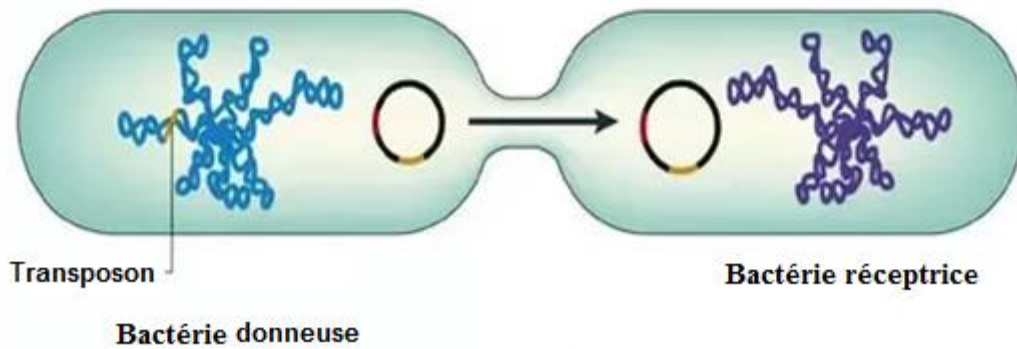
a. Transformation



b. Transduction



c. Conjugaison



Copyright © 2006 Nature Publishing Group
Nature Reviews | Microbiology

Figure 2: Mécanismes de transfert horizontal de gène chez la bactérie

D'après, [23]

Il peut se réaliser selon trois mécanismes : *i*) la transformation où un fragment d'ADN est libéré par une cellule bactérienne et absorbé par une autre, *ii*) la transduction où un fragment d'ADN est introduit dans une cellule bactérienne via un bactériophage (virus de la bactérie), et *iii*) la conjugaison où un fragment d'ADN est transmis via un contact direct entre une cellule bactérienne donneuse et une cellule receveuse (Figure 2).

La distribution de la diversité génétique peut être étudiée et quantifiée au sein de la population (diversité génétique intra-population). Les indices couramment utilisés pour décrire la diversité génétique intra-population sont : le nombre moyen d'allèles par locus (N_a), la richesse allélique A (mesurée par raréfaction) qui exprime le nombre d'allèles observés à un locus donné, le taux de polymorphisme P qui exprime le pourcentage de loci polymorphes dans la population étudiée, et l'indice de diversité non biaisé de Nei (H_{nb}) qui est défini comme étant la probabilité de tirer, au hasard, deux allèles différents à un même locus (Tableau 1).

De plus, au sein d'une espèce, différentes variables telles que la géographie et les plantes hôtes peuvent induire l'apparition de différentes populations dont les fréquences alléliques et génotypiques divergent entre elles [24]. Cette différenciation génétique entre populations (diversité inter-population) est influencée par cinq forces évolutives : la mutation, la dérive génétique et la sélection naturelle (permettant une adaptation locale) qui favorisent la divergence entre les populations tandis que la migration d'individus (flux de gènes), et la recombinaison favorisent une homogénéisation. Les indices couramment utilisés pour décrire la différenciation génétique entre populations sont : *i*) l'indice de différenciation F_{ST} de Wright [25] qui mesure la différenciation génétique en fonction des fréquences alléliques et assume que chaque mutation engendre une forme allélique unique (modèle de mutation IAM, Infinite Allele Model) [26,27]; et *ii*) le R_{ST} de Slatkin's [28] qui est une variante du F_{ST} et qui tient compte des variations alléliques et assume une mutation avec perte ou gain d'un motif (modèle de mutation SMM, Stepwise Mutation Model) [28,29] (Tableau 1).

D'après [25, 28, 30]

Tableau 1: Indices de diversité et différenciation génétique des populations

Indices de diversité génétique	Remarques
$Na = \frac{1}{l} \sum a$ <p>Avec : l = le nombre de loci étudiés a = le nombre d'allèles à un locus</p>	<p>Nombre moyen d'allèles par locus (N_a)</p>
$A = K_i - \sum_{k=1}^k P_{ik}^{(g)}$ <p>Avec : K_i = somme des allèles dans la population i $P_{ik}^{(g)}$ = probabilité pour un allèle k, observé N_k fois parmi les N gènes dans la population i, d'être absent dans un échantillon de taille réduite g telle que $g < N$.</p>	<p>La richesse allélique A (mesurée par raréfaction) exprime le nombre d'allèles observés à un locus donné.</p>
$P = \frac{\sum_{i=1}^i P_i}{N}$ <p>Avec : P_i = proportion de loci polymorphes dans une population N = nombre de populations</p>	<p>Le taux de polymorphisme P exprime le pourcentage de loci polymorphes dans la population étudiée.</p>
<p>Diploïde: $H_{nb} = \frac{2n(1 - \sum p_i^2)}{2n-1}$ Haploïde: $H_{nb} = \frac{n(1 - \sum p_i^2)}{n-1}$</p> <p>Avec : n = nombre d'individus étudiés p_i = la fréquence de l'$i^{\text{ème}}$ allèle à un locus</p>	<p>L'indice de diversité non biaisé de Nei (H_{nb}) est défini comme étant la probabilité de tirer, au hasard, deux allèles différents à un même locus.</p>
$F_{st} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}}$ <p>Avec : F_{IS} = différenciation des individus à l'intérieur de chaque population</p>	<p>Tous ces indices ont une valeur comprise entre 0 et 1.</p> <ul style="list-style-type: none"> • Si la valeur des indices de différenciation est égale à 0 alors les

Indices de diversité génétique	Remarques
F_{IT} = différenciation des individus par rapport au total	populations ne sont pas différenciées, et
$R_{st} = \frac{(S - S_w)}{S}$ <p>Avec : S = la moyenne de la différence au carré de la taille des allèles entre toutes les paires d'allèles</p> <p>S_w = moyenne de la somme au carré des différences de taille des allèles dans chaque sous-population</p>	<ul style="list-style-type: none"> • Plus la valeur tend vers 1, plus les populations sont différenciées.

Par ailleurs, selon leur processus d'évolution, les populations bactériennes peuvent présenter différentes structures : *i*) une structure clonale dans le cas de population présentant un fort déséquilibre de liaison, soit une association non aléatoire des allèles où les individus composant la population sont issus de génotypes dits fondateurs se reproduisant de manière asexuée, favorisant ainsi une faible diversité génétique au sein de la population ; *ii*) une structure panmictique dans le cas de population présentant un faible déséquilibre de liaison, soit une association aléatoire des allèles, induisant une faible différenciation génétique entre populations [31] ; et *iii*) une structure intermédiaire dite épidémique résultant d'une dominance explosive d'un génotype (clone épidémique) sur fond de population panmictique [32].

2. Intérêt de la méthode de génotypage MultiLocus Variable number of tandem repeats (VNTR) Analysis (MLVA) en épidémiologie moléculaire

De nombreuses méthodes de génotypage ont été développées pour différencier les souches impliquées dans une épidémie. Ces méthodes de génotypage peuvent être classées en deux groupes :

- les méthodes basées sur l'analyse des fragments d'ADN obtenus par digestion avec des enzymes de restriction, par amplification, ou la combinaison des deux [33]. Citons en exemple, la méthode d'analyse du polymorphisme de longueur des fragments de restriction (RFLP) où les fragments obtenus par digestion de l'ADN avec des enzymes de restriction sont séparés par électrophorèse puis hybridés

avec une sonde ADN (technique de Southern) [34], et la méthode d'analyse MLVA qui s'appuie sur l'amplification par réaction de polymérisation en chaîne (PCR) des séquences d'ADN répétées en tandem (VNTR) [35–37],

- les méthodes basées sur le séquençage de l'ADN [33] telles que les méthodes d'analyse de séquences multilocus (MLST/MLSA) qui consistent à amplifier par PCR puis à séquencer des fragments d'ADN correspondant à des gènes dits « ménage » codant des fonctions de base de la cellule.

Dans notre étude, nous allons nous intéresser particulièrement à la méthode MLVA. Elle a été initialement développée en 1997 pour tracer les isolats d'*Haemophilus Influenzae* à l'origine d'infection muqueuse chez l'Homme [38]. Depuis, elle a été développée pour pratiquement toutes les espèces bactériennes d'intérêt médical [39]. Chez les bactéries phytopathogènes, son utilisation est venue plus tardivement, pour différencier des souches de *Xylella fastidiosa* [40]. Ensuite, elle a été développée pour étudier l'épidémiologie d'autres bactéries phytopathogènes telles que : *Clavibacter michiganensis subsp. michiganensis* [41], *Pseudomonas syringae pv. tomato* et *maculicola* [42], *Xylella fastidiosa* [43], *Xanthomonas citri pv. citri* [44,45], *Xanthomonas oryzae* [46], *Xanthomonas arboricola pv. pruni* [47], *Xanthomonas phaseoli pv. manihotis* [48], *Xanthomonas vasicola pv. musacearum* [49], *Xanthomonas citri pv. viticola* [50], et le complexe d'espèces *Ralstonia solanacearum* [51–54]. La méthode MLVA consiste à faire une amplification par PCR de plusieurs VNTR répartis sur le génome bactérien, à déterminer la taille des amplicons par électrophorèse et à évaluer le nombre de répétitions à chaque locus. Ce nombre de répétitions constitue pour l'ensemble des loci le profil génétique d'une souche typée (Figure 3). Elle présente les avantages d'avoir un fort pouvoir discriminant, d'être robuste, reproductible, répétable et rapide [35–37].

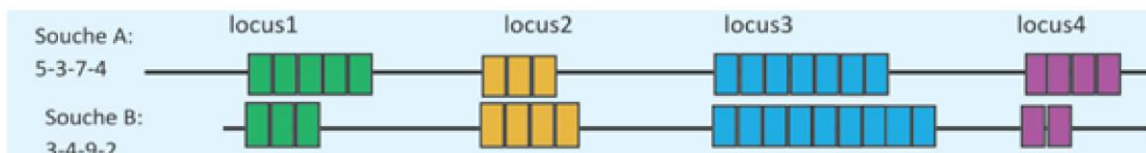


Figure 3: Profil génétique de deux souches (A et B) après analyse MLVA

D'après [55]

Le nombre de rectangles à chaque locus correspond au nombre de répétitions du VNTR. Par exemple, le locus 1 comprend 5 répétitions de VNTR chez la souche A et 3 répétitions chez la souche B.

La variation du nombre de copies des VNTR est due à deux mécanismes mutationnels : i) le crossing over inégal, fréquent pour les minisatellites¹ et ii) le glissement d'un brin d'ADN lors de la réplication suivi d'une absence de réparation (« Slipped Strand Mismatching » ou SSM) (Figure 4), fréquent pour les microsatellites² [56,57].

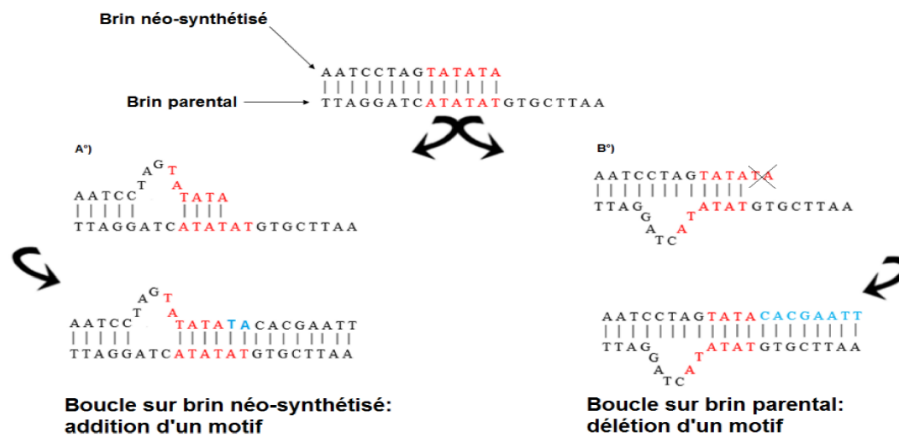


Figure 4: Modèle du Slipped Strand Mismatching (SSM)

D'après [58]

Lors de la réplication : a°) Le glissement de quelques paires de bases au niveau du brin néo-synthétisé entraîne l'addition d'un motif; b°) Le glissement de quelques paires de bases au niveau du brin parental entraîne la délétion d'un motif au niveau du brin néo-synthétisé.

Différents modèles ont été proposés pour modéliser l'évolution des marqueurs VNTR (Figure 5) : i) le modèle pas-à-pas ou Stepwise Mutation Model (SMM) [29] qui suppose que chaque mutation va créer, avec la même fréquence μ , soit la perte soit le gain d'un unique motif VNTR ; ii) le modèle de mutation par allèles infinis ou Infinite Alleles Model (IAM) qui se base sur l'idée que chaque mutation introduit un nouveau profil allélique à la fréquence μ , car il y aura ajout ou suppression de plusieurs unités de répétition nucléotidique [26]; iii) le modèle à deux phases ou Two Phase Model (TPM) qui suppose que les mutations augmentent ou diminuent la taille d'un allèle donné soit d'une unité de répétition avec la probabilité P, soit de plusieurs unités avec la probabilité 1-P [59]; iv) le modèle par pas généralisé ou Generalized Stepwise Mutation model (GSM) où la mutation

¹ Minisatellite: VNTR dont la longueur des motifs est entre 7 à 60 paires de bases (pb). [36]

² Microsatellite : VNTR dont la longueur des motifs est entre 2 à 6 pb. [36]

va ajouter ou retirer X répétitions, telle que X est une variable aléatoire suivant une loi géométrique de paramètre p (en pratique, p est de l'ordre de 0.2) [60].

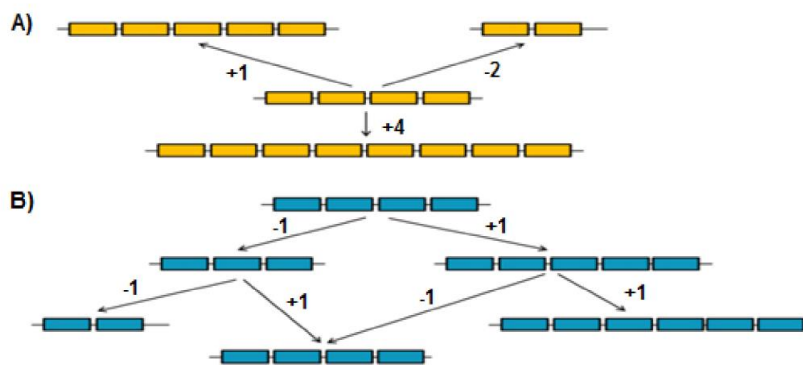


Figure 5: Principaux modèles d'évolution des VNTR. A) Modèle IAM et B) Modèle SSM

D'après [58]

Un rectangle correspond à un VNTR. Dans un modèle IAM (A), la mutation a entraîné l'ajout ou la suppression d'un ou de plusieurs VNTR. Dans un modèle SMM (B), la mutation a entraîné l'ajout ou la suppression d'un VNTR à la fois d'où la dénomination modèle pas à pas.

3. Facteurs impliqués dans l'émergence de lignée génétique phytopathogène

L'émergence d'une lignée génétique phytopathogène correspond à sa prolifération (prévalence) rapide dans un espace géographique donné [61,62]. Après l'identification des lignées génétiques les plus répandues, la détermination des facteurs pouvant expliquer ces émergences est essentielle pour l'épidémiologiste. Plusieurs cas d'étude nous montrent que les activités humaines telles que les introductions de matériels végétaux contaminés, les pratiques agricoles, et la perturbation des habitats sont à l'origine de l'émergence de nombreuses maladies infectieuses des plantes [1,61]. Citons en exemple : *i*) les cas d'introduction de la pyriculariose due à *Magnaporthe grisea* dans toutes les principales zones de production de riz, causant une perte estimée à 55 millions de dollars US/an dans le sud et sud-est de l'Asie [63]; *ii*) la pratique de la monoculture qui favorise une dispersion rapide des agents phytopathogènes en comparaison à un agroécosystème plus diversifié [1], *iii*) l'introduction d'un mutant avirulent (souche 1) du Tomato mosaic virus en Angleterre pour protéger les plants de tomates de souches plus virulentes (protection croisée) qui a entraîné l'émergence de formes virulentes de la souche 1 dans toutes les plantations commerciales où le mutant avait été libéré [64].

Les conditions météorologiques extrêmes comme les cyclones (vent violent, forte pluie) favorisent également la propagation rapide, et sur de larges distances, des agents phytopathogènes. C'est le cas par exemple de la dispersion par le vent et la pluie de *Xanthomonas axonopodis* pv. *citri* en Floride entre 2004 et 2005 [65].

Les facteurs de virulence sont des facteurs liés à la capacité des agents phytopathogènes à provoquer une maladie malgré les mécanismes de résistance de l'hôte [66]. Ils jouent donc un rôle crucial dans l'épidémiologie. Pour permettre l'infection, l'agent phytopathogène sécrète des enzymes de dégradations de la paroi cellulaire végétale, des effecteurs de type III, des sidérophores, des exopolysaccharides, des hormones et des toxines [67]. Certains facteurs de virulence sont liés à des traits d'adaptation et de survie dans les habitats non agricoles et les substrats comme le sol, les nuages, l'eau, les débris végétaux, etc [68]. C'est le cas, par exemple, de *Pseudomonas syringae* qui est capable de se maintenir dans les nuages, la pluie, la neige, et les plantes alpines sauvages en formant des biofilms, en produisant des toxines et des sidérophores [68]. Ainsi, les hôtes alternatifs peuvent constituer des réservoirs d'inoculum pour les agents phytopathogènes, mais également la capacité d'adaptation à ces hôtes alternatifs peuvent favoriser l'apparition de nouveaux traits susceptibles d'améliorer la pathogénicité et le fitness bactérien.

Le saut d'hôte qui désigne le procédé selon lequel un pathogène infecte une nouvelle espèce d'hôte [69] peut également expliquer l'émergence d'agent phytopathogène. C'est le cas par exemple, au sein du complexe d'espèces *Ralstonia solanacearum* de l'émergence de la lignée génétique phylotype IIB sequevar 4 NPB (Non Pathogenic on Banana) en Martinique qui a supplanté les lignées génétiques phylotype I et IIA historiquement présentes sur l'île [70]. En effet, jusqu'en 2002, ces souches IIB-4NPB étaient inféodées aux anthuriums et cucurbitacées, mais des missions de collectes réalisées en 2002 et 2003 ont révélé leurs émergences sur les Solanacées (30% des isolats prélevés sur ces hôtes). Ce phénomène de saut d'hôte pourrait être la conséquence du système de culture adoptée sur l'île, car les plants de tomates flétris par les souches IIB-4NPB étaient tous en succession avec des cultures de banane.

Les environnements naturels abritent une grande diversité d'espèces microbiennes, induisant une forte compétition entre les microorganismes pour l'accès aux ressources

alimentaires et aux niches écologiques [71]. Pour prospérer dans leur environnement, les bactéries déploient un large éventail de mécanismes qui nuisent et tuent leurs concurrents [72]. Citons en exemple les effecteurs de type VI qui sont sécrétés par la bactérie lorsqu'elle est au contact d'un concurrent, entraînant la lyse cellulaire [73] et les bactériocines qui sont des substances protéiniques ribosomiques à activité antimicrobienne [74]. Les bactériocines tuent (effet bactéricide) ou inhibent (effet bactériostatique) le développement d'autres bactéries de la même espèce ou d'une espèce apparentée [75]. Plusieurs études ont montré l'avantage compétitif que procure la production de bactériocine sur des souches sensibles. C'est le cas par exemple de la souche de *Rhizobium* VRF10, productrice de bactériocine inhibant la croissance de la souche sensible VRF76 en co-inoculation dans du bouillon nutritif. De plus, la co-inoculation des deux souches VRF10 et VRF76 dans des graines de *Vigna radiata* qui sont semées puis repotées et mises en croissance pendant 40 et 50 jours a montré que la souche VRF10, productrice de bactériocine, occupait un nombre significativement plus important de nodules que la souche sensible VRF76, suggérant le rôle de la production de bactériocine dans cette compétitivité [76]. De plus, des études ont montré l'influence des bactériocines dans la prévalence des souches de la race T3 de *Xanthomonas perforans* qui ont supplanté les souches T1 dans les cultures de tomate sous serre et en plein champ de Floride [77,78].

II. Bactériocines

1. Généralités sur les bactériocines

La première bactériocine a été décrite par Gratia en 1925 lorsqu'il a observé l'activité d'inhibition de souches d'*Escherichia coli* S par une substance thermostable provenant d'une culture d'*Escherichia coli* V. Cette substance a été nommée colicine en référence à la bactérie productrice. La principale différence entre les bactériocines et les antibiotiques est leur spectre d'action respectif, car les bactériocines ne tuent ou n'inhibent que la croissance des bactéries apparentées à la souche productrice contrairement à la majorité des antibiotiques [75,79].

La majorité des bactériocines est produite durant la phase exponentielle de croissance de la bactérie [80–82]. Cependant pour certaines bactériocines, il a été démontré que leurs productions avaient lieu pendant la phase stationnaire [83–85]. C'est le cas par exemple de la lactostrepcine [83] et la plantaricine T [85]. De plus, la production de bactériocines présente un coût pour la bactérie productrice, car sa biosynthèse induit souvent un ralentissement de la croissance ou à la mort cellulaire dans le cas où l'exportation de la toxine dans le milieu se fait par lyse cellulaire [86,87].

La biosynthèse des bactériocines est soit constitutive, comme pour la lacticine 3147 produite par *Lactococcus lactis* [88] et l'enterocine B produite par *Enterococcus faecium* BFE 900 [89], soit régulée par différents mécanismes dont le plus fréquent est la régulation par « quorum sensing » [90–95]. Ce dernier mode de production des bactériocines est le plus prévalent chez les bactéries [94]. Le quorum sensing est un mécanisme par lequel les bactéries régulent l'expression des gènes liés à la biosynthèse de bactériocines en fonction de leur densité cellulaire. Les bactéries mesurent la densité de leurs populations en détectant la concentration des molécules dites de signalisation qu'elles produisent elles-mêmes dans leur environnement. À un certain seuil de concentration, correspondant à une densité de population critique, ces signaux servent d'inducteurs, régulant la transcription des gènes liés à la production de bactériocines [95]. Une étude faite par Blanchard *et al.* [93] a démontré que la production régulée par quorum sensing offre un avantage à la bactérie productrice lors de la colonisation d'un nouvel espace et une meilleure capacité de défense face à une invasion. De plus, il est à noter que la production de bactériocines se réalise également en fonction de différents paramètres

tels que la composition du milieu de culture, le pH, et l'environnement (température, etc.) [96-98].

2. Classification des bactériocines

Il a été suggéré que la majorité des lignées des eubactéries et archaeobactéries produisent un ou plusieurs types de bactériocines [99-102]. La classification de ces bactériocines diffère suivant qu'elle soit produite par une bactérie de type gram positif, de type gram négatif, ou par une archaeobactérie. Dans cet état de l'art, nous allons nous intéresser particulièrement aux bactériocines produites par les bactéries de type gram positif et négatif.

2.1. Bactériocines produites par les bactéries de type gram positif

Les bactériocines produites par les bactéries de type gram positif peuvent-être classées en : classe I composée de peptides ayant subi des modifications post-traductionnelles, la classe II composée de peptides thermostables et n'ayant pas subi des modifications post-traductionnelles, et la classe III qui est composée de larges molécules sensibles à la chaleur [103,104].

En 2019, Kumaryia a proposé un tableau résumant la classification des bactériocines et leurs modes d'action [105]:

Tableau 2: Classifications des bactériocines produites par les bactéries de type gram positif

Classe		Caractéristiques	Modes d'action	Références
I	Ia	appelée lantibiotique (<5kDa), les peptides contiennent du lanthionine et du β -methyl lanthionine (d'où le nom lantibiotique)	Formation de pores au niveau de la membrane cytoplasmique, rendant la cellule perméable	[106-109]
	Ib	appelée labyrinthopeptine, les peptides contiennent du labyrinthine et du labionine	Non connus	[110]

Classe		Caractéristiques	Modes d'action	Références
	Ic	appelée sactibiotique, les peptides contiennent du soufre et du carbone alpha	Non connus	[111]
II	Ila	appelée "pediocin like", car ces bactériocines ont une forte activité contre <i>Lysteria monocytogenes</i> à l'instar de la pédiocine PA-1 (ou AcH). Ce sont des peptides thermostables, synthétisés sous formes de précurseurs, la région N-terminale renferme une séquence consensus surnommée "pediocin box".	Formation de pores au niveau de la membrane bactérienne rendant la cellule perméable	[112-116]
	Ilb	appelée "two peptides", ces bactériocines sont composées de deux peptides formant un complexe. Seuls, les peptides sont peu ou pas actifs.		[117-119]
	Ilc	Peptides circulaires.		[120-122]
	Ild	Peptides linéaires, certains peptides antimicrobiens sont synthétisés sans une séquence N-terminale d'où leur nom "leaderless bacteriocins".		[123-125]
III		Larges molécules sensibles à la chaleur.		[126-128]

2.2. Bactériocines produites par les bactéries de type gram négatif

Les bactériocines produites par les bactéries de type gram négatif peuvent-être classées en : microcine (<10 kDa), colicine (25–80 kDa) et tailocine (bactériocine de type queue de phage) [129].

Les **microcines** sont des toxines ayant une masse moléculaire inférieure à 10 kDa. Elles sont produites en situation de manque de nourriture, de manque d'oxygène, de surpopulation, ou de manque de disponibilité en fer dans le milieu qui induisent l'expression des régulateurs impliqués dans la transcription des gènes intervenants dans sa biosynthèse [130]. Le cluster de gènes impliqué dans cette biosynthèse est typiquement composé du gène « précurseur de la microcine », et des gènes codant l'« immunité », la « sécrétion des protéines », et la « production des enzymes de modification post-traductionnelle ». Ces gènes sont localisés au niveau du plasmide ou du chromosome [131]. Les microcines sont généralement produites lors de la phase stationnaire de croissance de la bactérie [130]. Elles sont réparties en deux classes [129,130]:

- la classe I qui est composée des peptides ayant une masse moléculaire inférieure à 5 kDa, et dont la structure est soumise à diverses modifications post-traductionnelles.
- la classe II qui est composée des peptides ayant une masse moléculaire comprise entre 5 et 10 kDa. Elle est encore divisée en deux classes : i) la classe IIa dont le cluster de gènes impliqué dans la biosynthèse de microcine se situe au niveau du plasmide, les peptides n'ont pas subi de modification post-traductionnelle et il est possible qu'il y ait des liaisons disulfures entre les peptides ; et ii) la classe IIb dont le cluster de gènes impliqué dans la biosynthèse de microcine se situe au niveau du chromosome, et la microcine a une structure linéaire portant un terminal C sidérophore dû à une modification post-traductionnelle.

Les microcines sont exportées hors du cytoplasme *via* les transporteurs ABC (ATP Binding Cassette) ou les pompes membranaires d'efflux [130]. Par la suite, elles pénètrent la bactérie cible en détournant des récepteurs impliqués dans l'absorption de nutriments essentiels, comme les récepteurs des sidérophores du fer (FhuA, FepA, Cir et Fiu) ou la porine OmpF [130]. Arrivée dans la bactérie, les toxines tuent en inhibant des enzymes vitales à la bactérie (cas des microcines de la classe I) ou en agissant contre la membrane interne de la bactérie (cas des microcines de la classe II). Par exemple, la MccJ25, une

microcine de la classe I, bloque la synthèse de l'ARN par obstruction du canal secondaire de l'ARN polymérase ; tandis que la microcine V, appartenant à la classe IIa, modifie la perméabilité membranaire de la bactérie cible [130].

Les **colicines** sont des toxines ayant une masse moléculaire comprise entre 25 et 80 kDa. Elles sont produites par des souches d'*Escherichia coli* possédant un plasmide colicinogénique appelé pcol [87]. Selon la taille et les caractéristiques du plasmide, on distingue les pcol de type I et II. Le pcol de type I est un plasmide de petite taille allant de 6 à 10 kb, présent jusqu'à 20 copies dans la cellule et pouvant être amplifié et mobilisé à l'aide de plasmide conjugatif [87]. Il code principalement les colicines du groupe A [87]. Le pcol de type II est quant à lui une large monocopie de plasmide d'une taille d'environ 40kb. Il est conjugatif et favorise le transfert horizontal de gènes. Ce large plasmide peut contenir un ou plusieurs opérons disposés côte à côte. Le pcol de type II code généralement les colicines du groupe B [87].

Un seul gène ou un cluster de gènes sont impliqués dans la biosynthèse des colicines. Le premier gène de tous les opérons colicines correspond au gène qui code la « production » de la toxine appelé *cx_a*, pour colicine X activité [87]. L'exception est le cas des colicines B porogènes (pore forming colicin) où le *cx_a* est l'unique gène impliqué dans la biosynthèse de la toxine. Pour les colicines nucléases, l'opéron est typiquement composé du gène qui code la « production » de la toxine *cx_a*, du gène « d'immunité » appelé *cx_i* pour colicine X immunité ou *imX*, qui code une protéine immunisant la bactérie productrice, et du gène « lyse » qui code pour une protéine de lyse cellulaire permettant la libération de la colicine (cas de toutes les colicines du groupe A et de quelques colicines du groupe B telles que les colicines 5, 10 et D) [87]. Ainsi, la libération de cette toxine est létale pour la bactérie productrice [87].

La biosynthèse des colicines est régulée par le système SOS qui est un mécanisme de réponse de la bactérie en condition de stress impliquant des dommages de l'ADN, ou un problème de réplication de l'ADN [132]. Deux protéines LexA et RecA servent respectivement de répresseur et d'activateur de la transcription des clusters de gènes impliqués dans la production des colicines [87,132].

Dans le milieu, les toxines se lient à des récepteurs spécifiques à la surface de la cellule sensible et sont transférées à l'intérieur de celle-ci via deux systèmes appelés Tol et Ton

[87]. Le système Tol est composé d'un complexe de 5 protéines tandis que le système Ton est composé de 3 protéines. Ces protéines formant le système Tol et Ton sont localisées au niveau de la membrane interne de la bactérie, et contribuent à l'entrée des colicines dans la cellule cible aidée de la force proton motrice (fournisseur d'énergie) [87]. Les colicines tuent *via* trois mécanismes : en formant des canaux dans la membrane cytoplasmique, en dégradant l'ADN et l'ARN cellulaire, ou en inhibant la synthèse de protéines [87]. Les bactéries cibles sont les autres souches d'*Escherichia coli* et les entérobactéries qui lui sont apparentées. Les colicines sont classées en deux grands groupes : la colicine A qui est transférée dans la bactérie sensible par le système Tol, qui est codée par un petit plasmide (pcol de type I), et qui est relâchée dans le milieu de culture après mort de la cellule productrice par lyse cellulaire et la colicine B qui est transférée dans la bactérie sensible par le système Ton, codé par un large plasmide (pcol de type II) [87]. Comme les bactéries qui sécrètent la colicine B n'ont pas de gènes codant pour une lyse cellulaire, la libération de la toxine dans le milieu de culture pourrait être due à un prophage qui est un virus bactérien incorporé dans le génome de la bactérie [133,134].

Diverses bactéries, autres qu'*Escherichia coli*, produisent également des bactériocines ayant des caractéristiques similaires aux colicines. Citons en exemple *Pseudomonas pyogenes* sécrétant la pyocine, *Yersinia pestis* sécrétant la pesticine, *Klebsiella* sécrétant la klebicine ou klebocine, et *Photobacterium luminescens* sécrétant la luminicine [135].

Les **tailocines** ou **bactériocines de type queue de phage** sont des bactériocines de grande taille ($2 \times 10^6 - 1 \times 10^7$ Da) ressemblant à des phages sans tête [136,137]. Elles sont résistantes aux protéases, sont thermolabiles et sédimentables par centrifugation [138,139]. Typiquement, le cluster de gènes impliqué dans sa biosynthèse est composé de gènes codant les protéines qui composent et favorisent l'assemblage des différentes structures de la particule, les gènes de régulation qui contrôlent sa transcription, et les gènes de lyse cellulaire permettant la libération de la bactériocine [136].

Il existe deux types de tailocines : le type R et le type F (Figure 6). Le type R est composé de particules contractiles apparentées aux queues de phages des *Myoviridae*. Sa structure est formée d'un long tube entouré d'une gaine, et à son bout se présente une plaque basale où est fixée des fibres caudales qui sont des protéines permettant à la bactériocine de reconnaître et de se fixer sur la membrane cellulaire de la cellule cible (RBP = receptor-

binding proteins) [136]. Le type F quant à lui est composé de particules non contractiles apparentées aux queues de phages des *Siphoviridae*. Sa structure est simple, uniquement formée d'un tube sans gaine, mais il possède également un RBP (receptor-binding proteins) qui lui permet de reconnaître et de se fixer sur la membrane cellulaire de sa cible [136]. La régulation de la biosynthèse des bactériocines de type queue de phage se fait probablement par le système de régulation SOS d'après les études menées sur les pyocines de type R et F [140] et la diffocine, une tailocine de type R, produite par *Clostridium difficile* [141,142]. Les tailocines tuent la bactérie cible en perforant la membrane cellulaire [136].

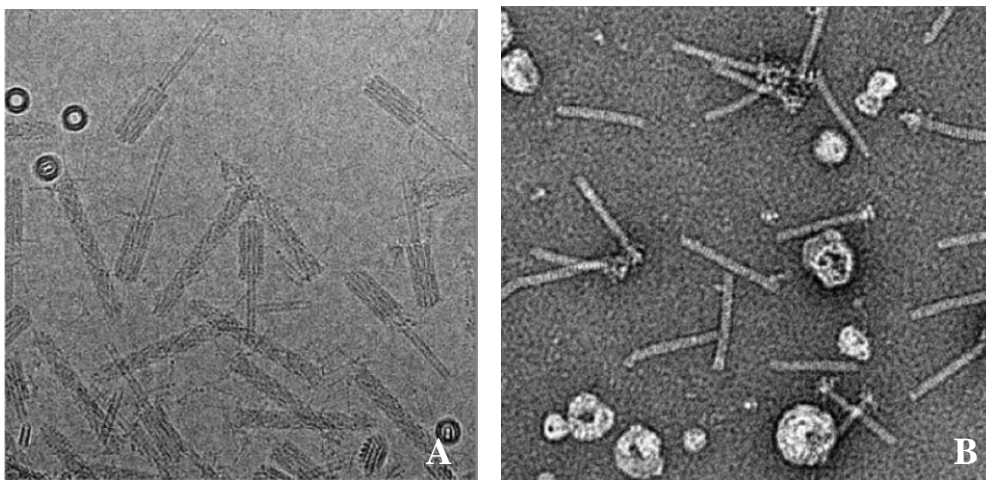


Figure 6: Micrographie électronique des tailocines. A) R2 pyocine; B) monocine de type F issue du *Listeria monocytogenes*

D'après, [136]

3. Rôle biologique des bactériocines

La production de toxines telles que les bactériocines est un des mécanismes utilisés par la bactérie pour établir sa dominance envers ses compétiteurs [143]. Compte tenu de son spectre d'action souvent étroit, il a été proposé que son rôle principal était d'assurer la médiation intraspécifique entre souches apparentées [102]. Les bactériocines peuvent permettre aux souches productrices d'envahir une communauté microbienne établie. Elles peuvent également jouer un rôle défensif et empêcher l'invasion d'autres souches ou espèces dans une niche occupée ou limiter la progression des cellules voisines [99]. Par exemple, Smidt et Vidaver [144] ont mis en évidence une forte inhibition de croissance d'une souche sensible n°16 à la bactériocine syringacine W-1 produite par la souche de *Pseudomonas syringae* PsW-1 en condition de co-inoculation dans des tiges de haricot rouge. De plus, Dorosky *et al.* [139] ont démontré que la production de R-tailocine

par *Pseudomonas chlororaphis* 30-84 lui confère un avantage compétitif par rapport à *Pseudomonas putida* KT2440 dans la rhizosphère et en biofilm. En outre, une étude faite par Kommineni *et al.* [145] a montré que la production de bactériocine 21 par la bactérie commensale *Enterococcus faecalis* portant le plasmide pPD1 favorise sa compétitivité vis-à-vis d'autres souches d'enterococci dans le tractus gastro-intestinal de la souris. Cependant, la production de bactériocines ne s'avère pas être constamment favorable pour la souche productrice. En effet, en cas de présence de souches résistantes aux bactériocines dans le milieu, les souches peuvent interagir suivant la théorie du jeu du "pierre-papier-ciseaux": les souches productrices de toxines peuvent envahir une population de cellules sensibles parce que la population sensible est empoisonnée par la toxine. Les souches productrices de toxines peuvent à leur tour être envahies par des souches résistantes qui ont un coût métabolique inférieur. Les souches résistantes, cependant, peuvent être envahies par les souches sensibles, qui n'ont pas besoin de "payer" les coûts de la résistance. Par conséquent, on peut supposer que trois souches ayant les propriétés susmentionnées présentent des oscillations dans leurs densités de population [146].

Par ailleurs, Huerta *et al.* [97] ont suggéré un rôle de la bactériocine dans la spécialisation de niche. En effet, son étude a montré que la souche GMI1000 (I-18) de *Ralstonia pseudosolanacearum* originaire de la Guyane française et la souche K60 (IIA) de *Ralstonia solanacearum* endémique du sud-est des Etats-Unis, produisent une bactériocine qui inhibe en condition tropicale (28°C) la croissance de la souche UW551 de *Ralstonia solanacearum* (IIB-1), endémique des hautes terres d'Amérique du Sud. Pourtant cette dernière se développe bien en condition tropicale (28°C) sans compétition avec les souches GMI1000 et K60.

4. Application des bactériocines

4.1. Dans le domaine de l'industrie agroalimentaire

Les bactériocines sont utilisées en industrie agroalimentaire en tant que bio-conservateurs. En effet, leur utilisation permet de réduire l'addition de conservateur chimique dans les aliments, et permet de diminuer l'intensité de leurs traitements à la chaleur, ce qui favorise une conservation naturelle et une plus grande richesse organoleptique et nutritionnelle des aliments [147]. Les bactériocines peuvent être

ajoutées sous forme de préparation concentrée ou bien produite *in situ* par la bactérie [147]. Les recherches sur les bactériocines utilisées en agroalimentaire se sont principalement focalisées sur les bactériocines produites par les bactéries lactiques [148,88]. Une bactérie lactique est une bactérie gram positif dont le produit final de fermentation est l'acide lactique. Ce groupe de bactéries est composé de plusieurs genres comme *Lactococcus*, *Lactobacillus*, *Enterococcus*, *Streptococcus*, *Leuconostoc* et *Pediococcus* [149]. L'intérêt des bactéries lactiques réside dans le fait qu'elles sont à l'origine de nombreuses transformations microbiennes retrouvées dans les produits alimentaires fermentés les plus courants comme la bière, le fromage cheddar et la sauce soja [149]. De plus, les acides organiques produits par ces bactéries sont défavorables à la croissance de divers microorganismes. Enfin, les bactériocines produites par les bactéries lactiques possèdent l'avantage d'être [149] : *i*) reconnues comme sans danger pour la santé humaine, *ii*) non toxiques pour les cellules eucaryotes, *iii*) désactivées par les protéases digestives donc ayant peu d'influence négative sur le microbiote intestinal, *iv*) généralement tolérantes à une large gamme de température et pH, *v*) actives contre une gamme relativement large de bactéries pathogènes alimentaires, et *vi*) les clusters de gènes à l'origine de la biosynthèse des bactériocines sont localisés au niveau du plasmide donc facilement manipulables.

La nisine est la première bactériocine à avoir été isolée et approuvée pour une utilisation alimentaire par la FAO/WHO et plus d'une cinquantaine de pays. Sa commercialisation date de l'année 1953 et elle figure dans la liste des additifs alimentaires en tant que conservateur E234 [149]. Cette bactériocine de la classe Ia est produite par *Lactococcus lactis* et possède une activité inhibitrice contre divers pathogènes alimentaires tels que *Listeria monocytogenes*, *Bacillus cereus*, *Staphylococcus aureus*, *Clostridium perfringens* et *Clostridium botulinum* à une faible dose (de l'ordre des nmol/L) [150,151]. De plus, la nisine est thermostable, non toxique et dégradée par les protéases digestives [150]. Sur le marché, elle est commercialisée sous différentes appellations telles que Nisaplin (2.5% nisine) (www.dupontnutritionandsciences.com) et Nisin Z®P (>95% nisine) (www.handary.com). La nisine est utilisée, par exemple, dans les tartines de fromages pasteurisés pour prévenir la croissance des spores et la production de toxines par *Clostridium botulinum*, pour prolonger la durée de conservation des produits laitiers, et pour prévenir la détérioration des produits en conserve [149].

4.2. Dans le domaine de la santé humaine et animale

Selon l'Organisation Mondiale de la Santé (OMS), la résistance aux antibiotiques constitue aujourd'hui l'une des plus graves menaces pesant sur la santé mondiale (<https://www.who.int/fr/news-room/fact-sheets/detail/antibiotic-resistance>). En effet, l'utilisation accrue d'antibiotiques à large spectre, sans respect des bonnes normes d'usage a favorisé l'adaptation de certaines souches bactériennes, les rendant ainsi résistantes (<https://www.who.int/fr/news-room/fact-sheets/detail/antibiotic-resistance>). D'où l'intérêt croissant pour les bactériocines qui sont des toxines capables de tuer ou ralentir la croissance des bactéries apparentées aux bactéries productrices [99].

Quelques études se sont focalisées sur l'utilisation des bactériocines pour lutter contre les bactéries pathogènes de l'homme et des animaux. Ces études ont montré que pour la santé humaine, les bactériocines sont efficaces pour inhiber la croissance de quelques agents pathogènes permettant ainsi d'arrêter une infection. Un exemple bien connu de l'antibiorésistance est la résistance à la méticilline, dérivé semi-synthétique de la pénicilline, pour le traitement des infections staphylococciques, qui sont apparues un an à peine après son utilisation en 1959. Par la suite, la vancomycine était considérée comme l'antibiotique de choix pour lutter contre ces bactéries. Cependant, en 1996 est apparue la première souche de *Staphylococcus aureus* résistante à cet antibiotique [152]. Depuis, quelques études se sont penchées sur la substitution de la méticilline et de la vancomycine par d'autres substances antimicrobiennes. En 2009, Kwaadsteniet *et al.* [153] ont montré une possible utilisation de la bactériocine Nisine F dans la lutte contre les infections staphylococciques. En effet, cette bactériocine inhibe la croissance des souches de *Staphylococcus aureus* K dans les voies respiratoires de rats de laboratoire. De plus, le *Clostridium difficile*, un agent pathogène responsable de la diarrhée, présente actuellement une résistance aux antibiotiques à larges spectres comme la vancomycine et le métronidazole [154]. Pour pallier à ce problème, une évaluation réalisée par Rea *et al.* [155] de l'activité antimicrobienne de la thuricine CD produite par *Bacillus thuringiensis* a montré qu'elle est aussi efficace que les antibiotiques pour lutter contre *C. difficile* et elle ne présente pas de dommage significativement important sur le microbiote intestinal. En outre, certaines bactériocines ont des effets mortels sur les cellules cancéreuses [156]. Par exemple, l'azurine produite par *Pseudomonas aeruginosa* a un effet cytotoxique et

provoque l'apoptose ou la mort programmée des cellules cancéreuses [156]. De plus, la nisine [157], la subtilosine [158], la fermenticine [159], et la lacticine 3147 [160] altèrent la motilité des spermatozoïdes humains montrant ainsi leurs potentialités en tant que méthode contraceptive [161].

Par ailleurs, les bactériocines ou souches sécrétant les bactériocines sont utilisées dans le domaine de la santé animale pour favoriser la croissance des animaux : *i)* en tuant les agents pathogènes infectieux, *ii)* en favorisant la compétition pour les ressources nutritionnelles ou l'espace entraînant l'exclusion des agents pathogènes, et/ou en *iii)* modulant la défense immunitaire de l'hôte [161,162]. Par exemple, la microcine J25 est utilisée en élevage avicole pour tuer les salmonelles [163] et la colicine E1 favorise la croissance des porcelets en inhibant la croissance de souches entérotoxiques d'*Escherichia coli* [164]. De plus, l'additif alimentaire Bactocell® à base de souche viable de *Pediococcus acidilactici*, sécrétant la pédiocine, est commercialisé pour favoriser la croissance des salmonidés, crevettes, poissons, poulets et porcelets [165–168].

4.3. Dans le domaine de la phytopathologie

L'utilisation des bactériocines dans la lutte contre les bactéries phytopathogènes est encore limitée principalement en raison d'un manque de recherche sur la diversité et les mécanismes d'action de ces protéines chez les bactéries, plutôt que par leurs potentiels [169,170].

Dans la littérature, quelques études mettent en évidence l'intérêt de l'utilisation des bactériocines dans la lutte contre les bactéries phytopathogènes. Kerr et Htay [171] ont observé une atténuation de la formation de la galle du collet causée par *Agrobacterium radiobacter var tumefaciens* lorsque celui-ci est co-inoculé avec une souche apparentée non pathogène *d'A. radiobacter var radiobacter* (souche 84). Ils ont montré que cette atténuation de la formation de galle avait une corrélation forte avec la sensibilité des souches d'*A. radiobacter var tumefaciens* à la bactériocine produite par la souche 84 montrant ainsi le rôle de la bactériocine nommée agrocine 84 dans ce biocontrôle. Les gènes de synthèse et d'immunité à l'agrocine 84 sont situés au niveau du plasmide pAgK84. Ce plasmide pouvant être transféré à d'autres bactéries pathogènes, une souche mutante K1026 portant une délétion au niveau du gène de transfert a été construite [172]. Aucun brevet n'a été déposé pour la souche 84 qui a commencé à être commercialisée à

partir de 1974 par deux compagnies australiennes sous forme de cultures gélosées ou d'inoculum purs à base de tourbe [173]. Par contre, la souche K1026 est brevetée. Elle est commercialisée en Australie, Etats-Unis, Canada, Turquie, Nouvelle Zélande, Chili, Japon et dans certains pays européens [173].

Lavermicocca *et al.*, [174] ont inoculé *Pseudomonas syringae subsp. savastanoi* sur des entailles faites à des tiges d'olivier. Le traitement de ces entailles infectées à l'aide d'une bactériocine produite par *Pseudomonas syringae pv. ciccaronei* a réduit de 51 à 81% la formation de galles sur les tiges. De plus, la pulvérisation de la bactériocine à la surface des feuilles et rameaux d'olivier a affecté négativement la survie de *P. syringae subsp. savastanoi*.

Hammami *et al.*, [175] ont identifié la bactériocine Bac IH7 produite par la souche IH7 de *Bacillus subtilis* qui a une activité bactéricide contre la souche d'*Agrobacterium tumefaciens* C58, et une activité fongistatique contre la souche *Candida tropicalis* R2 CIP203. De plus, cette bactériocine a un effet désinfectant contre *Xanthomonas campestris* et *Pseudomonas syringae* sur des semences de tomate (*Lycopersicon esculentum*) et melon (*Cucumis melo*) préalablement contaminées. En outre, elle réduit significativement l'incidence de la fonte de semis dû à *Alternaria solani* sur des semences préalablement contaminées.

Rooney *et al.*, [176] ont fait exprimer la bactériocine putidacine L1 (PL1) par des plants d'*Arabidopsis* et *Nicotiana benthamiana* leur conférant ainsi une forte résistance à *Pseudomonas syringae pv. syringae* LMG5084.

III. Modèle d'étude : le complexe d'espèces *Ralstonia solanacearum*

1. Distribution géographique, spectre d'hôtes et importance socioéconomique

Le complexe d'espèces *Ralstonia solanacearum* (ceRs) est responsable de la maladie du flétrissement bactérien sur plus de 200 espèces végétales cultivées et sauvages distribuées dans 54 familles botaniques différentes (monocotylédones et dicotylédones) [9]. Il attaque des cultures d'importance socio-économique majeure telles que la pomme de terre, la tomate, la banane, le haricot et l'arachide. Le ceRs est largement distribué dans les régions tropicales, subtropicales et tempérées [10] (Figure 7). Par ailleurs, aux États-Unis, il est considéré comme agent potentiel de bioterrorisme [177] et en Europe comme organisme de quarantaine A2 (<http://www.eppo.org/QUARANTINE/listA2.htm!>).

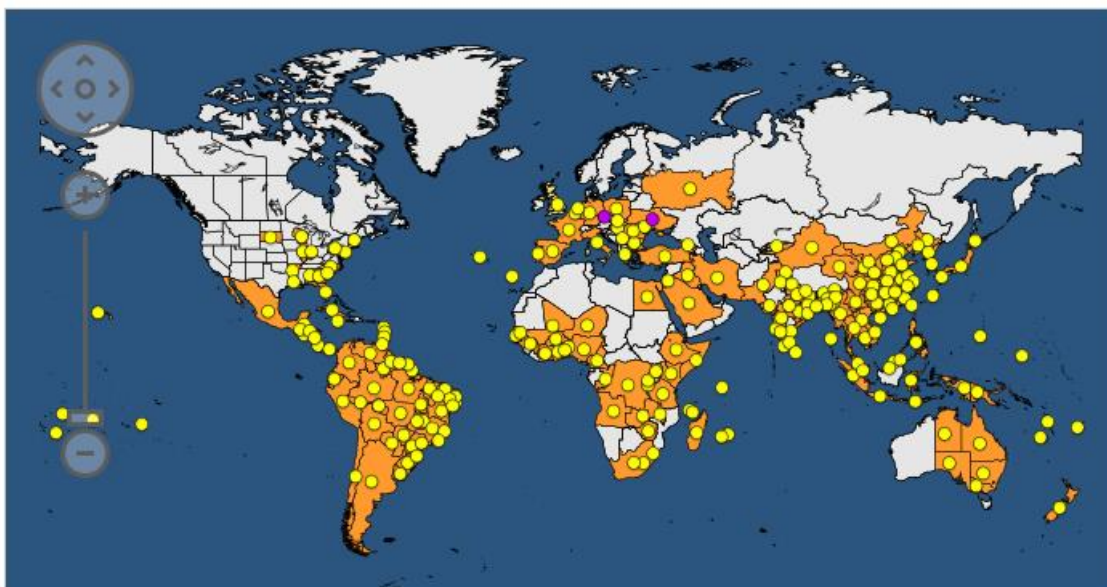


Figure 7: Répartition géographique du ceRs

D'après, <https://gd.eppo.int/taxon/RALSSO/distribution> (latest update : 03-06-22)

2. Caractéristiques morphologiques et génomiques

Les souches du ceRs sont des β -protéobactéries, aérobies de type gram négatif [178]. Elles se présentent sous forme de bâtonnet à contour arrondi, de 0.5 à 0.7 μm de largeur et de 1.5 à 2.5 μm de longueur [179]. La cellule bactérienne possède un ou plusieurs flagelles polaires [180] (Figure 8 A). Sur le milieu Kelman contenant du chlorure de triphényltétrazolium (TTC), après 2 à 3 jours d'incubation à 28°C, une colonie virulente

du ceRs a un contour irrégulier, est muqueuse et présente une tache rose au milieu de la colonie et un large bord blanc (Figure 8B), tandis que la colonie d'une souche avirulente de *R. solanacearum* est ronde, rugueuse et présente une tache rouge foncé au milieu de la colonie et un bord blanc étroit (Figure 8C) [181].

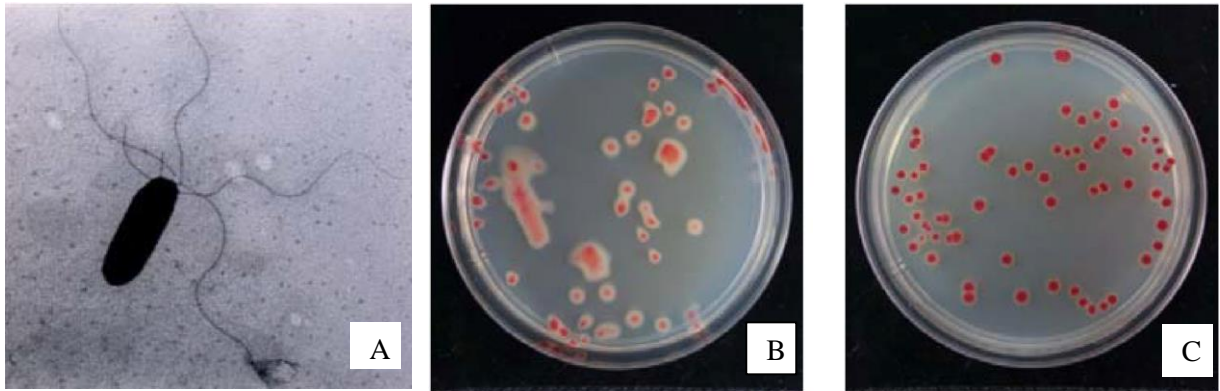


Figure 8: A) Image au microscope électronique du ceRs ; B) Colonies virulentes du ceRs sur un milieu contenant du TTC ; C) Colonies avirulentes du ceRs sur un milieu contenant du TTC.

D'après, A) [182]; B) et C) [181].

Le génome de la bactérie est constitué d'un chromosome d'environ 3,7 Mégabase (Mb) et d'un mégaplasmide d'environ 2,1 Mb [183]. Le chromosome héberge 80% des gènes de ménage qui sont des gènes indispensables pour le fonctionnement de la cellule bactérienne et le mégaplasmide héberge plusieurs gènes qui sont impliqués dans la virulence de la bactérie et l'adaptation à l'environnement [183].

3. Taxonomie et diversité phylogénétique

En 1995, l'agent pathogène responsable du flétrissement bactérien prend le nom de *Ralstonia solanacearum* [184]; il était auparavant nommé *Pseudomonas solanacearum* (1914) puis *Burkholderia solanacearum* (1992) [185]. En raison de sa très large diversité

génétique et phénotypique, *R. solanacearum* est considéré par la communauté scientifique comme un complexe d'espèces constitué de trois espèces et de cinq groupes phylogénétiques majeurs nommés phylotypes : i) *R. pseudosolanacearum* rassemble les souches du phylotype I originaire d'Asie et du phylotype III originaire d'Afrique ; ii) *R. solanacearum* regroupe les souches des phylotypes IIA et IIB originaires des Amériques ; iii) *R. syzygii* comprend les souches du phylotype IV originaire d'Indonésie - Australie - Japon [186,187] (Figure 9). À une échelle plus fine, chaque phylotype peut être subdivisé

en lignées phylogénétiques nommées *sequevars* qui rassemblent les souches dont les séquences partielles du gène codant l'endoglucanase (*egl*) divergent à moins de 1% [188]. Actuellement, 71 *sequevars* ont été définis (Poussier, communication personnelle).

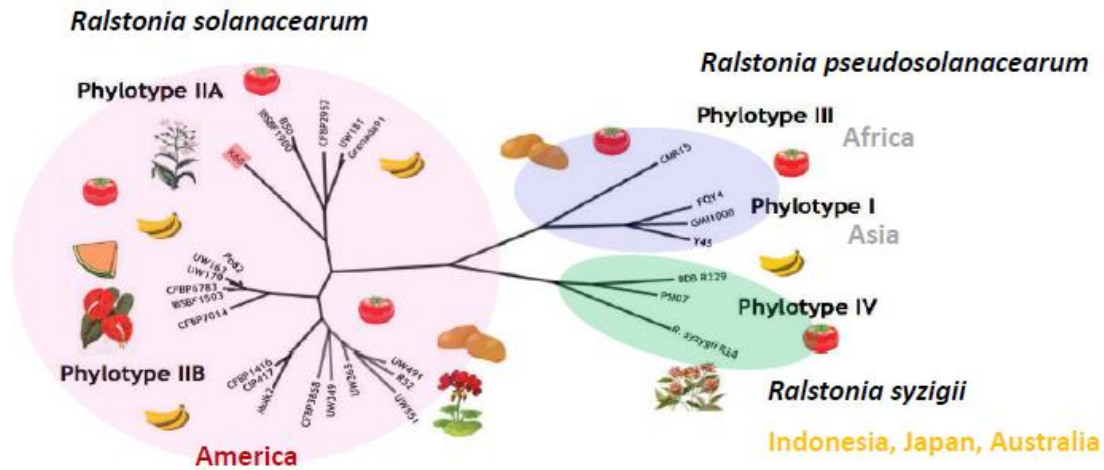


Figure 9: Classification du ceRs

D'après, [186,187]

4. Cycle infectieux et symptomatologie

Le ceRs est d'origine tellurique et infecte la plante au niveau de ses racines (Figure 10). Attiré par les exsudats racinaires de la plante hôte [189], il pénètre dans les racines *via* les ouvertures naturelles ou les blessures artificielles [190]. Dans la plante, le ceRs colonise d'abord les espaces intercellulaires du cortex interne racinaire puis envahit les vaisseaux du xylème. La bactérie s'y multiplie abondamment pouvant atteindre une densité de 10^{10} CFU/g de tige fraîche [191] et produit des exopolysaccharides (EPS) qui vont obstruer les vaisseaux et bloquer la circulation de l'eau dans la plante. Par conséquent, la plante flétrit irréversiblement et la bactérie retourne au sol pour vivre en mode saprophyte. Les symptômes dus à l'infection (Figure 11) peuvent varier suivant le type de plante hôte, mais plus fréquemment, les feuilles de la plante infectées par le ceRs flétrissent, rabougrissent et jaunissent [180]. Elles peuvent également se plier vers le bas en épinastie [192]. Des nécroses sont visibles au niveau des vaisseaux vasculaires infectés [180] et lorsque la tige est sectionnée transversalement un exsudat bactérien visqueux de couleur blanc-crème (EPS) peut apparaître.

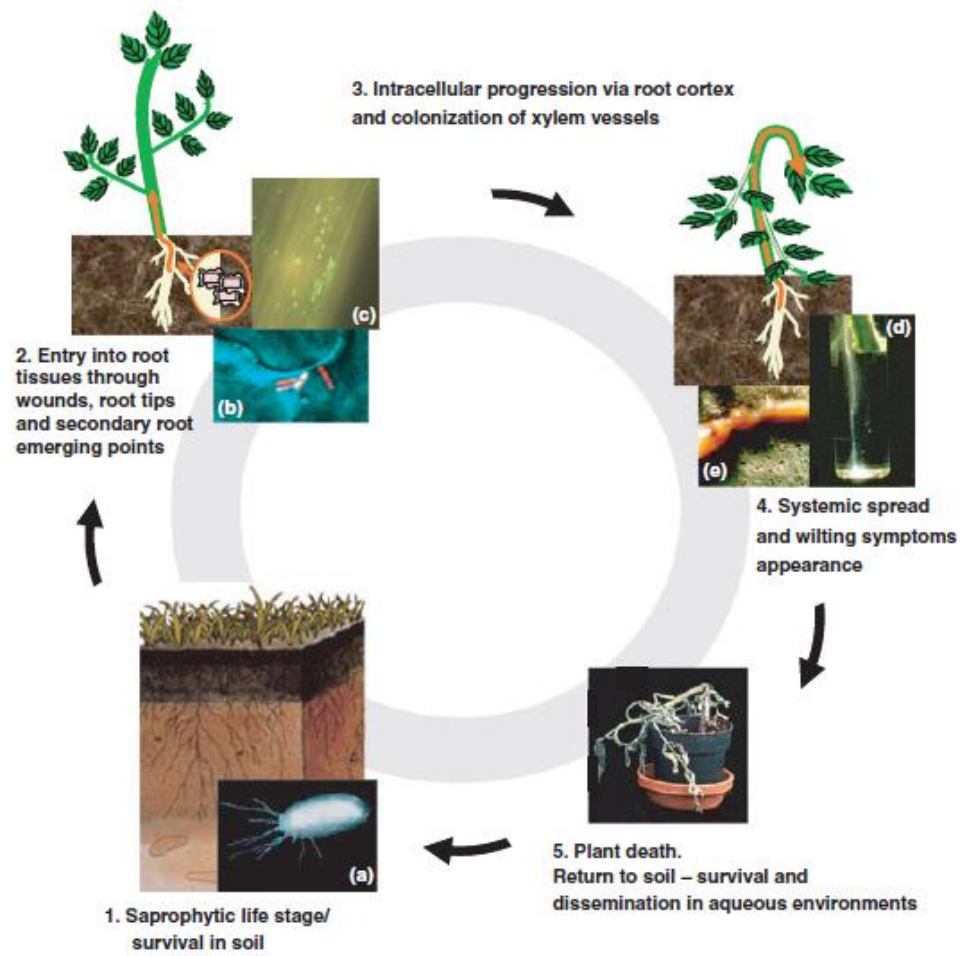


Figure 10: Cycle infectieux du ceRs

D'après, [193]

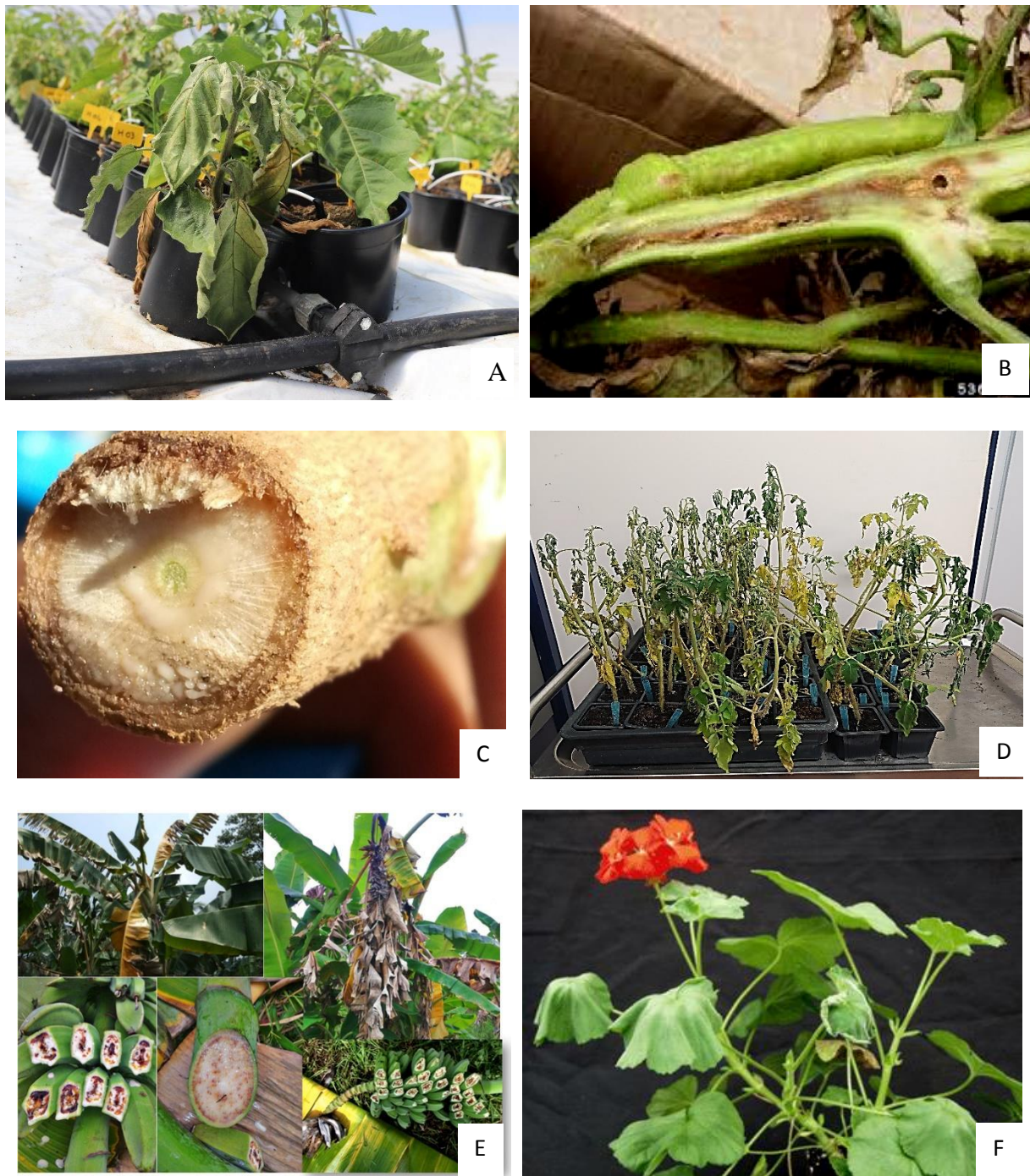


Figure 11: Symptômes du flétrissement bactérien causé par le ceRs sur différentes plantes hôtes.

A) Flétrissement irréversible d'un plant d'aubergine. B) Nécroses visibles au niveau des vaisseaux vasculaires d'un plant de tomate infecté. C) Exsudat bactérien au niveau d'une tige d'aubergine. D) Jaunissement et épïnastie des feuilles de tomates infectées. E) Symptômes du flétrissement bactérien sur bananier. F) Flétrissement irréversible de plant de géranium.

D'après, A) C) D) Auteur. B) [192]. E) [194]. F)

https://plantpath.ifas.ufl.edu/rsol/Trainingmodules/SWGeranium_Module.html

5. Stratégies de lutte et de gestion

Plusieurs mesures prophylactiques permettant de prévenir ou réduire fortement l'extension rapide du ceRs sont à considérer étant donné l'inefficacité des traitements chimiques et leurs impacts néfastes sur l'environnement. Tout d'abord, le matériel végétal à utiliser doit être sain, l'eau d'irrigation contrôlée et les matériels agricoles à utiliser désinfectés, car ils sont les principaux moyens de propagation de la bactérie. Ensuite, la pratique de la rotation de cultures avec des plantes non hôtes telles que le maïs (*Zea mays*), le riz (*Oryza sativa*), et le soja (*Glycine max*) sont à préconiser afin de réduire le potentiel d'inoculum présent dans le sol [195]. Cette mesure est particulièrement efficace contre les souches IIB-1 qui s'y conservent mal. En outre, l'utilisation de variétés résistantes est une stratégie de choix pour lutter contre le flétrissement bactérien. Depuis des décennies, les sélectionneurs ont recherché et exploité différentes sources de résistance chez les solanacées, afin de créer des cultivars présentant un niveau élevé de résistance au flétrissement bactérien, une bonne qualité agronomique et une bonne adaptation aux conditions tropicales [196]. Par exemple, à la Réunion, Lebeau a testé les interactions entre une collection de 30 accessions de Solanacées (tomate, aubergine, piment), appelée Core-TEP, et une collection de 12 souches représentatives de la diversité du ceRs (phylotypes I, II et III), appelée Core-Rs2 [196]. Ses travaux ont permis la détection dans la population d'aubergine MM738 (sensible) × AG9125 (résistante), d'un facteur de résistance majeur *ERs1* contrôlant trois souches du phylotype I (CMR134 (I-13), PSS366 (I-15) et GMI1000 (I-18)), mais contourné par une souche virulente de ce même phylotype (PSS4 (I-15)). Par la suite, Salgon a réalisé une cartographie génétique en se basant sur des données de phénotypage de la résistance de la population d'aubergine MM738 (sensible) × AG9125 (résistante) vis-à-vis de huit souches du ceRs (CMR134 (I-13), PSS366 (I-15), GMI1000 (I-18), PSS4 (I-15), TO10 (I-47), CFBP 2957 (IIA-36), CFBP3059 (III-23), CMR34 (IIB-1)). Cette analyse lui a permis d'identifier le QTL³, nommé EBWR9⁴, qui permet de contrôler trois souches du phylotype I CMR134 (I-13), PSS366 (I-15) et GMI1000 (I-18). En effet, le locus de EBWR9 correspond au locus du gène majeur de résistance *ERs1* identifié par Lebeau et il est localisé à l'extrémité du chromosome 9, une région riche en gènes de résistances envers divers agents

³ Quantitative Trait Loci

⁴ Eggplant Bacterial Wilt Resistance located in chromosome 9

phytopathogènes. La découverte de ce QTL majeur de résistance est importante pour les sélectionneurs, car il est facilement introgressé dans les cultivars commerciaux (travaux en cours à la Réunion). Cependant, afin de contrôler une large gamme de souches du ceRs, la recherche d'autres sources de résistance demeure primordiale, car les tests réalisés par Salgon ont montré qu'EBWR9 n'est pas efficace pour contrôler au moins deux autres souches du phylotype I PSS4 (I-15), TO10 (I-47) et il ne permet pas de contrôler les souches d'autres phylotypes CFBP 2957 (IIA-36), CFBP3059 (III-23), et CMR34 (IIB-1). Cet exemple montre la complexité du développement de variétés résistantes au ceRs étant donné sa grande diversité, rendant la résistance souche spécifique. Ainsi, le choix du déploiement d'une variété résistante dans une zone géographique donnée doit se faire en ayant une parfaite connaissance de la diversité génétique et du potentiel évolutif des souches présentes dans la zone [17]. Dans la littérature, divers moyens de lutte sont également en cours de recherche comme la lutte biologique par l'utilisation de bactéries antagonistes afin d'inhiber la croissance de l'agent pathogène [197–200]; et le biocontrôle du ceRs par des bactériophages à cycle lytique qui sont des virus infectant et tuant la cellule bactérienne [201].

6. Distribution de la diversité génétique du ceRs

6.1. À l'échelle mondiale

Le phylotype II présente la plus large distribution mondiale, car il est présent sur les cinq continents. Le phylotype I est lui aussi mondialement distribué, car il a été détecté récemment en Europe [202] ; le phylotype III a une distribution plus restreinte car il n'est rencontré qu'en Afrique, à La Réunion et à Madagascar ; et le phylotype IV n'est présent qu'au Japon, Indonésie, Australie et récemment identifié en Inde [160] et à Maurice [161] (Figure 12).

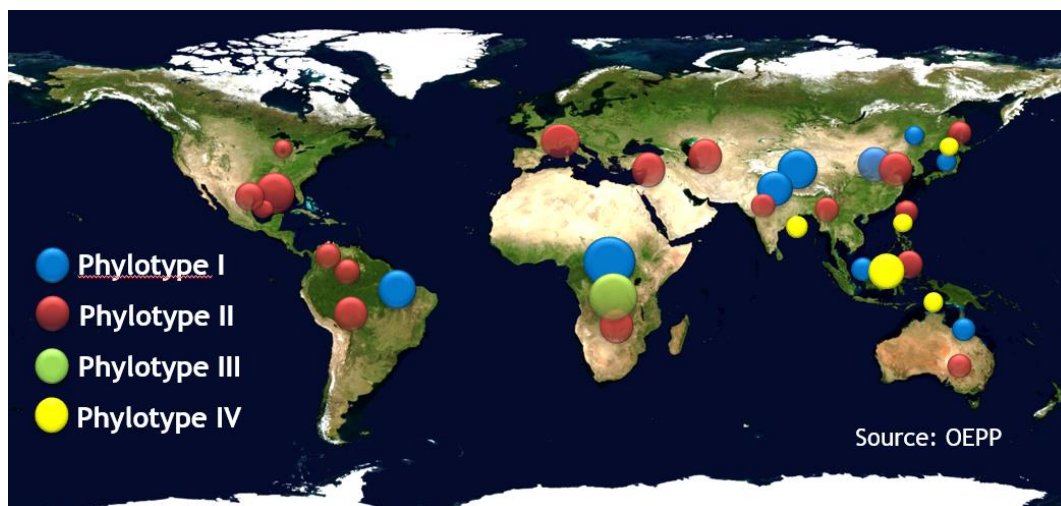


Figure 12: Distribution mondiale des quatre phylotypes du ceRs

6.2. Dans les petites îles du Sud-Ouest de l'Océan Indien

Dans les petites îles du SOOI (La Réunion, Comores, Mayotte, Maurice, Rodrigues, Seychelles), le flétrissement bactérien représente un des principaux facteurs limitant pour la culture des Solanacées (pomme de terre, tomate, piment, poivron, aubergine, etc.) [203]. Dans l'optique de mettre en place un réseau d'épidémiosurveillance dans la zone, une étude a été menée sur la diversité génétique du ceRs durant le projet Ralstotracing (2015-2018). Cette étude a montré que les quatre phylotypes du ceRs y sont présents avec une forte prévalence du phylotype I (représentant 87% des 1704 souches collectées) [162]. De plus, l'analyse des séquences *egl* a montré la prédominance du sequevar I-31 qui est présent dans toutes les îles que ce soit dans les basses ou moyennes altitudes [203]. Un schéma MLVA comprenant 17 marqueurs VNTRs a été appliqué sur les souches du phylotype I et a montré des niveaux de diversité génétique très contrastés selon les îles (Maurice présentant la plus forte diversité et les Comores la plus faible). Il a également révélé une forte prévalence d'un haplotype (MT035) présent dans toutes les îles [203].

6.3. À Madagascar

Depuis l'année 2009, de graves épidémies de flétrissement bactérien ont été signalées dans les principaux bassins de production de pomme de terre des hautes terres centrales malgaches [12]. Malgré la distribution de variétés tolérantes à la maladie par le Centre National de Développement Rural et de Recherche Appliquée (FIFAMANOR), la maladie

progresses car l'agent pathogène a réussi à contourner leurs résistances ne permettant plus aux paysans d'accéder à des tubercules sains. Ces épidémies laissent les producteurs dans le désarroi car le flétrissement bactérien est responsable d'une perte d'environ 30% de leurs productions [12]. Afin de mieux comprendre et maîtriser la progression du flétrissement bactérien, une étude épidémiologique a été menée par Ravelomanantsoa lors de sa thèse de doctorat (2013 – 2016) [11]. Elle a montré que les trois phylotypes (I, II et III) y sont présents avec une forte prévalence du phylotype IIB-1 (représentant 72% des 1224 souches collectées). Le phylotype IIB-1, probablement introduit *via* des tubercules infectés, présente une faible diversité génétique et est à l'origine des épidémies récentes de flétrissement bactérien. Le phylotype III présente une forte diversité génétique, suggérant son caractère endémique et l'absence de transmission par tubercules [12] (Figure 13). Le phylotype I quant à lui n'a pas pu être analysé lors de ces travaux dû à l'absence de schéma de génotypage spécifique à ces souches.

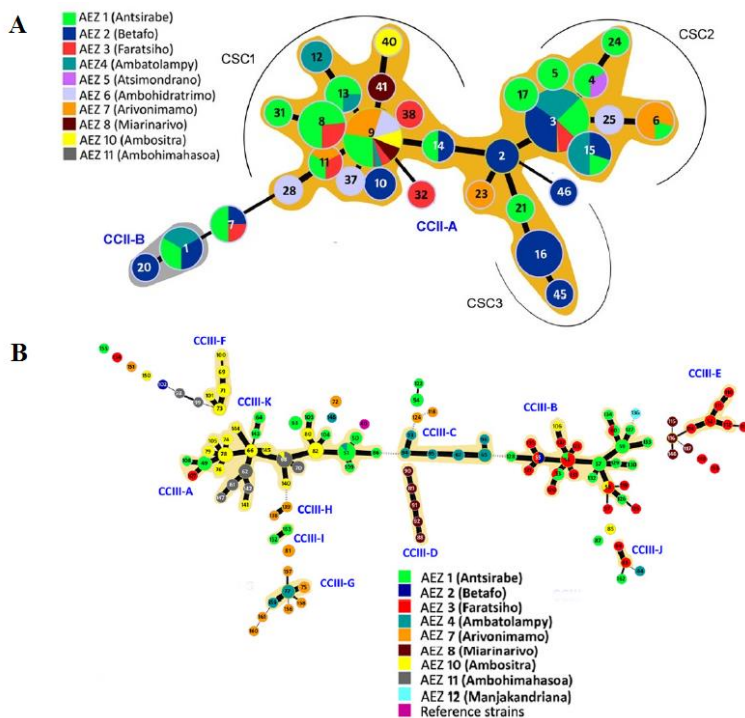


Figure 13: Minimum Spanning Trees (MSTs)

A) Minimum Spanning Tree représentant les souches du phylotype IIB-1 collectées dans les Hautes Terres Centrales malgaches. B) Minimum Spanning Tree (MST) représentant les souches du phylotype III collectées dans les Hautes Terres Centrales malgaches.

7. Etudes antérieures sur l'activité bactériocine du ceRs

C'est en 1954 qu'Okabe a reporté pour la première fois l'activité bactériocine du ceRs [204]. Depuis, quelques études se sont intéressées à leur caractérisation, dont celles de Cuppels et al. [205] et He et al. [206] qui ont identifiés des bactériocines non sédimentables et de faibles masses moléculaires ; ainsi que l'étude de Arwiyanto et al. [96], qui ont identifié des bactériocines sédimentables, résistantes à la trypsine (protéase), thermolabiles et de hautes masses moléculaires. En 1996, Frey et al. [207] ont effectué le typage bactériocine de souches provenant des Antilles françaises : 24 souches de référence et 65 souches issues de parcelles d'expérimentations. Ils ont mis en évidence que les bactériocines étaient thermolabiles et sensibles à la protéinase K, donc qu'elles étaient de nature protéinique. Ce résultat rejoint la caractérisation faite par Arwiyanto et al. [96]. De plus, il a été montré que le groupe génétique d'affiliation des souches est parfaitement corrélé au groupe généré par le typage bactériocine [207]. En outre, l'activité bactériocine des souches du ceRs a également été démontrée *in planta*, ce qui pourrait conférer un avantage sélectif pour la croissance des souches productrices. En 2015, les résultats obtenus par Huerta et al. [97] ont suggéré l'influence de la production de bactériocines sur la distribution géographique de souches de phylotype IIB-1, représentée par la souche UW551, qui est adaptée aux températures fraîches. En effet, ils ont démontré que la souche GMI1000 (sequevar I-18) adaptée à une température tropicale et la souche K60 (sequevar IIA-7) adaptée à une température subtropicale produisent une bactériocine qui inhibe la croissance de la souche UW551 en température chaude tropicale et subtropicale.

Par ailleurs, une étude préliminaire sur l'activité et le typage bactériocine des souches du ceRs provenant de Madagascar (phylotypes IIB-1 et III) et des petites îles du SOOI (phylotypes I, II, III, IV) a été menée par Fabre en 2016 afin d'expliquer l'émergence du sequevar I-31 [208]. Il en est ressorti une relation entre la variabilité phénotypique (profil bactériocine) et la variabilité génétique des souches (haplotype MLST). De plus, un résultat marquant de cette étude est le contrôle *in vitro* par les souches du sequevar I-31 (prévalent dans les petites îles du SOOI) d'un nombre important de souches. Ces résultats confortent l'hypothèse que les bactériocines pourraient donner un avantage compétitif aux souches productrices et qu'elles pourraient contribuer à la prévalence de certaines souches bactériennes par rapport à d'autres.

IV. Objectifs et questions de recherche

Cette thèse a comme principaux objectifs d'analyser la diversité et la structure génétique des populations du ceRs à Madagascar et d'étudier les caractéristiques biologiques des populations du ceRs à Madagascar et dans les petites îles du SOOI.

Pour cela, nous allons répondre à différentes questions de recherche. D'abord, celles concernant l'épidémiologie moléculaire du ceRs à Madagascar:

- Quel est le niveau de diversité et la structure génétique des populations du ceRs à Madagascar?
- Quel est le profil épidémiologique des populations du ceRs malgache ?
- Quels sont les liens génétiques entre les souches malgaches du ceRs et les souches provenant des petites îles du SOOI et du reste du monde ?

Ensuite, les questions de recherche concernent les caractéristiques biologiques des populations permettant une meilleure compréhension de la situation épidémiologique observée à Madagascar et dans les petites îles du SOOI:

- Quelle est l'activité bactériocine des souches du ceRs provenant de Madagascar, des petites îles du SOOI et des souches représentatives de la diversité mondiale du ceRs ?
- Quelles sont les différences au niveau génomique du répertoire de gènes codant potentiellement pour les bactériocines entre les différentes lignées génétiques du ceRs ?

Chapitre 1

Chapitre 1: Epidémiologie moléculaire du complexe d'espèces *Ralstonia solanacearum* à Madagascar

Le flétrissement bactérien (FB) causé par les bactéries du complexe d'espèces *Ralstonia solanacearum* ou ceRs constitue un fléau majeur de la culture maraîchère à Madagascar. Afin de proposer des moyens de lutte efficace contre cette maladie et pour éviter sa propagation, la connaissance de la diversité et structure génétique des populations du ceRs qui sévissent dans les parcelles de cultures s'avèrent nécessaire. En 2006, 2012 et 2013, des missions de prospections du FB ont été effectuées dans les hautes terres centrales malgaches et au niveau d'une région localisée en basse altitude. Ces prospections ont permis de collecter 1249 souches dont 12% appartiennent au phylotype I, 70% au phylotype II et 18% au phylotype III. Grâce aux schémas de génotypage RS2-MLVA9 et RS3-MLVA16, une étude d'épidémiologie moléculaire a pu être menée respectivement pour les souches du phylotype II et III [12]. Malheureusement, l'absence d'un outil de génotypage spécifique aux souches du phylotype I, n'a pas permis leurs analyses. Ainsi, pour décrire la situation épidémiologique du phylotype I, un schéma de génotypage MLVA a été développé. Au total 14 marqueurs VNTR ont été retenus, car ils se sont avérés suffisamment polymorphes, amplifiables et reproductibles pour un grand nombre de souches présentes dans la collection du Pôle de Protection des Plantes (Saint-Pierre, La Réunion). Ces VNTR sont localisés à la fois sur le chromosome et le mégaplasmide de la bactérie. Les résultats du développement de cet outil de génotypage a été valorisé dans l'**article 1**, publié dans Plos One. Cette étude démontre la pertinence du schéma de génotypage RS1-MLVA14 pour décrire la situation épidémiologique du phylotype I malgache collecté dans deux régions de la Haute Terre Centrale et d'une région localisée en basses altitudes. Or le FB sévit également dans d'autres régions de Madagascar situées en basses et moyennes altitudes. Ainsi, pour une étude plus exhaustive, de vastes campagnes de collecte d'échantillons de tige flétrie par le ceRs ont été menées en 2018, 2019 et 2021 (**article 2**). Les souches du phylotype I sont largement disséminées dans les 3 provinces représentant 8 régions de Madagascar. L'assignation phylogénétique de ces isolats a permis d'identifier quatre lignées génétiques

(I-18, I-31, I-33 et I-46), avec une forte prévalence du sequevar I-18. Par la suite, le schéma de génotypage RS1-MLVA14 a été appliqué aux isolats de phylotype I pour les différencier plus finement et leur diversité et structure génétique ont été décrites. Au total, 97 haplotypes ont été recensés et les souches collectées en basses altitudes (provinces de Toamasina et Mahajanga) présentent une plus grande diversité génétique que les souches collectées en hautes altitudes (province d'Antananarivo). De plus, la présence de certains haplotypes dans diverses parcelles parfois localisées dans des provinces différentes suggère une dissémination via des matériels végétaux infectés. En outre, pour expliquer la forte prévalence du I-18 à Madagascar, nous avons exploré l'influence des bactériocines qui sont des substances antimicrobiennes. Une corrélation existe entre l'activité bactériocine et la situation épidémiologique observée. Cependant, ce mécanisme biologique seul ne pourrait expliquer cette épidémiologie et d'autres pistes méritent d'être explorées.

Article 1

Contrasting genetic diversity and structure among Malagasy *Ralstonia pseudosolanacearum* phylotype I populations inferred from an optimized Multilocus Variable Number of Tandem Repeat Analysis scheme

Hasina Rasoamanana¹, Santatra Ravelomanantsoa², Noura Yahiaoui¹, Niry Dianzinga¹, Emeline Rébert¹, Miharisoa-Mirana Gauche¹, Yann Pecrix³, Laurent Costet³, Adrien Rieux³, Philippe Prior⁴, Isabelle Robène³, Gilles Cellier⁵, Fabien Guérin¹, Stéphane Poussier^{1*}

1 Université de La Réunion, UMR Peuplements Végétaux et Bioagresseurs en Milieu Tropical, F-97410 Saint-Pierre, Réunion, France, **2** Centre National de la Recherche Appliquée au Développement Rural FOFIFA, F-101 Antananarivo, Madagascar, **3** Centre de Coopération Internationale en Recherche Agronomique pour le Développement, UMR Peuplements Végétaux et Bioagresseurs en Milieu Tropical, F-97410 Saint-Pierre, Réunion, France, **4** Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, UMR Peuplements Végétaux et Bioagresseurs en Milieu Tropical, F-97410 Saint-Pierre, Réunion, France, **5** Anses - Plant Health Laboratory - Tropical Pests and Diseases unit, F-97410 Saint-Pierre, Réunion, France.

*: stephane.poussier@univ-reunion.fr

Abstract

The *Ralstonia solanacearum* species complex (RSSC) composed of three species and four phylotypes, are globally distributed soil-borne bacteria with a very broad host range. In 2009, a devastating potato bacterial wilt outbreak was declared in the central highlands of Madagascar, which reduced the production of vegetable crops including potato, eggplant, tomato and pepper. A molecular epidemiology study of Malagasy RSSC strains carried out between 2013 and 2017 identified *R. pseudosolanacearum* (phylotypes I and III) and *R. solanacearum* (phylotype II). A previously published population biology analysis of phylotypes II and III using two MultiLocus Variable Number of Tandem Repeats Analysis (MLVA) schemes revealed an emergent epidemic phylotype II (sequevar 1) group and

endemic phylotype III isolates. We developed an optimized MLVA scheme (RS1-MLVA14) to characterize phylotype I strains in Madagascar to understand their genetic diversity and structure. The collection included isolates from 16 fields of different Solanaceae species sampled in Analamanga and Itasy regions (highlands) in 2013 (123 strains) and in Atsinanana region (lowlands) in 2006 (25 strains). Thirty-one haplotypes were identified, two of them being particularly prevalent: MT007 (30.14%) and MT004 (16.44%) (sequence 18). Genetic diversity analysis revealed a significant contrasting level of diversity according to elevation and sampling region. More diverse at low altitude than at high altitude, the Malagasy phylotype I isolates were structured in two clusters, probably resulting from different historical introductions. Interestingly, the most prevalent Malagasy phylotype I isolates were genetically distant from regional and worldwide isolates. In this work, we demonstrated that the RS1-MLVA14 scheme can resolve differences from regional to field scales and is thus suited for deciphering the epidemiology of phylotype I populations.

Introduction

Knowledge of a pathogen's population biology is a key aspect of disease management [1,2]. Assessing the genetic diversity and population structure unravels the genetic links between isolates and is useful for developing preventive control strategies [3]. Multiple genotyping techniques have been developed to identify pathogenic isolates, determine phylogenetic relationships between isolates and reveal the genetic links between populations, e.g. for the plant pathogens *Erwinia amylovora* [4] and *Phytophthora ramorum* [5].

MultiLocus Variable Number of Tandem Repeat (VNTR) Analysis (MLVA) is a PCR-based genotyping method for amplifying VNTR loci. VNTR are tandemly repeated DNA sequences, whose copy numbers may vary among isolates. This variation can be due to two mutational mechanisms during DNA replication: i) the unequal crossing over, frequent among minisatellites with repeat units longer than 9 bp; and ii) the sliding of a strand of DNA followed by a lack of repair ("Slipped Strand Mispairing" or SSM), common for microsatellites with repeat units between 2 and 8 bp [6]. MLVA has the advantage of being highly discriminatory, robust, reproducible and rapid [6,7]. It has been widely used in medical microbiology [8] to trace the origin of epidemics or the spread of virulent or antibiotic-resistant variants. Later developed for plant pathogenic bacteria, MLVA schemes have proven to be useful for outbreak analysis, epidemiological surveillance and the study of the genetic

diversity and population structure of: *Clavibacter michiganensis* subsp. *michiganensis* [9], *Pseudomonas syringae* pv. *tomato* and *maculicola* [10], *Xylella fastidiosa* [11], *Xanthomonas citri* pv. *citri* [12,13], *Xanthomonas oryzae* [14], *Xanthomonas arboricola* pv. *pruni* [3], *Xanthomonas phaseoli* pv. *manihotis* [15], *Xanthomonas vasicola* pv. *musacearum* [16], *Xanthomonas citri* pv. *viticola* [17], and the *Ralstonia solanacearum* species complex [18–22].

The soil-borne and xylem-limited *Ralstonia solanacearum* species complex (RSSC) causes bacterial wilt (BW) disease on an unusually broad host range, which includes over 50 botanical families [23,24]. RSSC is one the most devastating plant pathogenic pathogens in terms of the damage it causes to crops [25], such as potato, tomato, tobacco, ginger and banana, as well as leguminous and ornamental crops [24]. BW causes economic losses estimated at US \$ 1 billion/year for the potato industry [23]. The RSSC is classified into four major phylogenetic groups (named phylotypes) according to their geographical origin [26] and comprises three species [27,28]: *R. pseudosolanacearum* includes phylotypes I and III originating from Asia and Africa, respectively; *R. solanacearum* includes phylotype II with IIA and IIB subdivisions from the Americas; and *R. syzygii* includes the phylotype IV from Indonesia, Japan and Australia. Each phylotype is subdivided into unique sequevars based on the partial endoglucanase gene (*egl*) sequencing [29].

In Madagascar, BW disease was first reported in 1934 by Bouriquet on tobacco [30] and artichoke [31]. Later, BW was described on potato, eggplant, groundnut, tomato [32], bean and cabbage [33]. BW was well established in the central highlands, as well as in western and eastern lowlands, where a few RSSC isolates were collected and identified as belonging to phylotypes I and III [34] (Philippe Prior, 2006, unpublished work,). In 2009, major outbreaks were reported in the main potato production areas of the central highlands [19]: BW, which usually occurs on rain-fed potato crops during the rainy season, appeared in the winter season on irrigated potato crops previously unaffected by the disease [19,35]. In addition, BW symptoms appeared very early during vegetative growth and the BW-tolerant cultivars, developed by the National Centre for Rural Development and Applied Research (FIFAMANOR) and distributed to farmers, became susceptible to RSSC [19].

In 2013, a wide sampling campaign was conducted in Madagascar's central highlands. A total of 1224 isolates were collected: 10% belonged to phylotype I, 72% to phylotype IIB-1 and

18% to phylotype III [19] . Only the population genetics of phylotypes IIB-1 and III were analysed, using the MLVA schemes RS2-MLVA9 for phylotype II isolates and RS3-MLVA16 for phylotype III isolates [19], derived from previously published studies [20,21]. This study revealed contrasted population structures between phylotype IIB-1 and phylotype III. Findings suggest that phylotype IIB-1 isolates, reported in Madagascar for the first time, were introduced and spread massively via latently infected potato seed tubers, whereas Malagasy phylotype III isolates appeared to be endemic [19].

The present study aims to develop and apply an optimized MLVA scheme specifically designed for phylotype I isolates in order to: (i) assess the level of genetic diversity and structure of phylotype I populations isolated in Madagascar's central highlands and lowlands, (ii) compare the population genetic structures of phylotype I with Malagasy phylotypes IIB-1 and III isolates and (iii) analyse the potential genetic links of the Malagasy phylotype I isolates with isolates from other countries.

Materials and Methods

1. Bacterial isolates and phylogenetic assignment

Two collections of Malagasy RSSC isolates belonging to phylotype I and maintained at the Plant Protection Platform (Saint-Pierre, Réunion, France) were used in this study. The first collection included 123 isolates from Analamanga (n = 28) and Itasy (n = 95), regions in the central highlands (mean elevation = 1057.23 m). The second collection included 25 isolates from the Atsinanana region in the lowlands (mean elevation = 28.31 m) (S1 Table). The regional difference in elevation provided the opportunity to study the influence of altitude on the genetic diversity and structure of phylotype I populations. In addition, the Analamanga and Itasy regions are in the main vegetable producing areas in Madagascar, where a devastating BW outbreak occurred in 2013. In the Atsinanana region, several cases of BW were also reported. Madagascar's main seaport is in this region (Toamasina), which allowed the study of the potential influence of frequent commercial exchanges of plant material on the genetic diversity of phylotype I isolates.

The 148 Malagasy isolates were collected from stem fragments of seven plant species (*Solanum lycopersicum*, *S. tuberosum*, *S. melongena*, *S. aethiopicum*, *S. nigrum*, *S. scabrum*,

Capsicum annuum) collected from 16 fields (Fig 1) in the Atsinanana region (4 plots), the Analamanga region (4 plots) and the Itasy region (8 plots). All the Malagasy isolates were checked by multiplex PCR [29] in order to confirm their assignment to RSSC and phylotype I.

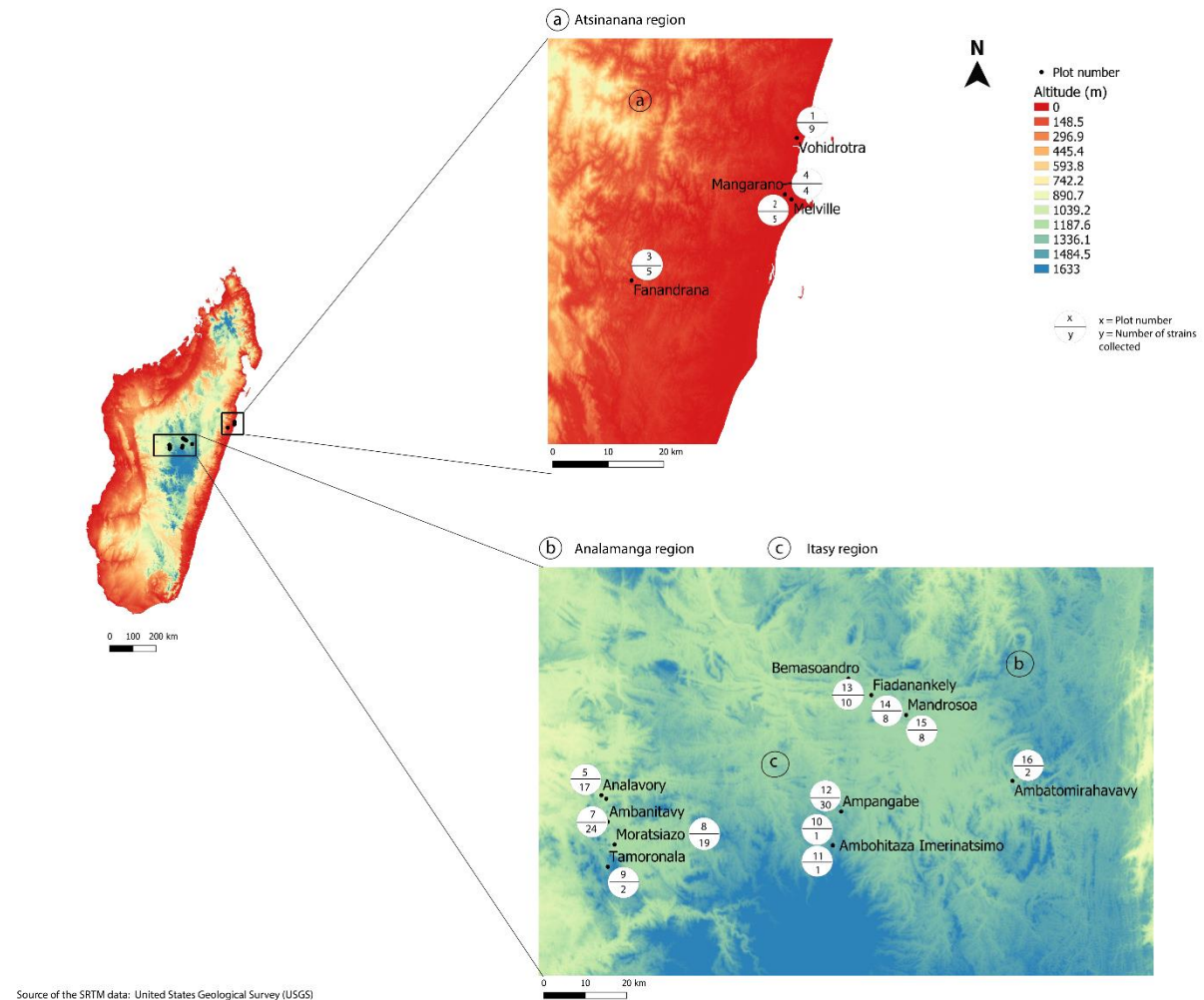


Fig 1. Collection area of RSSC phylotype I in Madagascar

The 16 sites where phylotype I isolates were collected in the three Malagasy regions (a: Atsinanana; b: Analamanga; and c: Itasy) are indicated by black dots. Each white circle contains information about the plot number and the number of isolates collected in each site.

In order to identify the potential genetic relationships between Malagasy and international RSSC phylotype I isolates, 145 isolates from a representative worldwide collection (South-West Indian Ocean, Africa, Asia, Americas) kept at the Plant Protection Platform (Saint-Pierre, Réunion, France) were also included in this study (S2 Table).

All RSSC isolates were stored on Cryobank® microbeads (Microbank®, PRO-LAB DIAGNOSTICS, Neston, Wirral, U.K.) at -80 °C. After growth on nutrient broth and Kelman media [36], 1 µl loop of a bacterial cell suspension was streaked onto agar plates to isolate colonies. Single colonies were suspended in 400 µl of sterile HPLC-grade water and used as templates for PCR amplification.

The phylogenetic assignment of Malagasy phylotype I isolates was based on the multiplex phylotype PCR [29] and the partial nucleotide sequences of the endoglucanase (*egl*) gene [29]. PCR amplification, sequencing and sequevar determination were performed as previously described [37]. The newly generated sequences were deposited in the GenBank database under accession numbers MN310771 to MN310886 and MW014321 to MW014325.

2. VNTR identification and selection for MLVA

In a first step, we worked with VNTR loci that had previously been identified throughout the genomes of RSSC isolates belonging to different phylotypes [19,20,22,38]. By using the Muscle algorithm in Geneious v10.0.9 [39], these VNTR, flanked by their primers, were searched and aligned with the reference genome of phylotype I (GMI1000) [40] and nine draft genome sequences of phylotype I isolates: CFBP7058 (sequevar 13), CIR011-208 (sequevar 17), RD13.01 (sequevar 31), TD13.01 (sequevar 31), CIV23 (sequevar 31), TF31.08 (sequevar 31), PSS04 (sequevar 15), TO10 (sequevar 47) and Y45 (sequevar 44) [41,42]. The purpose of the alignment was to check the presence of the VNTRs in the different genomes of phylotype I and confirm that their characteristics met the following requirements [38]: absence of *indel* in the VNTR, unicity of the VNTR sequence in the genome, non-overlapping VNTR, a very high typeability rate (100 % or almost 100%) for phylotype I isolates, 80% internal conservation of the VNTR in the genomes [6], and a repeat unit between 5 and 10 nucleotides [43].

In a second step, we screened the reference GMI1000 genome with the Tandem Repeat Finder v4.09 (<https://tandem.bu.edu/trf/trf.html>) [44] to detect new VNTR loci. New VNTRs were validated by using the Phobos Tandem Repeat Finder [45]. Then, an alignment with the nine draft genome sequences was conducted to localize the VNTR and confirm whether their characteristics met the requirements described above.

Lastly, to check the distribution and specificity of the selected VNTR markers among RSSC isolates, we screened with BLASTN [46] 114 publicly available RSSC genomes: 48 belonging to phylotype I, 50 to phylotype II, 3 to phylotype III and 13 to phylotype IV (<https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/490/>).

3. Primer design and assay optimization

PCR primers were designed from the flanking regions of VNTR loci using Primer3 v4.1.0 [47]. Selection was based on the following criteria: primer size from 18 to 23 bp, annealing temperature between 60°C and 65°C, percentage of guanine and cytosine from 50% to 70%, and total length of PCR product between 100 bp and 500 bp. The probability of dimerization and hairpin were evaluated using the OligoAnalyzer v3.1 software package (<https://eu.idtdna.com/calc/analyzer>). Oligonucleotides were synthesized by Macrogen, Inc. (Seoul, South Korea).

To perform a preliminary screening, simplex PCR were conducted to assess the amplifiability, reproducibility and polymorphism of the selected VNTR loci. We used eight phylotype I isolates (RUN0054, RUN0157, RUN0215, RUN0334, RUN0969, RUN1744, RUN1985 and RUN3014) representing several sequevars (13, 14, 15, 17, 18, 31 and 47) from a worldwide collection (French Guiana, Réunion, China, Thailand, Cameroon, Taiwan, Côte d'Ivoire). PCR amplifications were performed in 15 µl reaction volumes containing 7.5 µl Terra PCR Direct Buffer 2X (Terra™ PCR Direct Polymerase Kit, Clontech Laboratories, Inc.), 0.3 µl Terra PCR Direct Polymerase Mix - 1.25 U/µl, 1.5 µl 5X Q-solution (Qiagen®, Hilden, Germany), 2.3 µl of a forward and reverse primer mix (2 µM each), 2.4 µl sterile HPLC-grade water, and 1 µl of bacterial suspension as a template. PCRs were carried out in a GeneAmp PCR System 9700 Thermal Cycler (Foster City, CA 94404, USA) under the following conditions: an initial denaturation step at 98°C for 2 min, 30 cycles of denaturation at 98°C for 10s, annealing at 62°C for 15s, extension at 68 °C for 1 min, and a final extension step at 68°C for 30 min. Then, 6 µL of PCR product was mixed with 1 ml of loading dye solution and loaded into a 1.5% (w/v) SeaKem® LE Agarose (Lonza, Basel, Switzerland) gel for electrophoresis. Ethidium bromide was used to stain the gels and the G:BOX gel imaging system (Syngene, Cambridge, UK) enabled the visualization of the bands under ultraviolet. The molecular weights were estimated by comparison with a 100 bp DNA ladder (Promega, Madison, Wisconsin, USA). VNTR loci showing poor amplification and/or lacking diversity were removed at this step.

4. MLVA genotyping

A Multiplex PCR protocol was applied to analyse from 3 to 4 VNTR loci per reaction using the Multiplex PCR kit (Qiagen, Courtaboeuf, France). The primers (Table 1) were pooled in four different mixes according to their annealing temperature and the size of the PCR product. Each forward primer was marked with one of the fluorophore: 6-FAM, NED, PET and VIC (Applied Biosystems, Life Technologies, Saint Aubin, France). The conditions of amplification for the multiplex PCR were the same as for the simplex PCR, except the number of cycles (25 instead of 30). PCR products were diluted (at least 1:80) to avoid peak saturation.

In a 2 ml Eppendorf tube, 1080 µl formamide (Hi-Di™ formamide, Applied Biosystems) was mixed with 20 µl size marker (GeneScan™-500 LIZ® Size Standard, Applied Biosystems). Then, in each well of a 96-well plate, 11 µl of the mix was distributed with 1 µl of diluted PCR product. The samples were denatured at 95°C for 5 min, cooled immediately on ice and loaded onto an ABI Prism 3130XL Genetic Analyzer for capillary electrophoresis. The reproducibility of the MLVA genotyping was checked by including 3 phylotype I control isolates in each 96-well plate: RUN0054 (GMI1000, sequevar 18), RUN0157 (PSS04, sequevar 15) and RUN3014 (TFB1.08, sequevar 31).

5. Data analysis

The genotyping results were analysed using Geneious v10.0.9. Fragment sizes were estimated using the third order least squares algorithm and attributed to a bin size, which takes into account small size variation due to experimental variation. To confirm the validity and repetition number for each VNTR locus, PCR products of isolates RUN0320, RUN3012, RUN3216 and RUN3277 were sequenced (Genewiz, Leipzig, Germany) and analysed using Geneious v10.0.9. These sequences were also used to check for patterns in the flanking VNTR sequences and internal repeat variations (i.e. copy homology). When a VNTR array was truncated, the VNTR number was rounded to the nearest bin whole number [48–50]. The combination of alleles (i.e. number of repetitions) for the 14 VNTR loci was considered as MLVA type (MT, i.e. haplotype).

In order to assess the resolution of the RS1-MLVA14 scheme, a genotype accumulation curve (GAC) was built, based on the genotyping of 291 phylotype I isolates (146 from Madagascar and 145 from worldwide). The GAC represents the power of discrimination of our set of loci

in the population [51,52]. The discriminatory power of the RS1-MLVA-14 scheme was also evaluated by calculating the Hunter Gaston Discrimination Index (HGDI) [53]. It was compared to the recently developed MLST-7 scheme [37], by using 94 worldwide isolates (S1 Table and S2 Table) representing 9 sequevars of phylotype I and selected to maximize the phylogenetic diversity, as well as the diversity of country, host and date of isolation.

Using Arlequin v3.5.2.2 [54], the genetic diversity of Malagasy isolates was evaluated by calculating several indices: Nei's unbiased estimates of genetic diversity (H_{nb}), number of haplotypes, number of alleles, number of private alleles. The allelic richness and private allelic richness were evaluated using HP-RARE [55]. The intra-population diversity indices (H_{nb} , number of haplotypes, mean number of alleles, number of private alleles, allelic richness and private allelic richness) provide information on the level of genetic diversity in a population, the population's epidemiological profile (epidemic or endemic), and the mode of dispersion of the inoculum (stochastic dispersion or dissemination by infected plant material, etc.).

For the robustness of analyses, we considered that a minimum number of samples was required (size $n \geq 14$) to form a population. A population is defined a priori by a group of individuals that could exchange genes. Different parameters, such as plant host, altitude and collection area were considered for the analysis of the genetic diversity indices and to define putative populations. The F_{ST} and R_{ST} index calculated with Arlequin (10 000 permutations) provided information on the similarity or differentiation between populations due to the presence or absence of gene flows, mainly due to migration.

We also determined a genetic structure of populations without a priori via the R package Bios2mds [56]. This package is used for conducting metric multidimensional scaling (MDS), a method that represents measurements of similarity (or dissimilarity) among pairs of objects as distances between points of low-dimensional or multidimensional space [57]. Furthermore, the genetic relationships among Malagasy isolates and between Malagasy and worldwide isolates were computed using Phyloviz (<http://www.phyloviz.net/>) [58]. Different minimum spanning trees (MSTs) were built with the goeBURST full MST algorithm using global optimal eBURST (goeBURST) and Euclidean distances. They revealed the level of differentiation between haplotypes: Single, Double or Triple Locus Variant (SLV, DLV, or TLV), and enabled the identification of clonal complexes (CC). We defined CC as groups of

SLV haplotypes in which the founder(s) is (are) defined as the haplotype(s) comprising the largest number of SLV.

Results

1. Development of the RS1-MLVA14 scheme

We developed an optimized MLVA scheme capable of characterizing phylotype I isolates with greater reliability and discriminatory power. Consequently, we were able to overcome certain issues, namely the absence of polymorphism, a low typeability rate and the presence of different VNTRs in some loci [20,59]. These problems were highlighted by our in-laboratory preliminary tests with Malagasy isolates. Based on previously identified VNTR markers [20,22,38] and our present genome screening, we selected a final set of 14 VNTR loci to generate the RS1-MLVA14 scheme (Table 1). VNTR markers were distributed on the chromosome ($n = 4$) or the megaplasmid ($n = 10$) (S1 Fig), 5 being intergenic and 9 being intragenic. VNTR sizes varied from 5 to 9 nucleotides and were repeated in the genomes from 1 to 25 times. In our collection of 293 phylotype I isolates, all selected VNTR loci were polymorphic. The number of alleles per locus ranged from 2 to 16 alleles and the allelic richness varied from 1.179 to 5.027. The VNTR marker GMIch_1844 showed the highest genetic diversity ($H_{nb} = 0.738$), whereas the VNTR marker GMImp_1618 had the lowest ($H_{nb} = 0.027$). Our screening of 114 publicly available RSSC genomes (S3 Table) revealed that the 14 selected VNTR loci were found in all the 48 phylotype I isolates (except in some incomplete genomes, where 3% of loci could not be found, as already observed [20]). In addition, 10 VNTR loci were found in the 3 phylotype III isolates and 3 to 8 VNTR loci were found in the 50 phylotype II and 13 phylotype IV isolates. Four VNTR loci (GMIch_0581, GMIch_3461, GMIch_1844, GMImp_0266) appeared to be specific to phylotype I.

Table 1: Features of the 14 selected VNTR loci composing the RS1-MLVA14 scheme

Mix	VNTR marker, Locus name ^a	Localization	Code	VNTR sequence	Primers (5' - 3')		Range of repeat numbers ^{bc}	Number of alleles ^b	H _{nb} ^b	Allelic richness ^b	Reference
Mix A	GMlch_0581	Intergenic	VNTR01	TGGTT	F:GGGGCGTTGGTGTTTGGCTG	VIC	6 - 25	14	0.727	3.799	This study
					R:AACACAGGATGCACCACCAG						
	GMlch_2433	Intragenic	VNTR02	CAAGCACCT	F:AGTTCACTGTCGATCCAA	FAM	3 - 10	8	0.654	4.039	[20]
					R:ACATCCATGTCCGCACGC						
	GMlch_3461	Intragenic	VNTR03	AATGGTTG	F:CGAGGTCGCTCTCCAGAAGGCGA	NED	2 - 9	6	0.579	2.896	[22]
					R:CGAGAAGGCCAGTCCCGAGCTGA						
	GMlch_1457	intragenic	VNTR04	CGACCGACT	F:CGACCGTCGCCGACGTCC	PET	1 - 15	13	0.579	3.405	This study
					R:GAGCACGTTGGCGATGAA						
Mix B	GMImp_0487	Intragenic	VNTR06	CACCACGAG	F:CGCAGACGTTACAGAGACTT	FAM	3 - 5	3	0.359	3.411	This study
					R:CGGACAGCTGTGCTGATAC						
	GMlch_1844	Intergenic	VNTR07	CGGCATAC	F:CGGCTGAGACCGGATTGC	NED	1 - 10	10	0.738	4.805	[20]
					R:CTCATGCGAAAGTGTTTGTAAGTGTGT						
	GMImp_1343	Intragenic	VNTR08	GCCCAATCG	F:CACGACATCCACCGAAG	PET	0 ^d - 6	7	0.636	3.403	This study

Mix	VNTR marker, Locus name ^a	Localization	Code	VNTR sequence	Primers (5' - 3')		Range of repeat numbers ^{bc}	Number of alleles ^b	H _{nb} ^b	Allelic richness ^b	Reference
					R:CGGCTCGTTCATCAAGCA						
Mix C	GMImp_0449	Intragenic	VNTR09	CACCACCAG	F:GTAACCCACCCCAACGCTTA	VIC	5 - 6	2	0.496	2.000	This study
					R:TTGAGCGGATCCTGGTTCTT						
	GMImp_1828	Intergenic	VNTR10	TGCGGA	F:CTCGGGCCGTTTCATCGAC	FAM	3 - 20	16	0.689	5.013	[38]
					R:GATCGGGCAGCCGGATAC						
	GMImp_0740	Intergenic	VNTR12	GACCGCCAC	F:GCTGCATCACTTCGGCATT	PET	2 - 7	6	0.690	4.227	This study
					R:GAAGCGCAGGGAGATGGTAT						
GMImp_0618	Intragenic	VNTR13	GCCGCGACA	F:TGCTGTACCGCTGGCTTA	NED	2 - 3	2	0.495	2.000	This study	
				R:CCAGCCGAGCAGATCAT							
Mix D	GMImp_0266	Intragenic	VNTR17	CCGGGCAG	F:GGATCGACGTACACGCCTTT	FAM	1 - 6	6	0.568	2.906	This study
					R:GGACGATGTCGACGTTACGA						
	GMlch_0827	Intergenic	VNTR18	GGCGATGAG	F:TATGTCTCCTGTGCAAGTCGGTG	NED	3 - 6	4	0.520	2.310	[20]
					R:CCCAAGACCATCTCGGAAAG						
GMImp_1618	Intragenic	VNTR19	GGCATCGGC	F:GAATGCCGACGGAAAACTC	PET	1 - 2	2	0.027	1.178	This study	
				R: TGTCGTGGCACTGGATCTTC							

^aLoci were named according to the genome's name (GMI= GMI1000), (ch=chromosome, mp=megaplasmid), and physical position (kb).

^bThe VNTR number range, the number of alleles per locus, the Nei's unbiased diversity index (H_{nb}), and the allelic richness were calculated from the genotyping data derived from the 293 isolates used in this study.

^cThe number of repeats of the VNTR loci was calculated using the following formula:

$$\text{Number of repeats} = \frac{\text{Bin size} - \text{Forward and reverse primer size} - \text{Flanking region size}}{\text{VNTR size}}$$

^dThe VNTR was absent in 9 worldwide phylotype I isolates: 4 from Cameroon (RUN0119, RUN0194, RUN0205, RUN0215), 2 from Côte d'Ivoire (RUN1798, RUN1869), 1 from Taiwan (RUN0257), 1 from China (RUN0617), 1 from French Guiana (RUN1972).

2. Genotypic resolution of the RS1-MLVA14 scheme

The genotypic resolution of our MLVA scheme was assessed by a genotype accumulation curve (S2 Fig), which revealed that our set of loci is sufficient to discriminate between haplotypes in our collection, given that nearly 100% of the genotypes could be detected with 13 markers.

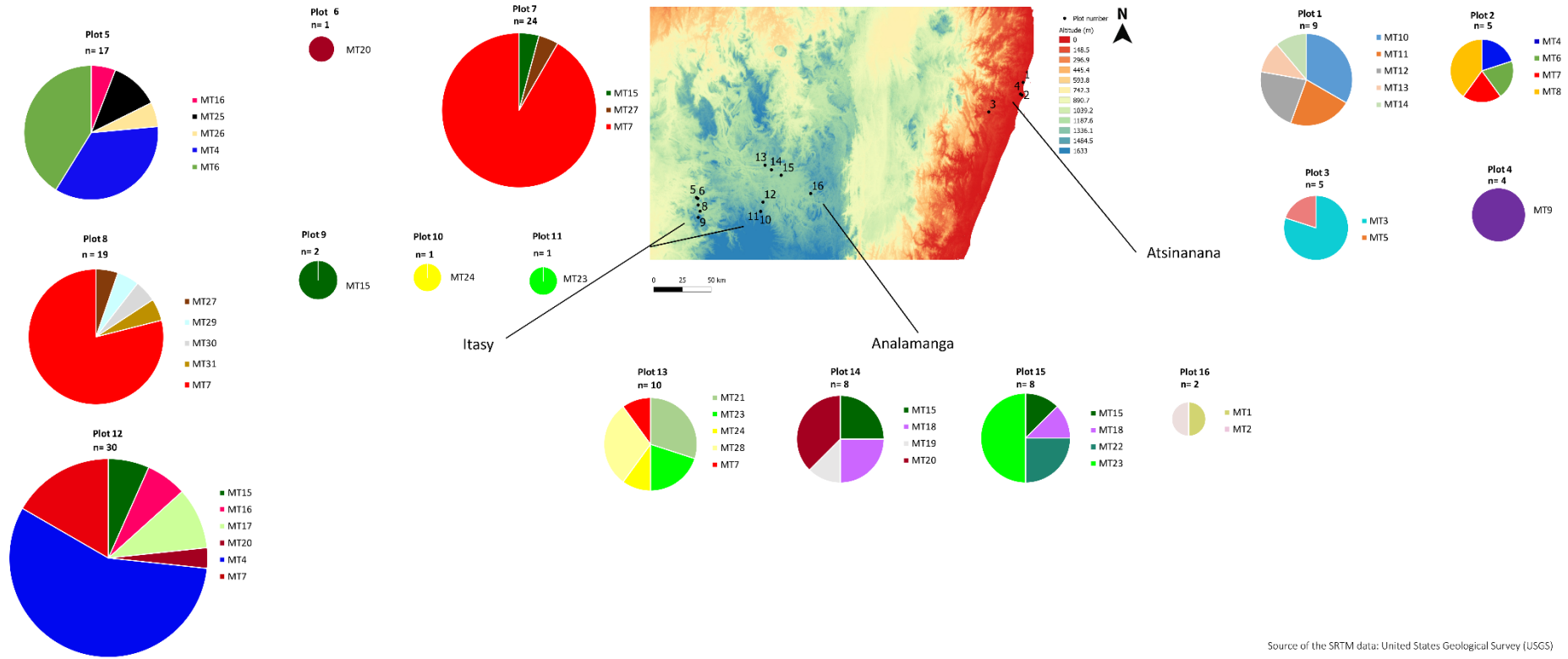
The comparative study of the discriminatory power of RS1-MLVA14 and MLST-7 typing schemes using 94 isolates representing 9 sequevars of phylotype I (S4 Table) showed that RS1-MLVA14 reveals twice as many haplotypes as MLST-7: 27 MLVA Type (MT) instead of 14 Sequence-Type (ST). The discriminatory power is 1.38 times higher with the RS1-MLVA14 than with the MLST-7 (HGDI = 0.793 vs 0.574). RS1-MLVA14 was able to split some sequevars more broadly than MLST-7 (S4 Table). However, it should be noted that both MLST-7 and RS1-MLVA14 schemes were unable to subdivide sequevars I-13, I-14, I-16, I-34, I-46, each of which were represented by only a single ST and a single MT.

3. RS1-MLVA14 revealed genetic diversity among Malagasy phylotype I populations and was discriminative at the field scale

The RS1-MLVA14 scheme was applied on 148 Malagasy isolates. The 14 VNTR loci were amplified from all isolates except the GMIch_3461 locus, which was not amplified, despite several assays in two isolates (RUN0306 and RUN0307). Nonetheless, the latter was retained for the analysis of diversity (H_{nb} , allelic richness, private allelic richness), genetic structure (F_{ST} , R_{ST} , MDS) and the genotype accumulation curve (GAC). Overall, 31 haplotypes were identified, including two major haplotypes: MT007 and MT004, which represented 30.14% and 16.44% of the isolates, respectively (S5 Table). The frequency of the other haplotypes varied from 0.68% to 5.48%. Regarding the host of isolation (S3 Fig), the majority of haplotypes (70.97%) were isolated from one host: 32% haplotypes were isolated from *S. lycopersicum*, 32% from *C. annuum*, 18% from *S. aethiopicum*, 14% from *S. tuberosum* and 4% from *S. melongena*. However, some haplotypes were found on several hosts: the most prevalent haplotype MT007 was found on four hosts (*S. aethiopicum*, *S. lycopersicum*, *S. melongena* and *S. tuberosum*); MT015 was found on three hosts (*S. aethiopicum*, *S. lycopersicum* and *S. tuberosum*); and seven haplotypes were found on two hosts, MT004 and MT006 (*S. melongena*, *S. lycopersicum*), MT009 (*S. melongena*, *S. scabrum*), MT020 (*S.*

lycopersicum, *C. annuum*), MT018 and MT022 (*S. lycopersicum*, *S. tuberosum*) and MT023 (*S. nigrum*, *S. tuberosum*) (S3 Fig).

The number of haplotypes per field varied from one to six (Fig 2). Interestingly, three haplotypes MT004, MT006 and MT007 were present at both high and low altitudes (Fig 2). The most prevalent haplotype MT007, identified in five fields (Fig 2), was found in the three studied regions (Fig 2). The haplotypes MT004 and MT006, identified in three and two fields, respectively (Fig 2), were found in the regions of Itasy and Atsinanana (Fig 2). Four haplotypes (MT015, MT020, MT023 and MT024) were present in the highland regions of Analamanga and Itasy (Fig 2). The 24 remaining haplotypes were identified in only one region: 9 haplotypes in the Atsinanana region, 8 haplotypes in the Itasy region and 7 haplotypes in the Analamanga region (Fig 2).



Source of the SRTM data: United States Geological Survey (USGS)

Fig 2. Distribution and frequency of phylotype I haplotypes collected in the 16 fields of the Atsinanana, Analamanga and Itasy regions

The 31 MLVA type (MT) are displayed. Each plot is indicated by a black dot and the number of isolates per plot (n) is mentioned.

The overall genetic diversity of Malagasy phylotype I isolates was $H_{nb} = 0.226$. However, there were marked differences between the three regions, with a H_{nb} value of 0.109, 0.318, and 0.430 for Itasy, Analamanga and Atsinanana regions, respectively. The average allelic richness over loci and the average private allelic richness over loci also differ between the three regions, 1.54 and 0.05 in Itasy, 2.32 and 0.49 in Analamanga, and 2.98 and 1.28 in Atsinanana (S6 Table).

Genetic differentiation among a priori populations was very important and significant between Itasy and Atsinanana regions ($F_{ST} = 0.288^{***}$, $R_{ST} = 0.381^{***}$), important and significant between the Itasy and Analamanga regions ($F_{ST} = 0.166^{***}$, $R_{ST} = 0.252^{***}$), and moderate but significant between the Atsinanana and Analamanga regions ($F_{ST} = 0.109^{***}$, $R_{ST} = 0.0897^*$) (S7 Table). An important and significant genetic differentiation ($F_{ST} = 0.242^{***}$, $R_{ST} = 0.336^{***}$) was shown between populations from lowlands (0-100 m) and highlands (1200-1500 m) and a moderate but significant differentiation according to the $F_{ST} = 0.079^{***}$ (R_{ST} value was not significant) between populations from mid-elevation (900-1200 m) and highland areas (1200-1500 m), and between lowland (0-100 m) and mid-elevation (900-1200 m) ($F_{ST} = 0.095^{**}$; R_{ST} value was not significant) (S7 Table).

4. RS1-MLVA14 revealed genetic structure among Malagasy phylotype I populations and showed the singularity of the most prevalent genetic cluster

The MDS representation showed that the Malagasy haplotypes could be grouped into two clusters with a probability of 67.5% (Fig 3). Cluster 1 and cluster 2 included 25 and 6 haplotypes, respectively. Haplotypes were assigned to these clusters with high probability (90.9 %). Axis 1 and Axis 2 explained 81.7% and 9.2% of the variance, respectively (Fig 3).

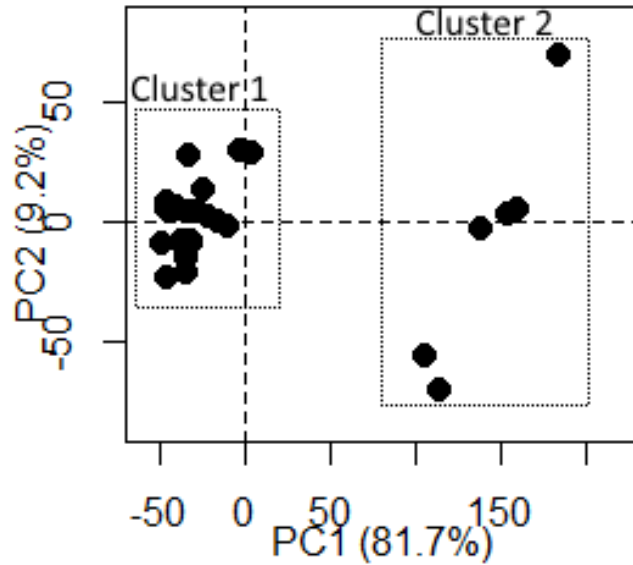


Fig 3. Multidimensional scaling (MDS) of the Malagasy phylotype I (total = 31 haplotypes, 291 isolates)

Cluster 1 and cluster 2 represent the two groups of haplotypes.

A network representation (MST) revealed the genetic relationships between the 31 Malagasy haplotypes (Fig 4). A major Malagasy clonal complex (MCC1), composed of isolates that differed by only one VNTR, was identified. MCC1 gathered 23 haplotypes (cluster 1, 123 isolates), from which two founder haplotypes (the most prevalent haplotypes, MT007 and MT004) were predicted. Three minor Malagasy clonal complexes were identified: MCC2 (cluster 1, 3 haplotypes, 12 isolates), MCC3 (cluster 2, 3 haplotypes, 6 isolates), and MCC4 (cluster 2, 2 haplotypes, 5 isolates).

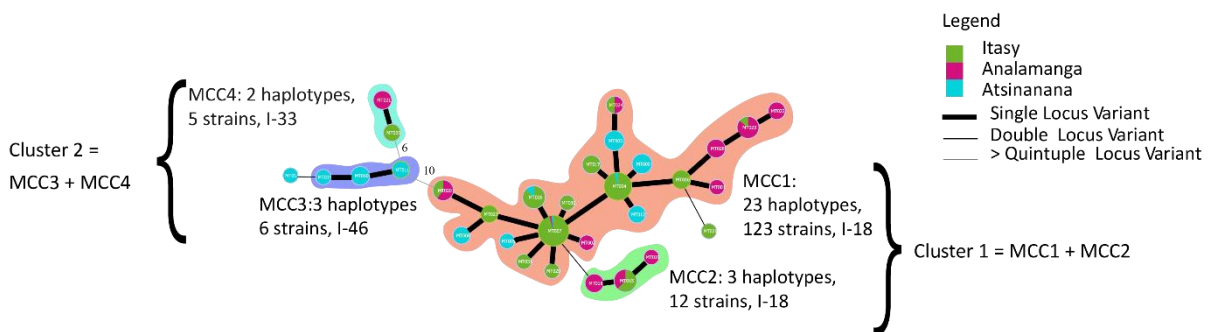


Fig 4: Minimum spanning tree (MST) of the Malagasy phylotype I according to the region of collect

The haplotypes were identified by using goeBURST full MST in PHYLOVIZ. Each MLVA type (MT) is displayed as a circle, the size of which is proportional to the number of isolates represented. The different colours indicate the region where data was collected.

The branch thickness indicates the number of locus differences between the neighbouring haplotypes. MCC1, MCC2, MCC3, MCC4 represent the Malagasy Clonal Complexes 1, 2, 3 and 4. A clonal complex is composed of haplotypes that differ only by one VNTR locus. I-18, I-33 and I-46 represents the sequevar of the isolates.

Interestingly, the MDS and MST representations were congruent with the phylogenetic assignment (i.e. sequevar) of the isolates based on *egl* partial sequences. Indeed, cluster 1 (MCC1, MCC2) grouped the isolates belonging to the sequevar 18, while cluster 2 gathered isolates belonging to sequevar 33 (MCC4) and sequevar 46 (MCC3). Moreover, MCC1 was the only CC which included isolates isolated from both highlands and lowlands. The remaining CCs comprised isolates isolated from the highlands only (MCC2, MCC4) or from the lowlands only (MCC3) (Fig 4). All CCs included isolates collected from several hosts except MCC3, for which isolates were only collected from *C. annuum* (S3 Fig).

A global MST was built to display the genetic relationships between phylotype I isolates from Madagascar (146 isolates, 31 haplotypes) and from worldwide isolates (145 isolates, 76 haplotypes) (Fig 5). Only the seven worldwide haplotypes, which had genetic links with the 31 Malagasy haplotypes, are shown in Fig 5.

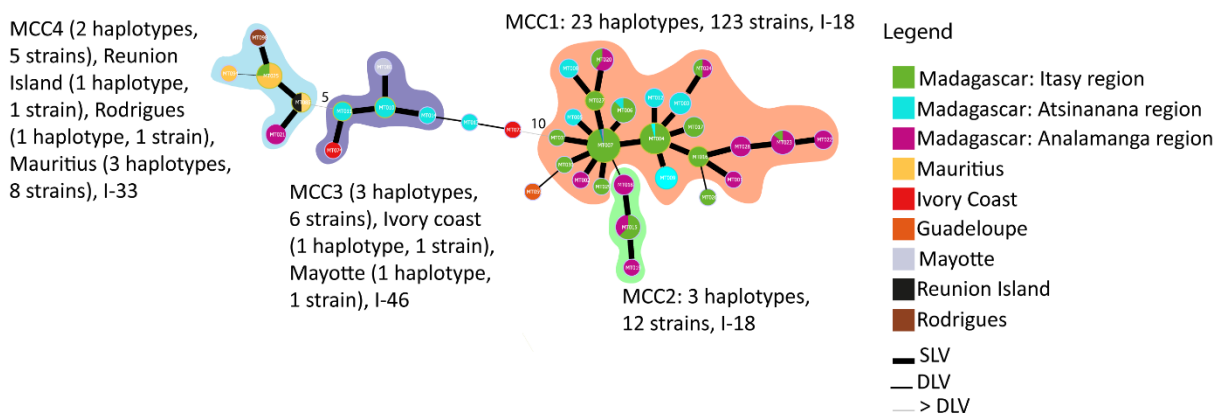


Fig 5. Minimum spanning tree of the Malagasy and worldwide phylotype I

The haplotypes were identified by using goeBURST full MST in PHYLOVIZ. Each MLVA type (MT) is displayed as a circle, the size of which is proportional to the number of isolates represented. The different colours indicate the country. The branch thickness indicates the number of locus differences between the neighbouring haplotypes. MCC1, MCC2, MCC3, MCC4 represent the Malagasy Clonal Complex 1, 2, 3 and 4. A clonal complex

is composed of haplotypes that differ only by one VNTR locus. The sequevar of the isolates are represented as I-18, I-33 and I-46.

The global MST revealed 4 clonal complexes. Strikingly, Malagasy haplotypes in cluster 1 (MCC1, MCC2), corresponding to sequevar 18, appeared genetically distant from worldwide phylotype I isolates. Indeed, the major Malagasy MCC1 differed to international haplotypes by at least ten loci (except one strain from Guadeloupe, MT059, which differed by 2 loci). The Malagasy minor MCC2 appeared closely related to MCC1 and unrelated to other worldwide haplotypes. In contrast, Malagasy haplotypes in cluster 2 (MCC3 and MCC4, which correspond to sequevars 46 and 33, respectively) appeared related to other worldwide haplotypes. Indeed, the three Malagasy haplotypes (MT010, MT011, and MT014) belonging to MCC3 were very closely related to haplotypes from Mayotte and Côte d'Ivoire (single-locus variant). Thus, they were grouped in the same CC. Interestingly, the haplotypes from Madagascar (MT010, MT011 and MT014) and Mayotte (MT081) were isolated from the same host species, *C. annuum*. The two Malagasy haplotypes, MT021 and MT025, belonging to MCC4 were included in a CC, which gathered isolates from Rodrigues, Mauritius and Réunion. Interestingly, within this CC, the haplotype MT025 isolated from *S. lycopersicum* was found both in Mauritius and in Madagascar. *S. tuberosum* was also the host plant shared by the haplotypes collected in Madagascar (MT021) and Mauritius (MT025, MT094).

Discussion

A growing interest in the use of MLVA on plant pathogenic bacteria for studying population biology has been observed in recent years [4,11,13–17]. As far as the RSSC is concerned, MLVA schemes have been applied to the global surveillance of phylotype III [38], to trace potential sources of contamination by phylotype IIB-1 in England and Ethiopia [21,60]. Very recently, they have been used to study the genetic diversity and structure of phlotypes I and IIB-1 populations in Uganda [59]. The study of RSSC populations (phlotypes I, II and III) collected in the central highlands of Madagascar in 2013 was the first project of such magnitude (1224 isolates collected) [19]. Remarkably, this work revealed contrasting epidemiological patterns between phylotype IIB-1 and phylotype III populations. The population biology of Malagasy phylotype I had not yet

been studied due to the absence of reliable MLVA markers to amplify phylotype I isolates [19]. The objective of the present study was to: develop and apply an optimized MLVA scheme to study the population genetics of Malagasy phylotype I; analyse the genetic relationships with worldwide phylotype I; and compare the epidemiological situation of phylotype I with Malagasy phylotypes IIB-1 and III.

1. RS1-MLVA14 scheme, a genotyping tool adapted to study the population genetics of phylotype I, both at field and global scales

By using previously identified VNTR markers [20,22,38], as well as by screening genomes for new VNTR markers, we developed an optimized MLVA scheme adapted to RSSC phylotype I isolates. The 14 selected VNTR markers were all present in phylotype I, to a lesser extent in phylotype III and rarely found in phylotypes II and IV. This is consistent with the species delineation of phylotypes I and III (belonging to *R. pseudosolanacearum*) compared to the phylotype II (*R. solanacearum*) and phylotype IV (*R. syzygii*) [27,28]. The RS1-MLVA14 scheme showed a very high typeability because only one VNTR could not be amplified from 2 out of the 293 phylotype I isolates tested. We assessed whether the RS1-MLVA14 scheme improved genotypic resolution compared to MLST-7 [37]. MLST is the gold standard for epidemiological surveillance and outbreak investigations of bacterial diseases. The value of the discriminatory power of the RS1-MLVA14 scheme (HGDI = 0.793) was higher than for the MLST-7 scheme (HGDI = 0.574). This revealed greater genetic and haplotypic diversity and highlighted the interest of this new MLVA scheme. VNTR loci with low diversity values are useful to establish phylogenetic relationships. Those with a high diversity have a strong discriminatory power [61]. VNTR are supposed to have high mutation rates, which increases the likelihood of homoplasy [62]. Another point is that molecular samples with lower values of HGDI may indicate high levels of homoplasy. However, low values may also be due to low mutation rates [62]. The RS1-MLVA14 scheme combines VNTR loci with variable diversity values and retains the phylogenetic signal, since our MDS and MST analyses showed that the isolates were grouped according to their phylogenetic assignment (sequevar). These results suggest that the probability of homoplasy is low. They support previous studies, such as the work on phylotype III of the RSSC [38] or other plant pathogens [3,17]. Lastly, we showed that the RS1-MLVA14 scheme can be used to explore the genetic diversity at the field scale

(several haplotypes were disclosed in single fields), as well as at a global scale. Thus, RS1-MLVA14 is well suited to deciphering the epidemiology of phylotype I populations.

2. RS1-MLVA14 unveiled contrasting genetic diversity and epidemiology among Malagasy phylotype I populations

The analysis of phylotype I populations revealed contrasting genetic diversity depending on the three Malagasy regions. Genetic diversity is greater in the lowlands than in highland areas. In the lowlands, the Atsinanana region imports huge amounts of vegetables from other Malagasy producing areas, such as Ambatondrazaka or wholesale markets in Antananarivo, the Malagasy capital [63]. In addition, this region has the Toamasina seaport, where there is a great deal of trade in plant material with foreign countries, which could be responsible for the introduction of contaminated agricultural products in this lowland region. It might also explain why the level of genetic diversity of phylotype I isolates is greater.

Our findings showed a lower genetic diversity in phylotype I populations in the central Malagasy highlands. Some haplotypes were found in different fields in the same region (for example MT007 was detected in plots 7, 9 and 12 in the Itasy region), which suggests possible field contamination from exchanges of infected plant material. Indeed, producers do not always have access to healthy seeds [64]. Farmers often produce their own seeds or exchange seeds with other farmers, with no sanitary guarantee. Trade in agricultural products might encourage the propagation of haplotypes between regions. Most vegetables from the Itasy region are sold in the capital, Antananarivo (Analamanga region), at Madagascar's wholesale markets, where large volumes of agricultural products are traded [64].

Another interesting point shown by our study is that Malagasy phylotype I populations (sequevars 18, 33 and 46) are genetically differentiated depending on the elevation. Remarkably, a similar situation was observed in China, where the genetic structure of phylotype I isolates (for sequevars 13, 14, 15, 17, 34, 44, 54 and 55) was associated with elevation [65]. Sequevar 18 seems to be adapted to both warm and cool temperatures in Madagascar, since it was isolated in the lowlands at sea level in the Atsinanana region (temperature range is from 17°C to 30.1°C), as well as in the highlands at an altitude of up to 1500m in the Analamanga region (from 8.9°C to 26.6°C) and the Itasy region (from

9.1°C to 28°C) [66]. Interestingly, sequevar 31 isolated in the lowlands and up to altitudes above 1000m in Réunion [37], as well as in Côte d'Ivoire [67], exhibited a similar environmental distribution. So far, sequevar 18 has always been reported from lowlands in the Americas [68–70], Africa [67], Asia [71–73], Oceania [74] and on islands in the South-West Indian Ocean [37,75]. In contrast, even though a very limited number of isolates was used in our study, sequevar 33 was only isolated at moderate and high elevations (above 900m) and sequevar 46 was only isolated in lowlands (below 300m). Thus, in addition to the studies in China [65], Réunion [37] and Côte d'Ivoire [67], our study in Madagascar shed new light on the ecology of phylotype I isolates, which are often reported as only being composed of tropical and subtropical isolates adapted to lowland areas. Altogether, the ecology of phylotype I isolates differs from that of phylotype IIB-1 and phylotype III isolates, which are considered as cold-tolerant and adapted to highland areas [19,23,24,76,77]. Further research and massive sampling of phylotype I isolates would certainly provide a better understanding of the differences in the prevalence of sequevars according to elevation and the factors determining their differential environmental adaptation.

Our study of population structure showed that Malagasy phylotype I isolates were distributed in two genetic clusters, which displayed different epidemiological features. A striking result is the singularity of the most prevalent Malagasy isolates belonging to sequevar 18 (corresponding to cluster 1). Indeed, despite the fact that sequevar 18 is distributed worldwide [37,67–73,75], our study revealed no genetic links between Malagasy isolates and worldwide isolates (South-West Indian Ocean, Africa, Americas, Asia isolates). The closest haplotype (one strain) was a double-locus variant from Guadeloupe. Cluster 1 could be derived from the evolution of an ancient introduction of isolates that evolved locally. This would explain the differentiation from worldwide isolates, with short-distance dissemination restricted to Madagascar. This work must be continued by integrating more Malagasy and global phylotype I isolates to further our understanding of the singularity of isolates belonging to cluster 1. In contrast, cluster 2, which included the least prevalent isolates in Madagascar (sequevars 33 and 46), showed genetic links with isolates from South-West Indian Ocean islands (Mayotte, Mauritius, Rodrigues and Réunion). So far, apart from the recent introduction of phylotype I via rose cuttings in the Netherlands [78], no long-distance dissemination of phylotype I isolates

has been reported. Only the spread of phylotype IIB isolates via bananas, potato tubers and geranium cuttings [79–81] has been clearly documented. Interestingly, based on the *egl* sequence analysis, the phylotype I isolates introduced in the Netherlands were assigned to sequevar 33 [82], the sequevar that is likely to have been disseminated between Madagascar and other South-West Indian Ocean islands. Moreover, the *egl* sequences from the Dutch isolates appeared 100% identical to *egl* sequences from isolates from Madagascar, Mauritius [37], Rodrigues [37,83]

and India [84]. This highlights the phylogenetic links between isolates from distant geographical areas. As far as the sequevar 46 is concerned, it has now been reported in Madagascar and Mayotte (South-West Indian Ocean) [75], as well as in more geographically distant areas, such as Côte d’Ivoire [67] and Myanmar [85]. All these data strongly support the theory of the global dissemination of phylotype I, which is probably linked to human activities and the transport of contaminated material [86].

3. Comparative epidemiology of Malagasy phylotype I, IIB-1 and III populations

Our study showed that the level of genetic diversity of Malagasy phylotype I populations was globally low ($H_{nb} = 0.226$). This result was surprising since, according to McDonald and Linde [1], phylotype I is considered to have a high evolutionary potential due to its natural ability to transform and recombine [37,86,87]. The level of genetic diversity of Malagasy phylotype I populations was quite similar to that of Malagasy phylotype IIB-1 populations ($H_{nb} = 0.19$) and clearly lower than that of Malagasy phylotype III populations ($H_{nb} = 0.40$) [19]. As determined for Malagasy phylotype IIB-1 [19], the Malagasy phylotype I haplotypes appeared to be closely related genetically and gathered into clonal complexes. Some haplotypes were shared between fields in the same agroecological region and even between agroecological areas (at low and high elevations). In contrast to the endemic nature of phylotype III [19], these results suggest that the phylotype I isolates have a similar epidemic population pattern to phylotype IIB-1 isolates.

In Madagascar’s central highlands above 1000 m elevation, phylotype I populations co-occurred with populations of phylotype IIB-1 and III [19]. In addition, as for phylotypes IIB-1 and III [19], phylotype I populations were isolated from many solanaceous crops (*S. lycopersicum*, *S. tuberosum*, *S. melongena*, *S. gilo* and *C. annuum*), and weeds (*S. nigrum*,

S. scabrum). Weeds play an important role. They often remain in the soil after harvesting and harbour pathogenic populations, which survive in the environment as a result [24,88,89]. Our data supports this because although our weed sampling was very limited, RS1-MLVA14 did not reveal specific haplotypes on weeds. However, in the central Malagasy highlands, phylotype I populations were the least prevalent (10% instead of 18% for phylotype III and 72% for phylotype IIB-1) [19]. This observation could be explained by the fact that phylotype IIB-1 populations are reported to be better adapted to low temperatures compared to phylotype III populations and even more so compared to phylotype I populations [76,77,90]. Interestingly, a comparative pathogenicity test of Malagasy phylotypes I, IIB-1 and III isolates on eight potato varieties showed that all isolates were strongly pathogenic at tropical lowland temperatures, but that IIB-1 isolates were the most aggressive [18]. Similar comparative pathogenicity tests should now be performed on different hosts at cold temperatures to determine whether temperature influences the capacity of the three phylotype isolates to co-occur in the same cropping areas and cause disease.

In conclusion, during this study, we developed an optimized MLVA scheme dedicated to phylotype I populations. We showed that the RS1-MLVA14 scheme is highly resolutive at global, regional and field scales, which makes it suitable for epidemiological studies. This work on Malagasy phylotype I represents a first step. Our research was limited to three vegetable growing regions in Madagascar. A broader study has been launched with a wider spatial scale, including all the vegetable cropping areas in Madagascar. The goal is to further our understanding of the migration routes and population biology of phylotype I in Madagascar and between Madagascar and other countries.

Supporting information

S1 Table. Characteristics of the 148 Malagasy isolates used in the study. The isolates used for the comparative analysis of MLST-7 and RS1-MLVA14 are in bold.

S2 Table. Characteristics of the 145 worldwide isolates used in the study. The isolates used for the comparative analysis of MLST-7 and RS1-MLVA14 are in bold.

S3 Table. Screening of the 14 selected VNTR loci in the 114 publicly available RSSC genomes.

S4 Table. Correspondence between sequevar, ST and haplotype for the comparative analysis of MLST-7 and RS1-MLVA14.

S5 Table. Frequency of the Malagasy MLVA Type.

S6 Table. Diversity intra-population indexes.

S7 Table. F_{ST} and R_{ST} between populations from the three regions where data was collected and between populations from lowlands, mid-elevation and highlands.

S1 Fig. Localization of the 14 VNTRs on the chromosome and megaplasmid of the reference phylotype I strain GMI1000.

S2 Fig. Genotype accumulation curve for 291 strains of RSSC phylotype I strains genotyped over 14 loci. The horizontal axis represents the number of loci randomly sampled without replacement up to $n-1$ loci, the vertical axis shows the number of unique MLVA types observed ($n = 107$) in the data set. The level of 100% of unique MLVA types detected is indicated with a dotted red line.

S3 Fig. Minimum spanning tree (MST) of the Malagasy phylotype I according to the host of collect. The haplotypes were identified by using goeBURST full MST in PHYLOVIZ. Each MLVA type (MT) is displayed as a circle, the size of which is proportional to the number of isolates represented. The different colours indicate the sample host. Branch thickness indicates the number of locus differences between the neighbouring haplotypes. MCC1, MCC2, MCC3, MCC4 represent the Malagasy Clonal Complexes 1, 2, 3 and 4. A clonal complex is composed of haplotypes that differ only by one VNTR locus. I-18, I-33 and I-46 represent the phylotype-sequevar of the isolates.

Acknowledgments

The authors wish to thank Olivier Pruvost for advice on the development of the MLVA scheme and data analyses, Laurence Humeau for advice on genotype accumulation curve analysis, Jean-Jacques Chéron for helping with the microbiology lab work and Karine Boyer for assisting with the molecular biology lab work. We would like to thank all the scientific and technical teams, who took part, including those from CIRAD-3P (Saint Pierre, Réunion), ANSES (Saint-Pierre, Réunion), FOFIFA (Antananarivo, Madagascar),

the CSA agricultural service centres from the Ministry of Agriculture (Madagascar) and CEFFEL/FERT Centre (Antsirabe, Madagascar).

References

1. McDonald BA, Linde C. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu Rev Phytopathol.* 2002;40: 349–379. doi:10.1146/annurev.phyto.40.120501.101443
2. Milgroom MG. Population biology of plant pathogens: genetics, ecology, and evolution. The American Phytopathological Society; 2017. doi:10.1094/9780890544525
3. López-Soriano P, Boyer K, Cesbron S, Morente MC, Peñalver J, Palacio-Bielsa A, et al. Multilocus Variable Number of Tandem Repeat Analysis reveals multiple introductions in Spain of *Xanthomonas arboricola* pv. *pruni*, the causal agent of bacterial spot disease of stone fruits and almond. Vinatzer BA, editor. *PLOS ONE.* 2016;11: e0163729. doi:10.1371/journal.pone.0163729
4. Bühlmann A, Dreo T, Rezzonico F, Pothier JF, Smits THM, Ravnikar M, et al. Phylogeography and population structure of the biologically invasive phytopathogen *Erwinia amylovora* inferred using minisatellites. *Environ Microbiol.* 2014;16: 2112–2125. doi:doi:10.1111/1462-2920.12289
5. Mascheretti S, Croucher PJP, Vettraino A, Prospero S, Garbelotto M. Reconstruction of the sudden oak death epidemic in California through microsatellite analysis of the pathogen *Phytophthora ramorum*: genetic structure of *P.ramorum* in California. *Mol Ecol.* 2008;17: 2755–2768. doi:10.1111/j.1365-294X.2008.03773.x
6. Vergnaud G, Pourcel C. Multiple Locus Variable Number of Tandem Repeats Analysis. In: Caugant DA, editor. *Molecular epidemiology of microorganisms.* Totowa, NJ: Humana Press; 2009. pp. 141–158. doi:10.1007/978-1-60327-999-4_12
7. Pourcel C, Vergnaud G. Strain typing using Multiple "Variable Number of Tandem Repeat" Analysis and genetic element CRISPR. 2011.

8. Van Belkum A. Tracing isolates of bacterial species by Multilocus Variable number of tandem repeat Analysis (MLVA). *FEMS Immunol Med Microbiol.* 2007;49: 22–27. doi:10.1111/j.1574-695X.2006.00173.x
9. Zaluga J, Stragier P, Van Vaerenbergh J, Maes M, De Vos P. Multilocus Variable Number Tandem Repeats Analysis (MLVA) distinguishes a clonal complex of *Clavibacter michiganensis subsp. michiganensis* strains isolated from recent outbreaks of bacterial wilt and canker in Belgium. *BMC Microbiol.* 2013;13: 126. doi:10.1186/1471-2180-13-126
10. Gironde S, Manceau C. Housekeeping gene sequencing and Multilocus Variable Number Tandem Repeat Analysis to identify subpopulations within *Pseudomonas syringae pv. maculicola* and *Pseudomonas syringae pv. tomato* that correlate with host specificity. *Appl Environ Microbiol.* 2012;78: 3266–3279. doi:10.1128/AEM.06655-11
11. Francisco CS, Ceresini PC, Almeida RPP, Coletta-Filho HD. Spatial genetic structure of coffee associated *Xylella fastidiosa* populations indicates that cross infection does not occur with sympatric citrus orchards. *Phytopathology.* 2017;107: 395–402. doi:10.1094/PHYTO-08-16-0300-R
12. Bui Thi Ngoc L, Verniere C, Jarne P, Brisse S, Guerin F, Boutry S, et al. From local surveys to global surveillance: three high-throughput genotyping methods for epidemiological monitoring of *Xanthomonas citri pv. citri* pathotypes. *Appl Environ Microbiol.* 2009;75: 1173–1184. doi:10.1128/AEM.02245-08
13. Pruvost O, Magne M, Boyer K, Leduc A, Tourterel C, Drevet C, et al. A MLVA genotyping scheme for global surveillance of the Citrus Pathogen *Xanthomonas citri pv. citri* suggests a worldwide geographical expansion of a single genetic lineage. Vinatzer BA, editor. *PLoS ONE.* 2014;9: e98129. doi:10.1371/journal.pone.0098129
14. Poulin L, Grygiel P, Magne M, Gagnevin L, Rodriguez-R LM, Forero Serna N, et al. New Multilocus Variable Number Tandem Repeat Analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by

- Xanthomonas oryzae*. Goodrich-Blair H, editor. Appl Environ Microbiol. 2015;81: 688–698. doi:10.1128/AEM.02768-14
15. Rache L, Blondin L, Flores C, Trujillo C, Szurek B, Restrepo S, et al. An Optimized Microsatellite Scheme for Assessing Populations of *Xanthomonas phaseoli* pv. *manihotis*. Phytopathology. 2019;109: 859–869. doi:10.1094/PHYTO-06-18-0210-R
 16. Nakato GV, Fuentes Rojas JL, Verniere C, Blondin L, Coutinho T, Mahuku G, et al. A new Multi Locus Variable Number of Tandem Repeat Analysis scheme for epidemiological surveillance of *Xanthomonas vasicola* pv. *musacearum*, the plant pathogen causing bacterial wilt on banana and enset. Chiang T-Y, editor. PLOS ONE. 2019;14: e0215090. doi:10.1371/journal.pone.0215090
 17. Ferreira MASV, Bonneau S, Briand M, Cesbron S, Portier P, Darrasse A, et al. *Xanthomonas citri* pv. *viticola* affecting grapevine in Brazil: emergence of a successful monomorphic pathogen. Front Plant Sci. 2019;10: 489. doi:10.3389/fpls.2019.00489
 18. Ravelomanantsoa S. Biologie des populations du complexe d'espèces *Ralstonia solanacearum* appliquée à l'épidémiologie du flétrissement bactérien de la pomme de terre à Madagascar. Université de La Réunion; 2016.
 19. Ravelomanantsoa S, Vernière C, Rieux A, Costet L, Chiroleu F, Arribat S, et al. Molecular epidemiology of bacterial wilt in the Madagascar highlands caused by Andean (Phylotype IIB-1) and African (Phylotype III) brown rot strains of the *Ralstonia solanacearum* species complex. Front Plant Sci. 2018;8. doi:10.3389/fpls.2017.02258
 20. N'Guessan CA, Brisse S, Le Roux-Nio A-C, Poussier S, Koné D, Wicker E. Development of variable number of tandem repeats typing schemes for *Ralstonia solanacearum*, the agent of bacterial wilt, banana Moko disease and potato brown rot. J Microbiol Methods. 2013;92: 366–374. doi:10.1016/j.mimet.2013.01.012
 21. Parkinson N, Bryant R, Bew J, Conyers C, Stones R, Alcock M, et al. Application of Variable-Number Tandem-Repeat Typing to discriminate *Ralstonia solanacearum*

- strains associated with English watercourses and disease outbreaks. *Appl Environ Microbiol.* 2013;79: 6016–6022. doi:10.1128/AEM.01219-13
22. Guinard J, Latreille A, Guérin F, Poussier S, Wicker E. New Multilocus Variable Number Tandem Repeat Analysis (MLVA) scheme for fine scale monitoring and microevolution related study of *Ralstonia pseudosolanacearum* phylotype I populations. Drake HL, editor. *Appl Environ Microbiol.* 2017;83. doi:10.1128/AEM.03095-16
 23. Elphinstone J. The current bacterial wilt situation : a global overview. *Bacterial wilt disease and the Ralstonia solanacearum species complex.* 2005; 9–28.
 24. Hayward AC. The hosts of *Pseudomonas solanacearum*. *Bacterial wilt: the disease and its causative agent, Pseudomonas solanacearum.* 1994; 9–24.
 25. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, et al. Top 10 plant pathogenic bacteria in molecular plant pathology: Top 10 plant pathogenic bacteria. *Mol Plant Pathol.* 2012;13: 614–629. doi:10.1111/j.1364-3703.2012.00804.x
 26. Prior P, Fegan M. Recent developments in the phylogeny and classification of *Ralstonia solanacearum*. *Acta Hortic.* 2005; 127–136. doi:10.17660/ActaHortic.2005.695.14
 27. Prior P, Ailloud F, Dalsing BL, Remenant B, Sanchez B, Allen C. Genomic and proteomic evidence supporting the division of the plant pathogen *Ralstonia solanacearum* into three species. *BMC Genomics.* 2016;17: 90. doi:10.1186/s12864-016-2413-z
 28. Safni I, Cleenwerck I, De Vos P, Fegan M, Sly L, Kappler U. Polyphasic taxonomic revision of the *Ralstonia solanacearum* species complex: proposal to emend the descriptions of *Ralstonia solanacearum* and *Ralstonia syzygii* and reclassify current *R. syzygii* strains as *Ralstonia syzygii subsp. syzygii subsp. nov.*, *R. solanacearum* phylotype IV strains as *Ralstonia syzygii subsp. indonesiensis subsp. nov.*, banana blood disease bacterium strains as *Ralstonia syzygii subsp. celebesensis subsp. nov.* and *R. solanacearum* phylotype I and III strains as *Ralstonia pseudosolanacearum sp. nov.* *Int J Syst Evol Microbiol.* 2014;64: 3087–3103. doi:10.1099/ijs.0.066712-0

29. Fegan M, Prior P. How complex is the “*Ralstonia solanacearum* Species Complex.” APS Press. Bacterial wilt disease and the *Ralstonia solanacearum* species complex. APS Press. Minnesota: Allen, C., Prior, P., Hayward, A.C.; 2005. pp. 449–461.
30. Bouriquet G. Les maladies du tabac à Madagascar. Annu Cryptogam Exot. 1934;VII: 97–112.
31. Bouriquet G. Madagascar : list of the parasites and diseases of cultivated plants. Int Bull Plant Prot. 1934;VIII: 99–100.
32. Bouriquet G. Les maladies des plantes cultivées à Madagascar. Encycl Mycol. 1946;XII: 545.
33. Dadant R, Rasolofo R, Baudin P. Liste des maladies des plantes cultivées à Madagascar. 1960;17: 94.
34. Lallmahomed GM, Rakotobe Rabehevitra E, Rakotondramanana. Biovars and races of *Pseudomonas solanacearum* in Madagascar a preliminary study. FAO Plant Prot Bull. 1988;36: 54–59.
35. Rabakoarihanta A, Rakotondramanana. “Potato production status in Madagascar,” in Potato development and transfer of technology in tropical africa. International Potato Center. Nairobi - Kenya; 1984.
36. Kelman A. The relationship of pathogenicity of *Pseudomonas solanacearum* to colony appearance in a tetrazolium medium. Phytopathology. 1954;54: 693–695.
37. Yahiaoui N, Chéron J-J, Ravelomanantsoa S, Hamza AA, Petrousse B, Jeetah R, et al. Genetic diversity of the *Ralstonia solanacearum* species complex in the South-West Indian Ocean islands. Front Plant Sci. 2017;8: 2139. doi:10.3389/fpls.2017.02139
38. Ravelomanantsoa S, Robène I, Chiroleu F, Guérin F, Poussier S, Pruvost O, et al. A novel multilocus variable number tandem repeat analysis typing scheme for African phylotype III strains of the *Ralstonia solanacearum* species complex. PeerJ. 2016;4: e1949. doi:10.7717/peerj.1949

39. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28: 1647–1649. doi:10.1093/bioinformatics/bts199
40. Salanoubat M, Genin S, Artiguenave F, Gouzy J, Mangenot S, Arlat M, et al. Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature*. 2002;415: 497–502. doi:10.1038/415497a
41. Li Z, Wu S, Bai X, Liu Y, Lu J, Liu Y, et al. Genome sequence of the tobacco bacterial wilt pathogen *Ralstonia solanacearum*. *J Bacteriol*. 2011;193: 6088–6089. doi:10.1128/JB.06009-11
42. Guinard J, Vinatzer BA, Poussier S, Lefeuvre P, Wicker E. Draft genome sequences of nine strains of *Ralstonia solanacearum* differing in virulence to eggplant (*Solanum melongena*). *Genome Announc*. 2016;4: e01415-01415.
43. Mrazek J, Guo X, Shah A. Simple sequence repeats in prokaryotic genomes. *Proc Natl Acad Sci*. 2007;104: 8472–8477. doi:10.1073/pnas.0702412104
44. Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res*. 1999;27: 573–580. doi:10.1093/nar/27.2.573
45. Mayer C. Phobos: a tandem repeat search tool. 2007.
46. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic Local Alignment Search Tool. *J Mol Biol*. 1990;215: 403–410.
47. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth B, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 2012;40: e115–e115.
48. Zhao S, Poulin L, Rodriguez-R LM, Serna NF, Liu S-Y, Wonni I, et al. Development of a Variable Number of Tandem Repeats Typing scheme for the bacterial rice pathogen *Xanthomonas oryzae pv. oryzicola*. *Phytopathology*. 2012;102: 948–956. doi:10.1094/PHYTO-04-12-0078-R

49. Yuan H, Qing LB, Zhi JD, Hua HL, Xia TX, Zhong ZJ. Identification of Variable-Number Tandem-Repeat (VNTR) sequences in *Acinetobacter pittii* and development of an optimized Multiple-Locus VNTR Analysis Typing scheme. *Biomed Env Sci*. 2015; 9.
50. Monteil M, Durand B, Bouchouicha R, Petit E, Chomel B, Arvand M, et al. Development of discriminatory multiple-locus variable number tandem repeat analysis for *Bartonella henselae*. *Microbiology*. 2007;153: 1141–1148. doi:10.1099/mic.0.2006/001164-0
51. Kamvar ZN, Tabima JF, Grünwald NJ. *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*. 2014;2: e281. doi:10.7717/peerj.281
52. Kamvar ZN, Brooks JC, Grünwald NJ. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front Genet*. 2015;6. doi:10.3389/fgene.2015.00208
53. Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol*. 1988;26: 2465–2466. doi:10.1128/JCM.26.11.2465-2466.1988
54. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 2010;10: 564–567.
55. Kalinowski ST. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes*. 2005;5: 187–189. doi:10.1111/j.1471-8286.2004.00845.x
56. Pelé J, Bécu J-M, Abdi H, Chabbert M. Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling. *BMC Bioinformatics*. 2012;13: 133. doi:10.1186/1471-2105-13-133
57. Borg I, Groenen PJ. *Modern multidimensional scaling: theory and applications*. Springer-Verlag New York; 2005.

58. Francisco AP, Vaz C, Monteiro PT, Melo-Cristino J, Ramirez M, Carriço JA. PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. BMC Bioinformatics. 2012;13: 87. doi:10.1186/1471-2105-13-87
59. Abdurahman A, Parker ML, Kreuze J, Elphinstone J, Struik PC, Kiggundu A, et al. Molecular epidemiology of *Ralstonia solanacearum* species complex strains causing bacterial wilt of potato in Uganda. Phytopathology. 2019; PHYTO-12-18-0476-R. doi:10.1094/PHYTO-12-18-0476-R
60. Abdurahman A, Griffin D, Elphinstone J, Struik PC, Schulz S, Schulte-Geldermann E, et al. Molecular characterization of *Ralstonia solanacearum* strains from Ethiopia and tracing potential source of bacterial wilt disease outbreak in seed potatoes. Plant Pathol. 2017;66: 826–834. doi:10.1111/ppa.12661
61. Keim P, Van Ert MN, Pearson T, Vogler AJ, Huynh LY, Wagner DM. Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. Infect Genet Evol. 2004;4: 205–213. doi:10.1016/j.meegid.2004.02.005
62. Estoup A, Jarne P, Cornuet J-M. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol. 2002;11: 1591–1604. doi:10.1046/j.1365-294X.2002.01576.x
63. Samad HM. Analyse de la filière tomate au lac Alaotra. Université d'Antananarivo; 2008.
64. FERT. Etude de la filière légumes sur les Hautes Terres de Madagascar régions Analamanga, Itasy, Vakinankaratra, Amoron'i Mania: pomme de terre, tomate, oignon, carotte, haricot vert, et chou. 2012 p. 88.
65. Liu Y, Wu D, Liu Q, Zhang S, Tang Y, Jiang G, et al. The sequevar distribution of *Ralstonia solanacearum* in tobacco-growing zones of China is structured by elevation. Eur J Plant Pathol. 2017;147: 541–551. doi:10.1007/s10658-016-1023-6
66. Rakotofiringa FMD. Caractérisation climatique des 22 chefs-lieux de région de Madagascar en utilisant les indices climatiques. Université d'Antananarivo; 2019.

67. N'Guessan CA, Abo K, Fondio L, Chiroleu F, Lebeau A, Poussier S, et al. So near and yet so far: the specific case of *Ralstonia solanacearum* populations from Côte d'Ivoire in Africa. *Phytopathology*. 2012;102: 733–740. doi:10.1094/PHYTO-11-11-0300
68. Santiago TR, Lopes CA, Caetano-Anolles G, Mizubuti ESG. Genetic structure of *Ralstonia solanacearum* and *Ralstonia pseudosolanacearum* in Brazil. *Plant Dis*. 2019; PDIS-09-19-1929-RE. doi:10.1094/PDIS-09-19-1929-RE
69. Deberdt P, Guyot J, Coranson-Beaudu R, Launay J, Noreskal M, Rivière P, et al. Diversity of *Ralstonia solanacearum* in French Guiana expands knowledge of the “emerging ecotype.” *Phytopathology*. 2014;104: 586–596. doi:10.1094/PHYTO-09-13-0264-R
70. Ramsubhag A, Lawrence D, Cassie D, Fraser R, Umaharan P, Prior P, et al. Wide genetic diversity of *Ralstonia solanacearum* strains affecting tomato in Trinidad, West Indies: Tomato bacterial wilt strains in Trinidad. *Plant Pathol*. 2012;61: 844–857. doi:10.1111/j.1365-3059.2011.02572.x
71. Lin C-H, Tsai K-C, Prior P, Wang J-F. Phylogenetic relationships and population structure of *Ralstonia solanacearum* isolated from diverse origins in Taiwan. *Plant Pathol*. 2014;63: 1395–1403. doi:10.1111/ppa.12209
72. Jiang G, Wei Z, Xu J, Chen H, Zhang Y, She X, et al. Bacterial wilt in China: history, current status, and future perspectives. *Front Plant Sci*. 2017;8: 1549. doi:10.3389/fpls.2017.01549
73. Carstensen GD, Venter SN, Wingfield MJ, Coutinho TA. Two *Ralstonia* species associated with bacterial wilt of *Eucalyptus*. *Plant Pathol*. 2017;66: 393–403. doi:10.1111/ppa.12577
74. Pastou D, Chéron JJ, Cellier G, Guérin F, Poussier S. First report of *Ralstonia pseudosolanacearum* phylotype I causing bacterial wilt in New Caledonia. *Plant Dis*. 2019; PDIS-05-19-1068-PDN. doi:10.1094/PDIS-05-19-1068-PDN

75. Chesneau T, Maignien G, Boyer C, Chéron J-J, Roux-Cuvelier M, Vanhuffel L, et al. Sequenar diversity and virulence of *Ralstonia solanacearum* phylotype I on Mayotte Island (Indian Ocean). *Front Plant Sci.* 2018;8. doi:10.3389/fpls.2017.02209
76. Milling A, Meng F, Denny TP, Allen C. Interactions with hosts at cool temperatures, not cold tolerance, explain the unique epidemiology of *Ralstonia solanacearum* race 3 biovar 2. *Phytopathology.* 2009;99: 1127–1134. doi:10.1094/PHYTO-99-10-1127
77. Cellier G, Prior P. Deciphering phenotypic diversity of *Ralstonia solanacearum* strains pathogenic to potato. *Phytopathology.* 2010;100: 1250–1261. doi:10.1094/PHYTO-02-10-0059
78. Tjou-Tam-Sin NNA, van de Bilt JLJ, Westenberg M, Gorkink-Smits PPMA, Landman NM, Bergsma-Vlami M. Assessing the pathogenic ability of *Ralstonia pseudosolanacearum* (*Ralstonia solanacearum* phylotype I) from ornamental *Rosa spp.* plants. *Front Plant Sci.* 2017;8: 1895. doi:10.3389/fpls.2017.01895
79. Janse JD. Potato brown rot in Western Europe – history, present occurrence and some remarks on possible origin, epidemiology and control strategies. *EPP0 Bull.* 1996;26: 679–695. doi:10.1111/j.1365-2338.1996.tb01512.x
80. Kim SH, Olson TN, Schaad NW, Moorman GW. *Ralstonia solanacearum* race 3, biovar 2, the causal agent of brown rot of potato, identified in geraniums in Pennsylvania, Delaware, and Connecticut. *Plant Dis.* 2003;87: 450–450. doi:10.1094/PDIS.2003.87.4.450C
81. Blomme G, Dita M, Jacobsen KS, Pérez Vicente L, Molina A, Ocimati W, et al. Bacterial diseases of bananas and enset: current state of knowledge and integrated approaches toward sustainable management. *Front Plant Sci.* 2017;8: 1290. doi:10.3389/fpls.2017.01290
82. Bergsma-Vlami M, van de Bilt JLJ, Tjou-Tam-Sin NNA, Westenberg M, Meekes ETM, Teunissen HAS, et al. Phylogenetic assignment of *Ralstonia pseudosolanacearum* (*Ralstonia solanacearum* phylotype I) isolated from *Rosa spp.* *Plant Dis.* 2018;102: 2258–2267. doi:10.1094/PDIS-09-17-1345-RE

83. Yahiaoui N, Chéron JJ, Jeetah R, Benimadhu S, Félicité J, Cellier G, et al. First report of *Ralstonia pseudosolanacearum* phylotype I causing bacterial wilt on Rodrigues Island, Indian Ocean. *Plant Dis.* 2016;100: 2522. doi:10.1094/PDIS-06-16-0811-PDN
84. Ramesh R, Gaitonde S, Achari G, Asolkar T, Singh NP, Carrere S, et al. Genome sequencing of *Ralstonia solanacearum* biovar 3, phylotype I, strains Rs-09-161 and Rs-10-244, isolated from eggplant and chili in India. *Genome Announc.* 2014;2: e00323-14, 2/3/e00323-14. doi:10.1128/genomeA.00323-14
85. Kyaw HWW, Tsuchiya K, Matsumoto M, Iiyama K, Aye SS, Zaw M, et al. Genetic diversity of *Ralstonia solanacearum* strains causing bacterial wilt of solanaceous crops in Myanmar. *J Gen Plant Pathol.* 2017;83: 216–225. doi:10.1007/s10327-017-0720-0
86. Wicker E, Lefeuvre P, de Cambiaire J-C, Lemaire C, Poussier S, Prior P. Contrasting recombination patterns and demographic histories of the plant pathogen *Ralstonia solanacearum* inferred from MLSA. *ISME J.* 2012;6: 961–974. doi:10.1038/ismej.2011.160
87. Coupat B, Chaumeille-Dole F, Fall S, Prior P, Simonet P, Nesme X, et al. Natural transformation in the *Ralstonia solanacearum* species complex: number and size of DNA that can be transferred. *FEMS Microbiol Ecol.* 2008;66: 14–24. doi:10.1111/j.1574-6941.2008.00552.x
88. Pradhanang PM, Elphinstone JG, Fox RTV. Identification of crop and weed hosts of *Ralstonia solanacearum* biovar 2 in the hills of Nepal. *Plant Pathol.* 2000;49: 403–413. doi:10.1046/j.1365-3059.2000.00480.x
89. Tusiime G, Adipala E, Opio F, Bhagsari AS. Weeds as latent hosts of *Ralstonia solanacearum* in highland Uganda: implications to development of an integrated control package for bacterial wilt. Elphinstone J. *Bacterial Wilt Disease*. Elphinstone J. Berlin, Heidelberg; 1998.
90. Ciampi L, Sequeira L, French ER. Latent infection of potato tubers by *Pseudomonas solanacearum*. *Am Potato J.* 1980;57: 377–386. doi:10.1007/BF02854329

Article 2

The epidemiology of the Malagasy *Ralstonia pseudosolanacearum* phylotype I genetic lineages might be impacted by bacteriocin production

Hasina Rasoamanana¹, Santatra Ravelomanantsoa², Nomenjanahary Marie-Véronique¹, Miharisoa-Mirana Gauche¹, Philippe Prior⁴, Fabien Guérin¹, Isabelle Robene³, Yann Pecrix³, Stéphane Poussier¹

¹ University of Reunion Island, UMR PVBMT, 97410 Saint-Pierre, Reunion Island, France

² CENRADERU/FOFIFA, 101 Antananarivo, Madagascar

³ CIRAD, UMR PVBMT, 97410 Saint-Pierre, Reunion Island, France,

⁴ INRAE, UMR PVBMT, 97410 Saint-Pierre, Reunion Island, France

Abstract

Bacterial wilt caused by the *Ralstonia solanacearum* species complex (RSSC) is a major threat to the vegetable crops in Madagascar. To effectively manage the disease, various surveys were carried out in the main vegetable production areas of the country leading to the collection of 401 new RSSC isolates. Phylogenetic assignment of the isolates revealed a high prevalence of phylotype I sequevar 18 in Madagascar. This result highly contrasts with the epidemiological pattern of RSSC in its neighboring islands as phylotype I sequevar 31 is widely distributed in La Réunion, Comoros, Mayotte, Mauritius, Rodrigues and Seychelles. Molecular typing characterization of the malagasy isolates allowed the identification of 97 haplotypes. Some of them are shared by various plots, sometimes even located in different provinces, which shows their probable dissemination by infected plant materials. To explain the observed epidemiological pattern, we have explored the bacteriocin production ability of the malagasy strains. Interestingly, the highly prevalent genetic lineages I-18 produces bacteriocins active against all the genetic lineages present

in the country. This work sheds light on the potential impact of bacteriocins on the epidemiology of malagasy RSSC.

Introduction

Regularly exposed to pathogenic microbes, plants may develop infectious diseases that spread rapidly over a given area under favorable environmental conditions and lead to an epidemic (Rapilly, 1991). To develop relevant control strategies against plant pathogens, knowledge of the genetic diversity and the dynamic structure of their populations is a prerequisite (Pruvost *et al.*, 2014; Poulin *et al.*, 2015; Yahiaoui *et al.*, 2017; Ravelomanantsoa *et al.*, 2018; Nakato *et al.*, 2019; Rasoamanana *et al.*, 2020). A fine identification of the genetic lineages responsible for an epidemic is fundamental to manage plant disease. The phenotypic methods based on the characterization of strains according to their expressed traits have long been used. For example, bacterial strains were differentiated according to the morphology of their colonies on medium culture, biochemical and serological tests, sensitivity or resistance to antibiotics (Li *et al.*, 2009). However, these old genotyping methods present the main inconvenience of not being sufficiently discriminative between related strains (Li *et al.*, 2009). Thereafter, the development of molecular tools based on markers that are stable, highly discriminative, and reproducible revolutionized the genotyping methods (Picard, 1999). MultiLocus Variable Number of Tandem Repeat (VNTR) Analysis (MLVA) is among the genotyping tools that have proven to be valuable for outbreak analysis and epidemiological surveillance of phytopathogenic bacteria such as: *Clavibacter michiganensis* subsp. *michiganensis* (Zaluga *et al.*, 2013), *Pseudomonas syringae* pv. *tomato* and *maculicola* (Gironde and Manceau, 2012), *Xylella fastidiosa* (Francisco *et al.*, 2017), *Xanthomonas citri* pv. *citri* (Pruvost *et al.*, 2014), *Xanthomonas vasicola* pv. *musacearum* (Nakato *et al.*, 2019) and the *Ralstonia solanacearum* species complex (Ravelomanantsoa *et al.*, 2016; Rasoamanana *et al.*, 2020).

However, a fine description of the epidemiological situation is not sufficient to manage an outbreak. Indeed, understanding the factors and mechanisms promoting the epidemic is also important to deploy relevant disease control strategies. It is assumed that a genetic lineage with a high ability to survive and proliferate in a competitive environment is likely

to thrive in its niche. Several weapons are used by bacteria to prosper in their environment like the secretion of antimicrobial compounds called bacteriocins which are ribosomally synthesized proteinaceous compounds active against closely related strains (Drider and Rebuffat, 2011). Several studies had highlighted the influence of bacteriocins in the competitiveness and prevalence of a genetic lineage over others (Bradley, 1967; Smidt and Vidaver, 1982; Goel *et al.*, 1999; Tudor-Nelson *et al.*, 2003; Hert *et al.*, 2005). For instance, in Florida, it was demonstrated that race T3 of *Xanthomonas euvesicatoria* supplanted race T1 in tomato fields due to its ability to produce bacteriocins (Hert *et al.*, 2005).

The *Ralstonia solanacearum* species complex (RSSC) comprises globally distributed plant pathogenic bacteria that causes bacterial wilt (BW) on nearly 54 botanical families (Denny, 2006). This species complex is composed of three species that can be further subdivided into four phylotypes according to the geographical origin of the strains. *R. pseudosolanacearum* comprises phylotypes I and III originating from Asia and Africa, respectively; *R. solanacearum* includes phylotypes IIA and IIB from America; and *R. syzygii* includes the phylotype IV originating from Indonesia, Japan and Australia (Fegan and Prior, 2005; Safni *et al.*, 2014; Prior *et al.*, 2016). Each phylotype can be subdivided into sequevars that group isolates whose partial endoglucanase (*egl*) gene sequence diverges at less than 1% (Fegan and Prior, 2005). BW is considered as one of the most devastating diseases on solanaceous crops in the South-West Indian Ocean (SWIO) islands (Madagascar, La Réunion, Mauritius, Rodrigues, Mayotte, Comoros, and Seychelles). Knowing the genetic diversity of the RSSC populations is a first step in the development of adapted BW control strategies in the SWIO. From 2014 to 2016, a wide sampling survey has been performed in the small SWIO islands (Comoros, Mauritius, La Réunion, Rodrigues, and Seychelles). Genetic diversity analysis and phylogenetic assignment of the isolates highlighted the prevalence of the *R. pseudosolanacearum* phylotype I sequevar 31 in these islands (Yahiaoui *et al.*, 2017). In 2020, a newly developed MLVA scheme, specific to phylotype I (RS1-MLVA14), allowed the genotyping of 148 Malagasy isolates from the central highlands (1,000-1,600 m) and a region located in the lowlands (6-50 m). Phylogenetic assignment of the isolates revealed the prevalence of phylotype I sequevar 18 and genetic diversity analysis showed a significant contrasting level of diversity according to elevation as strains were more diverse at low than at high

altitudes (Rasoamanana *et al.*, 2020). However, a more exhaustive study covering diverse crop production zones located at middle and low altitudes appeared necessary to analyze more exhaustively the genetic diversity and structure of RSSC populations in Madagascar.

Bacteriocin production by RSSC strains was first reported by Okabe (Okabe, 1954). Several studies partially characterized the bacteriocins as non sedimentable and belonging to a low molecular weight group (Cuppels *et al.*, 1978; He *et al.*, 1988) or on the contrary sedimentable, thermolabile, trypsin resistant and belonging to a high molecular weight group (Arwiyanto *et al.*, 1993). In 1996, bacteriocin profiles of French indies RSSC strains revealed a complete correlation with their group membership (Frey *et al.*, 1996). In addition, bacteriocin activity was also demonstrated *in planta*, which could confer a selective advantage for the growth of producing strains (Frey *et al.*, 1996). In 2015, the influence of bacteriocin production on the geographical distribution of phylotype IIB sequevar 1 strains (IIB-1), represented by strain UW551, was suggested. Indeed, the authors demonstrated that the tropical strain GMI1000 (I-18) and the subtropical strain K60 (IIA-7) produce bacteriocins that inhibit the growth of UW551 at warm temperatures (Huerta *et al.*, 2015). In Madagascar, no study has yet been done on the bacteriocin-producing ability of RSSC.

In this work, we first analyzed the genetic diversity and population structure of 401 RSSC isolates newly collected from various crop productions in Madagascar. Then, we tested for the first time the bacteriocin activity of isolates representative of the RSSC genetic lineages found in Madagascar to assess if the bacteriocin production might influence the observed epidemiological patterns.

Material and methods

1. Bacterial collections and phylogenetic assignment

Three collections of RSSC isolates were used in this study. All isolates were collected from portions of stem of symptomatic and asymptomatic plants belonging to Solanaceae (n = 601), Geraniaceae (n = 8), Fabaceae (n = 7), Araceae (n = 5), Cucurbitaceae (n = 5), Asteraceae (n = 4), Amaranthaceae (n = 4), Mimosoideae (n = 2), Musaceae (n = 2), Brassicaceae (n = 2), Rosaceae (n = 2), Zingiberaceae (n = 1), Begoniaceae (n = 1),

Apiaceae (n = 1), Casuarinaceae (n = 1), Vitaceae (n = 1), and non reported plant host (n = 11) (Table S1 +S2).

The two first collections (C1 + C2) (Table S1) were collected from three provinces of Madagascar representing various crop production zones from sea level up to 1503 m (Fig. 1). The difference in elevation provided the opportunity to study the influence of altitude on the genetic diversity and structure of RSSC populations. The first collection, C1, was composed of 148 phylotype I isolates isolated in 2006 and 2013, from the provinces of Toamasina and Antananarivo (Rasoamanana *et al.*, 2020). The province of Toamasina is represented by the region of Atsinanana (n = 25, 4 plots) located in the lowlands (mean elevation = 28.31 m). The province of Antananarivo is represented by the regions of Analamanga (n = 28, 4 plots) and Itasy (n = 95, 8 plots) located in the central highlands (mean elevation = 1057.23 m). The second collection, C2, comprises 401 isolates collected in 2018 and 2019 during this study. These isolates originated from the provinces of Toamasina, Mahajanga and Antananarivo. The province of Toamasina is represented by the regions of Atsinanana (n = 71, 7 plots) located in the lowlands (mean elevation = 12.61 m), Analanjirofo (n = 9, 4 plots) located in the lowlands (mean elevation = 158 m), and Alaotra-Mangoro (n = 43, 7 plots) located at mid-elevation (mean elevation = 779.42m). The province of Mahajanga is located in the lowlands and is represented by the regions of Betsiboka (n = 10, 1 plot, elevation = 44 m), Boeny (n = 30, 3 plots, elevation = 12 m), and Sofia (n = 212, 7 plots, mean elevation = 41.2m). The province of Antananarivo is located in the central highlands of Madagascar. It is represented by the regions of Analamanga (n = 12, 3 plots, mean elevation = 1294.65 m), Bongolava (n = 7, 2 plots, mean elevation = 905.14 m), and Itasy (n = 5, 1 plot, mean elevation = 1503 m). Two isolates are from non reported regions. In order to identify the genetic relationships between the Malagasy and international phylotype I isolates, a third collection, C3, corresponding to 165 worldwide isolates from the small islands of the SWIO (n = 83), Africa (n = 30), Asia (n = 17), America (n = 16), West-Indies (n = 12), Pacific Islands (n = 6), and Australia (n = 1) was added to the analysis (Table S2).

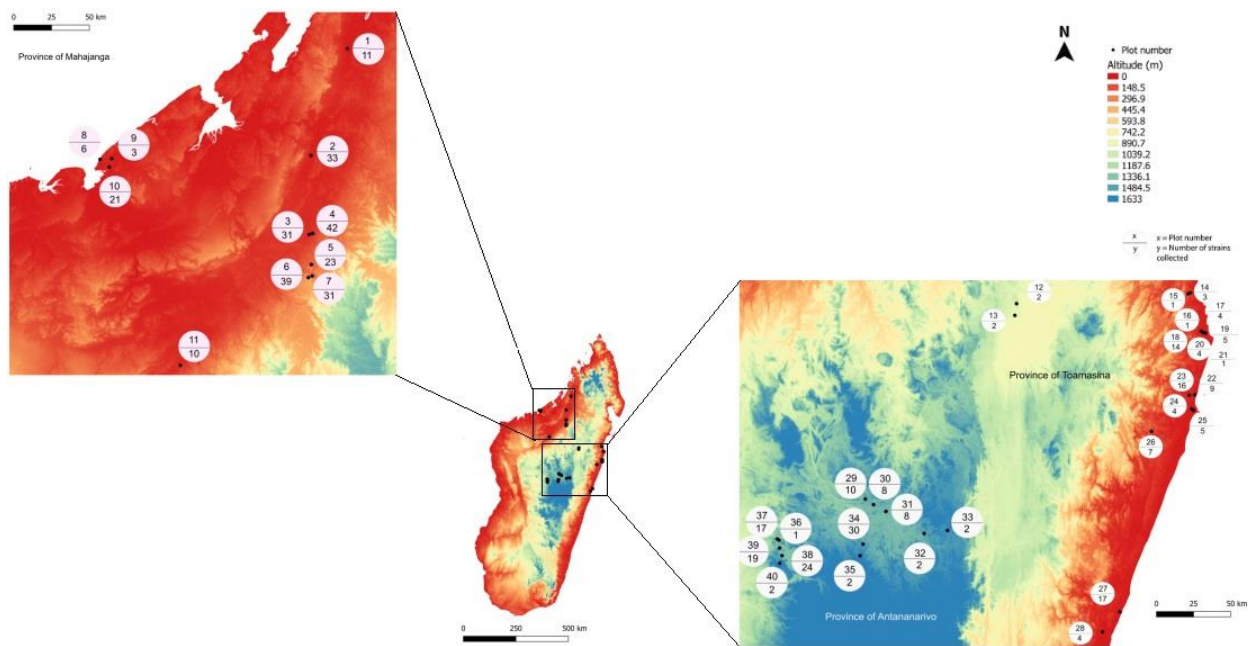


Fig. 1: Collection area of RSSC phylotype I in Madagascar.

Each black dot indicates the site of collection, the white circle contains information about the plot number and the number of isolates collected in each site.

Phylogenetic assignments were performed to verify that the isolates belonged to RSSC and to determine their phylogenetic groups. Firstly, multiplex PCR amplification (Fegan and Prior, 2005) was performed in 25 μ l reaction volumes containing 1X Green GoTaq® Reaction Buffer, 1.5 mM MgCl₂, 0.2 mM dNTP, 2.4 mM of each primers Nmult21:1F specific to phylotype I - Nmult21:2F specific to phylotype II - Nmult22:InF specific to phylotype IV - Nmult22:RR as a reverse primer, 7.2 mM of Nmult23:AF primer specific to phylotype III, 1.6 mM of each primers 759F/760R specific to RSSC, 1.25U μ l GoTaq flexi DNA polymerase (Promega), 2 μ l of DNA suspension and 10.88 μ l sterile HPLC-grade water. PCR was carried out in a Verity technology thermocycler (Applied Biosystems thermal cyclers) under the following conditions: an initial denaturation step at 96°C for 5 min, 30 cycles of denaturation at 94°C for 15s, annealing at 59°C for 15s, extension at 72°C for 30s, and a final extension step at 72°C for 10 min. Then, 10 μ l of PCR product was loaded into a 1.5% (w/v) Sea-Kem1 LE Agarose (Lonza, Basel, Switzerland) gel for electrophoresis. Ethidium bromide was used to stain the gels and the G:BOX gel imaging system (Syngene, Cambridge, UK) enabled the visualization of the bands under ultraviolet. The molecular weights were estimated by comparison with a 100 bp DNA ladder (Promega, Madison, Wisconsin, USA). Secondly, the partial endoglucanase (*egl*)

gene of RSCC isolates was amplified in 50 µl reaction volumes containing 1X Green GoTaq® Reaction Buffer, 1.5 mM MgCl₂, 0.2 mM dNTP, 0.25 µM of each primer EndoF and EndoR, 1.25U µl GoTaq flexi DNA polymerase (Promega), 2 µl of DNA suspension and 31.25 µl sterile HPLC-grade water. PCR was performed in a Verity technology thermocycler (Applied Biosystems thermal cyclers) under the following conditions: an initial denaturation step at 96°C for 9 min, 30 cycles of denaturation at 95°C for 1 min, annealing at 70°C for 1min 30s, extension at 72°C for 2 min, and a final extension step at 72°C for 10 min. Afterwards, 5 µL of PCR product was separated on 1% agarose gel to visualize the amplification. Positive PCR products were sent to Genewiz (Leipzig, Germany) for DNA double-stranded sequencing (forward and reverse). Forward and reverse chromatograms were analyzed using the Geneious v10.0.9 software (Kearse *et al.*, 2012). Isolates whose *egl*/sequence diverge less than 1% were grouped to form a sequevar (Fegan and Prior, 2005).

2. MLVA typing

The objective of the MLVA typing was to determine the allelic profile, named haplotype or Multilocus Type (MT), of the isolates allowing their discrimination at a fine level. This was accomplished by applying the RS1-MLVA14 scheme on phylotype I isolates following previously described protocol (Rasoamanana *et al.*, 2020). Briefly, a multiplex PCR amplification was carried out in a 15 µl reaction volumes containing 7.5 µl Terra PCR Direct Buffer 2X (Terra™ PCR Direct Polymerase Kit, Clontech Laboratories, Inc.), 0.3 µl Terra PCR Direct Polymerase Mix—1.25 U/µl, 1.5 µl 5X Q-solution (Qiagen1, Hilden, Germany), 2.3 µl of a forward and reverse primer mix (2 µM each), 2.4 µl sterile HPLC-grade water, and 1 µl of bacterial suspension as a template. PCR was performed in a Verity technology thermocycler (Applied Biosystems thermal cyclers) under the following conditions: an initial denaturation step at 98°C for 2 min, 25 cycles of denaturation at 98°C for 10s, annealing at 62°C for 15s, extension at 68°C for 1 min, and a final extension step at 68°C for 30 min. Then, PCR products were diluted (at least 1:80) to avoid peak saturation. In a 2 ml Eppendorf tube, 1080 µl formamide (Hi-Di™ formamide, Applied Biosystems) was mixed with 20 µl size marker (GeneScan™-500 LIZ1 Size Standard, Applied Biosystems). In each well of a 96-well plate, 11 µl was distributed with 1 µl of diluted PCR product. The samples were denatured at 95°C for 5 min, cooled immediately

on ice and loaded onto an ABI Prism 3130XL Genetic Analyzer for capillary electrophoresis.

3. Genetic structure and diversity analysis

The genotyping results were visualized and analysed with Geneious v10.0.9 (Kearse *et al.*, 2012). They were presented as electrophoregrams (peaks) of different colors. Each color corresponds to the fragment size of a VNTR locus that is estimated using the third order least squares algorithm implemented in Geneious v10.0.9, then a bin size is attributed. The bin size takes into account small size variation due to experimental variation. When a VNTR array was truncated, the VNTR number was rounded up to the nearest bin (Pourcel and Vergnaud, 2011; Pruvost *et al.*, 2014; Nakato *et al.*, 2019). Besides the number of repeats of the VNTR was calculated using the following formula:

$$\text{Number of repeats} = \frac{\text{Bin size} - \text{Forward and reverse primer size} - \text{Flanking region size}}{\text{VNTR size}}$$

The genetic profile of each isolate, called Haplotype or Multilocus Type (MT), was defined based on the number of repeats at each VNTR loci.

The genetic diversity of the Malagasy phylotype I isolates belonging to the C1 and C2 collections were analyzed according to their province and host of collect. Arlequin v3.5.2.2 (Excoffier and Lischer, 2010) was used to compute Nei's unbiased estimates of genetic diversity (Hnb). The allelic richness was estimated using HP-RARE (Kalinowski, 2005). A standardized measure of the genetic variance among the populations, Fst and Rst, were also estimated using Arlequin v3.5.2.2 to provide information on the possible gene flow between the Malagasy phylotype I populations.

The presence or absence of genetic structure among the Malagasy phylotype I isolates (C1 + C2) were estimated without a priori by computing multidimensional scaling (MDS) with the R package *Bios2mds* (Pelé *et al.*, 2012). MDS is a method that represents measurements of similarity (or dissimilarity) among pairs of objects as distances between points of low-dimensional or multidimensional space (Borg and Groenen, 2005). Afterwards, the level of differentiation between Malagasy phylotype I isolates (C1 + C2) were represented by minimum spanning trees (MSTs) according to their province of collect. Moreover, MST was also built with the Malagasy (C1 + C2) and worldwide (C3) phylotype I isolates to represent their genetic relationships. All the MSTs were built using

global optimal eBURST (goeBURST) and Euclidean distances implemented in PhyloViz (<http://www.phyloviz.net/>). A clonal complex (CC) was defined as a cluster of haplotypes, in which all haplotypes are linked as Single Locus Variants (SLVs) to at least another haplotype. In this model, the founder haplotype first increases in number before mutation generates SLVs (Vernière *et al.*, 2014).

4. Inhibition activity assays

All the isolates used in this study are representative of the phylotype I genetic lineages encountered in Madagascar, from the C1+C2 collections. Direct antagonism of 18 isolates were tested on 27 indicator isolates using an overlay plate growth inhibition assay modified from Gross and Vidaver, 1990 and Frey *et al.*, 1994. A log-phase aliquot (25 µl) of each indicator isolate grown in nutrient broth (approximately 10⁹ CFU/ml) was added to 4.5 ml of molten Kelman soft agar (0.75% [wt/vol] agar) cooled to 45°C and poured over a 2,3,5-triphenyltetrazolium chloride-agar plate. The soft agar layer was allowed to solidify, and 10 µl of filtrate (filter pore size, 0.2 µm) of the culture supernatant of each producer isolate grown in nutrient broth for 24 h was spotted onto the plate. Plates were incubated at 28°C for 48 h to allow observation of inhibition zones. The experiment was repeated twice to verify the repeatability of the results.

5. Characterization of inhibitory activity in culture supernatants

Supernatants from the 18 producer isolates used in growth inhibition activity assays were characterized by treating 100 µl at 100°C for 10 min. Fifty-microliter of proteinase K at a final concentration of 50 µg/ml was also added to 50 µl of supernatants and incubated at 37°C for 2 h. After treatment, they were tested for inhibitory activity against an indicator strain. To differentiate inhibition activity from bacteriophage or bacteriocin, supernatants were serially diluted in water and spotted onto soft agar plates following the formerly described protocol. Performing a serial dilution on a supernatant containing bacteriophage will result in individual plaques becoming less in number with greater dilution, whereas serial dilution of a supernatant containing bacteriocin will result in a clearing zone that becomes uniformly more turbid with greater dilution (Hockett and Baltrus, 2017). Inhibition activity of the supernatants were also tested on 19 strains belonging to bacterial species more or less closely related to RSSC: *Ralstonia eutropha*, *Ralstonia pickettii*, *Pseudomonas savastanoi*, *Pseudomonas fluorescens*, *Pseudomonas*

savastanoi pv. *savastanoi*, *Pseudomonas cichorii*, *Pseudomonas corrugata*, *Pseudomonas syringae* pv. *tomato*, *Pseudomonas putida*, *Erwinia chrysanthemi*, *Burkholderia andropogonis*, *Xanthomonas axonopodis* pv. *dieffenbachiae*, *Xanthomonas vesicatoria*, *Xanthomonas citri* pv. *citri*, *Xanthomonas gardneri*, *Xanthomonas euvesicatoria*, *Xanthomonas perforans*, *Clavibacter michiganensis*, *Escherichia coli*.

6. Data analysis

The bacterial lawn in the center of the Petri dish was manually removed to serve as a positive control in the calculation of inhibitory activity. Data was obtained by placing the plates on a luminous source before photographing them. Image analysis was performed using imageJ software (<https://imagej.nih.gov/>). The light intensity through the inhibition zones, the bacterial lawn and the positive control were compared using the following formula:

$$\text{Inhibitory activity} = \frac{\text{Light intensity through the halo} - \text{Light intensity through the bacterial lawn}}{\text{Light intensity through the positive control} - \text{Light intensity through the bacterial lawn}}$$

Two data sets were obtained from the two biological replicates of overlay plate growth inhibition assay. Thus, to verify the repeatability of the results, they were combined and analyzed together. A dendrogram was built on R v.4.1.3 to visualize the clustering of similar data based on the Minkowski distance which is a generalization of both the Euclidean distance and the Manhattan distance. Then, a heatmap based on the sensitivity of the indicator isolates was created on iTOL v6 (<https://itol.embl.de>) by importing the dendrogram and datasets.

Results

1. Phylogenetic assignments of the Malagasy isolates

During 2018 and 2019, wide samplings were conducted in 3 provinces of Madagascar: Mahajanga (3 regions, 11 plots) located at low elevation from 8 to 49 m, Toamasina (3 regions, 18 plots) located at low and mid elevations from 5.2 to 792 m, and Antananarivo (3 regions, 13 plots) located at high elevation from 902 to 1503 m. These samplings

allowed to collect 401 isolates, named C2 collection (Table S1), whose phylogenetic assignment revealed the presence of 3 phylotypes in those areas. Combined with the data from C1 collection (Rasoamanana *et al.*, 2020), the phylotype I (n = 491) represented 89.44% of the isolates. They were widely distributed in the provinces of Antananarivo (n = 125), Mahajanga (n = 252) and Toamasina (n = 112) representing 8 regions of collect, 40 plots dispersed from sea level up to 1484 m. The phylotype I was isolated from a wide range of plant hosts representing 9 botanical families: Solanaceae (n = 452), Cucurbitaceae (n = 1), Araceae (n = 2), Musaceae, (n = 1), Fabaceae (n = 3), Brassicaceae (n = 1), Asteraceae (n = 3), Mimosoideae (n = 2), and non reported plant hosts (n = 10). The phylotype II (n = 53) represented 9.65% of the isolates. They were sampled in the provinces of Antananarivo and Toamasina representing 4 regions, 11 plots, located at mid and high elevations from 764 m to 1503 m. The phylotype II was collected from Solanaceae, mostly *Solanum tuberosum* (n = 48). The phylotype III (n = 5) represented 0.91% of the isolates. They were collected in Toamasina and Antananarivo representing 2 regions, 2 plots, located in mid and high elevations from 787 m to 1240 m. The phylotype III isolates were sampled from Fabaceae (*Phaseolus vulgaris* (n = 1)) and Solanaceae (*Capsicum annuum* (n = 4)).

The analysis of the endoglucanase (*egl*) gene of 360 isolates (C1+C2) representative of the sampling provinces allowed their classification into sequevars. Four sequevars belonging to the phylotype I were identified: the sequevar 18 (n = 276) represented 76.67% of the isolates, the sequevar 46 (n = 41) represented 11.39%, the sequevar 33 (n = 5) (1.39%) and the sequevar 31 was only represented by two isolates. The sequevar 1 (n= 31) belonging to phylotype II represented 8.61% of the isolates. The sequevar 19 was represented by only one isolate and the sequevar 60 represented 1.11% of the isolates, both belonging to the phylotype III.

2. Genetic diversity of the Malagasy phylotype I isolates

A total of 475 phylotype I isolates were genotyped using the RS1-MLVA14 scheme: 148 isolates from a previous study (C1) (Rasoamanana *et al.*, 2020) and 343 in this study (C2). Four hundred forty-one isolates were genotyped with all of the 14 markers and 34 isolates could be genotyped with only 13 markers despite several assays. The latter isolates were nonetheless retained for the genetic diversity analysis. A total of 34 isolates belonging to

the sequevar 18 (n = 22), 46 (n = 5) and undetermined sequevar (n = 7) were removed from the analysis since they did not amplify on more than one locus. Two loci (GMich_3461 and GMImp_0618) of one of the two isolates belonging to sequevar 31 detected in Madagascar, RUN6340, did not amplify. As this strain is an isolate of interest (a representative of the sequevar 31 identified in Madagascar), the number of tandem repeats was determined after Sanger sequencing.

A total of 97 haplotypes were identified, with a high prevalence of two haplotypes MT004 (18.82%) and MT007 (18.37%). The overall genetic diversity of the Malagasy collections was $H_{nb} = 0.39$ varying from 0.18 to 0.48 for Antananarivo and Toamasina provinces respectively (Table 1). The province of Toamasina also has the highest allelic richness ($A = 2.67$) whereas Antananarivo has the lowest ($A = 1.83$) (Table 1). The major haplotypes encountered in Antananarivo and Mahajanga are MT004 and MT007 whereas MT014 and MT334 are the most present in Toamasina (Table 1).

According to the Malagasy provinces, three populations were defined : Mahajanga (n = 250), Antananarivo (n = 125) and Toamasina (n = 99). Genetic differentiations among the three populations are high and significant: $F_{ST} = 0.19^{***}$, $R_{ST} = 0.24^{***}$ between Mahajanga and Antananarivo; $F_{ST} = 0.28^{***}$, $R_{ST} = 0.19^{***}$ between Toamasina and Antananarivo; $F_{ST} = 0.19^{***}$, $R_{ST} = 0.29^{***}$ between Toamasina and Mahajanga (Table S3). Nei's unbiased genetic diversity (H_{nb}) ranges from 0.18 to 0.48 and the allelic richness varies from 1.83 to 2.67 corresponding respectively to Antananarivo and Toamasina.

According to the major plant hosts, we defined five populations composed of isolates collected from *Solanum lycopersicum* (n = 207), *Solanum melongena* (n = 46), *Capsicum annuum* (n = 41), *Solanum tuberosum* (n = 20), and *Solanum aethiopicum* (n = 99). Nei's unbiased genetic diversity (H_{nb}) ranges from 0.22 to 0.45 and the allelic richness varies from 1.98 to 2.57 corresponding respectively to *Solanum aethiopicum* and *Capsicum annuum* (Table 1). It is worth noting that the majority of isolates collected from *Capsicum annuum* are from the province of Toamasina (n = 38 out of 41 isolates) (Table 1). The genetic differentiation between *Solanum aethiopicum* and *Solanum lycopersicum* is low ($F_{ST} = 0.02^*$, $R_{ST} = 0.00015^*$). It is moderate but significant between *Solanum tuberosum* and *Capsicum annuum* ($F_{ST} = 0.08^{***}$, $R_{ST} = 0.007^{***}$). It is important and significant

between *Solanum tuberosum* and *Solanum lycopersicum* ($F_{ST} = 0.13^{***}$, $R_{ST} = 0.29^{***}$);
Capsicum annuum and *Solanum lycopersicum* ($F_{ST} = 0.16^{***}$, $R_{ST} = 0.26^{***}$),

Table_ 1: Genetic diversity estimated from RS1-MLVA14 data of C1 and C2 RSSC collection

Population		Elevation	n	A	G	H _{nb}		Major haplotypes		
Provinces	Toamasina	5.2 - 792m	99	2.67	42	0.39	0.48	MT014 (n = 8) MT334 (n = 6)		
	Antananarivo	1176.6 - 1484 m	125	1.83	22		0.18	MT004 (n = 23) MT007 (n = 43)		
							0.36	MT004 (n = 58) MT007 (n = 37)		
	Mahajanga	8 - 49 m	249	2.15	46					
	Hosts	<i>Solanum lycopersicum</i>		207	2.04		37	0.34	0.28	MT004 (n = 60) MT007 (n = 32)
									0.45	MT003 (n = 4) MT334 (n = 6)
<i>Solanum melongena</i>		46		2.48	20	0.38	MT236 (n = 14) MT266 (n = 4) MT249 (n = 4)			
							0.34		MT007 (n = 6) MT023 (n = 6)	
									0.22	MT007 (n = 47) MT004 (n = 18)
<i>Solanum tuberosum</i>		20		2.31	8					
<i>Solanum aethiopicum</i>		99		1.98	18					

n, number of isolates; A, allelic richness calculated by rarefaction method; G, number of haplotypes; H_{nb}, Nei's unbiased estimates of genetic diversity.

Solanum melongena and *Solanum lycopersicum* ($F_{ST} = 0.32^{***}$, $R_{ST} = 0.26^{***}$), *Solanum melongena* and *Capsicum annuum* ($F_{ST} = 0.26^{***}$, $R_{ST} = 0.12^{***}$), *Solanum aethiopicum* and *Capsicum annuum* ($F_{ST} = 0.19^{***}$, $R_{ST} = 0.16^{***}$), *Solanum tuberosum* and *Solanum melongena* ($F_{ST} = 0.34^{***}$, $R_{ST} = 0.28^{***}$), *Solanum aethiopicum* and *Solanum melongena* ($F_{ST} = 0.4^{***}$, $R_{ST} = 0.27^{***}$), *Solanum aethiopicum* and *Solanum tuberosum* ($F_{ST} = 0.16^{***}$, $R_{ST} = 0.33^{***}$) (Table S4).

3. Population structure analysis of the Malagasy phylotype I isolates

A multidimensional scale (MDS) analysis was performed to identify a possible clustering of the 97 Malagasy haplotypes. The MDS plot revealed two clusters with a probability of 52.08%. Haplotypes were assigned to these clusters with a probability of 77.57%. Axis 1 and Axis 2 explained 62.4% and 15.2% of the variance, respectively. Cluster 1 groups 72 haplotypes while cluster 2 comprises 25 haplotypes (Fig. 2).

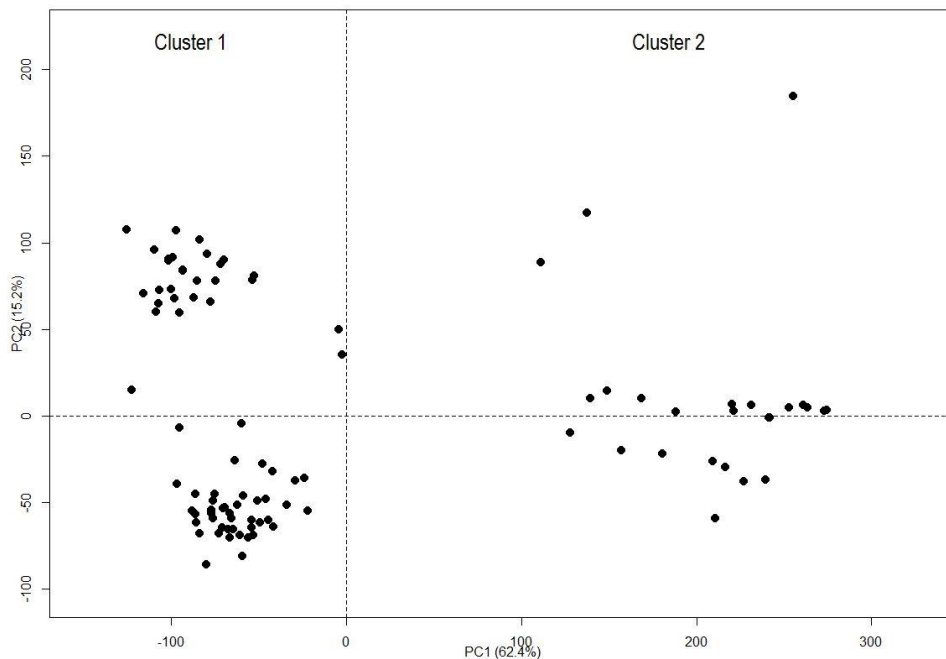


Fig. 2: Multidimensional scaling (MDS) of the 97 Malagasy haplotypes

Cluster 1 and 2 represent the two groups of haplotypes.

A minimum spanning tree (MST) was built to visualize the genetic links between the Malagasy haplotypes. According to the structure of the MST, 12 Malagasy clonal complexes (MCC) were identified (Fig. 3). The major clonal complex MCC1 ($n = 270$) regroups 28 haplotypes belonging to sequevar 18, sampled from the three provinces of

Antananarivo, Mahajanga and Toamasina. The second major clonal complex (MCC2) (n = 83) regroups 19 haplotypes belonging to sequevar 18, sampled from the province of Mahajanga. The third major clonal complex (MCC3) (n = 18) regroups 6 haplotypes belonging to sequevar 46, sampled from the province of Toamasina. The 9 remaining clonal complexes (MCC4 to MCC12) are minors, regrouping two to three strains sampled from the three provinces. We observed that the cluster 1 defined by the MDS analysis, regrouping the majority of strains (n = 373, G = 72), is composed of the MCC1, MCC2, MCC5, MCC6, MCC7 that belonged to sequevar 18 except an isolate from MCC6 that was identified as sequevar 46. The cluster 2, regrouping 120 isolates and 25 haplotypes, is composed of MCC3, MCC4, MCC9 and MCC10 which belonged to sequevar 46, MCC11 belonging to sequevar 31, and MCC12 belonging to sequevar 33.

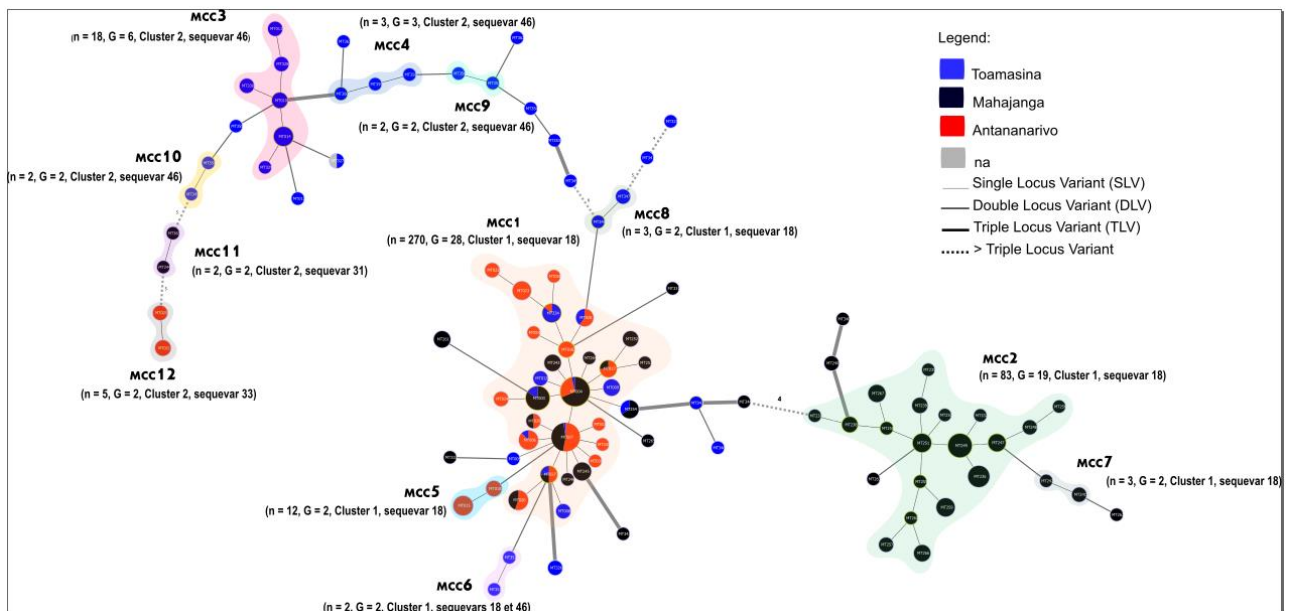


Fig. 3. Minimum spanning tree (MST) of the Malagasy phylotype I according to the province of collect

The haplotypes were identified by using goeBURST full MST in PHYLOVIZ. Each MLVA type (MT) is displayed as a circle, the size of which is proportional to the number of isolates represented. The different colors indicate the sampling province. The branch thickness depends on the number of locus differences between the neighboring haplotypes. MCC1 to MCC12 represent the Malagasy clonal complexes 1 to 12. A clonal complex is composed of haplotypes that differ only by one VNTR locus. n represents the number of strains, G represents the number of haplotypes, the cluster number corresponds to the MDS result.

4. Genetic links between the Malagasy and worldwide phylotype I isolates

A global MST was built with the 475 Malagasy isolates and the 165 worldwide isolates from the Indian Ocean (n = 83), Africa (n = 30), America (n = 16), Asia (n = 17), Australia (n = 1), Pacific Islands (n = 6), and West Indies (n = 12). The worldwide isolates that are close to the Malagasy strains were represented in Fig.4.

Surprisingly, the main Malagasy clonal complex MCC1 (n = 270) seems specific to Madagascar, only the major haplotype MT004 is shared with an isolate sampled in South Africa. The MCC2 forms a clonal complex with MT080 from Mayotte and MCC3 forms a clonal complex with MT074 from Ivory Coast and MT081 from Mayotte. Unsurprisingly, the MCC11 (including only two sequevar 31 isolates) is close (double locus variant) to a clonal complex combining sequevar 31 isolates, sampled from the small islands of the South-West Indian Ocean and Africa. MCC12 forms a clonal complex with haplotypes sampled from Mauritius (MT025, MT085, MT094), Reunion Island (MT085), and Rodrigues (MT098). The minor Malagasy clonal complexes (MCC4, MCC5, MCC6, MCC7, MCC9, MCC10) seem specific to Madagascar.

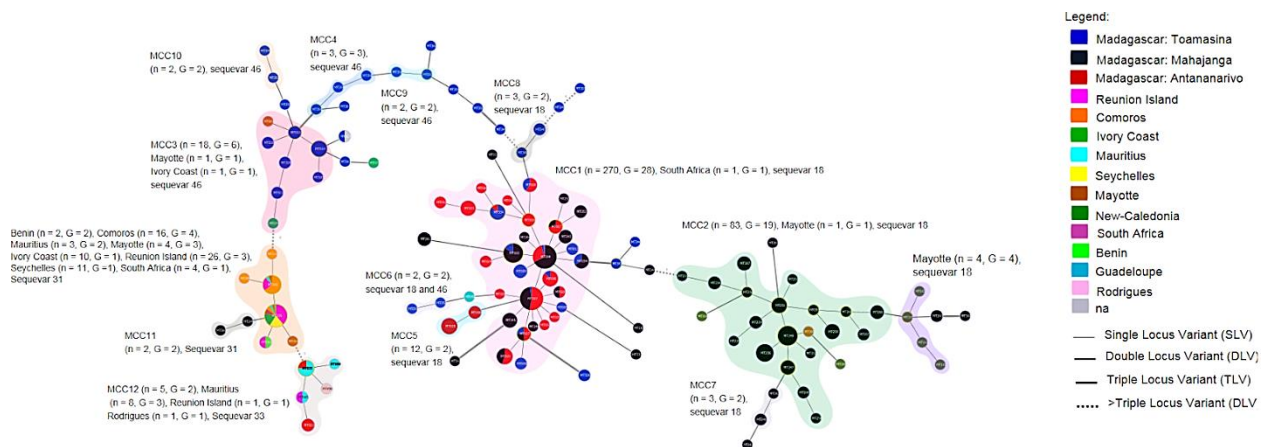


Fig. 4. Minimum spanning tree (MST) of the Malagasy and worldwide phylotype I

The haplotypes were identified by using goeBURST full MST in PHYLOVIZ. Each MLVA type (MT) is displayed as a circle, the size of which is proportional to the number of isolates represented. The different colors indicate the sampling country. The branch thickness depends on the number of locus differences between the neighboring haplotypes. MCC1 to MCC12 represent the Malagasy clonal complexes 1 to 12. A clonal complex is composed of haplotypes that differ only by one VNTR locus. n represents the number of strains, G represents the number of haplotypes.

5. Growth inhibition activity

The growth inhibition activity of 18 isolates representative of the 4 genetic lineages encountered in Madagascar was tested: sequevar 18 (n = 11), 31 (n = 1), 33 (n = 2), 46 (n = 4) against 27 indicator isolates representative of the sequevar 18 (n = 19), 31 (n = 2), 33 (n = 2), and 46 (n = 4).

According to their patterns of growth inhibition activity, isolates from the sequevar 18 could be classified into two clusters (Fig.4). The first cluster regroups five isolates belonging to the second major Malagasy clonal complex (MCC2, n = 3) and its closely related clonal complex (MCC7, n = 2). They were collected in the province of Mahajanga, regions of Boeny (n = 3) and Sofia (n = 2) and they produce an antimicrobial substance highly active against the two sequevar 31 isolates identified in Madagascar, collected in the province of Mahajanga, Betsiboka (n = RUN6285) and Sofia (RUN6340) regions. These sequevar 18 supernatants are also moderately active against strains belonging to the sequevars 33 and 46 and they are even weakly active against other sequevar 18 isolates. The second cluster regroups six isolates belonging to the major Malagasy clonal complex (MCC1, n = 4), its closely related clonal complex (MCC5, n = 2) and an isolate belonging to the MCC2 (RUN3130). They were collected in the provinces of Antananarivo (Itasy and Analamanga regions, n = 4), Toamasina (Alaotra-Mangoro region, n = 1), and Mahajanga (Sofia region, n = 1). Their supernatants are moderately to weakly active against the sequevar 31 isolates and a sequevar 33 isolate. Regarding the remaining sequevars, supernatants from sequevars 33 and 46 present no growth inhibition activity against all the isolates tested, whereas supernatant from the sequevar 31 (RUN6340 strain) is moderately to weakly active against a wide range of isolates belonging to the sequevars 18, 33, and 46 (Fig.5).

Partial characterization of all the supernatants revealed that their inhibitory activities are sensitive to high temperature (100°C, 10 mn), degraded by proteinase K at a final concentration of 50 µg/ml except for those from the MCC2 (n = 3) and MCC7 (n = 2), and only active against RSSC isolates. Their serial dilution resulted in a clearing zone becoming uniformly more turbid with greater dilution, which is characteristic of supernatants containing bacteriocins except for the RUN3130, which resulted in individual plaques probably due to bacteriophages.

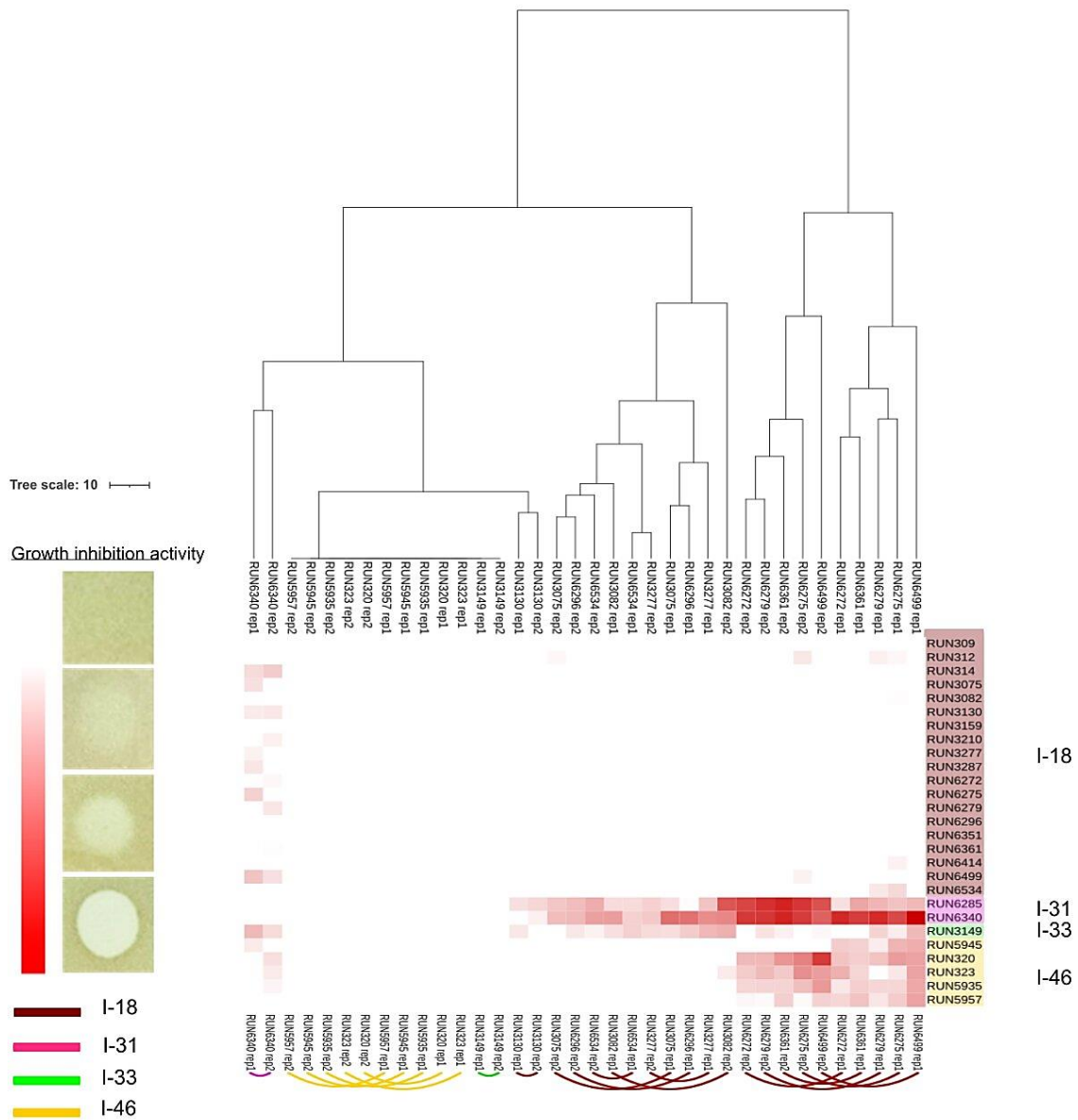


Fig. 5. Inhibitory activity of 18 producing isolates (horizontal) on 27 indicator isolates (vertical).

The dendrogram was constructed based on the similarity/dissimilarity of the inhibitory activity patterns of the 18 producing isolates. Links between the different patterns are colored according to the sequevar of the isolates (I-18, I-31, I-33 and I-46). The heatmap was based on the sensitivity of the indicator isolates to the antimicrobial compounds. The indicator isolates (vertical) were grouped based on their sequevar.

Discussion

1. High prevalence of the phylotype I sequevar 18 in Madagascar

The genetic analysis of Malagasy RSSC isolates collected in 2018 and 2019 globally confirmed the analysis of Malagasy RSSC isolates collected in 2006-2013 (C1) (Rasoamanana *et al.*, 2020), but also provides new information. This work extended to 6 new regions located in the provinces of Antananarivo, Toamasina and Mahajanga revealed that RSSC is well established in Madagascar with the presence of three phylotypes (I, II and III), collected from sea level up to 1503m. We observed a high prevalence of phylotype I in the country, representing 89.44% of the isolates. Usually encountered in tropical and subtropical zones, the Malagasy phylotype I isolates showed greater genetic diversity at lowlands, corresponding to the provinces of Mahajanga ($H_{nb} = 0.36$) and Toamasina ($H_{nb} = 0.48$) (mean elevation = 44.22 m). These provinces are also characterized by the presence of two seaports. The seaport of Mahajanga is mainly used for local traffic on Madagascar's west coast and its neighboring islands, mostly the Comoros and Mayotte. The seaport of Toamasina is the nation's main port and is used for national and international traffic. It is well-known that trade of plant products contributed greatly to the dispersion of their pests and diseases (Anderson *et al.*, 2004). Thus we can assume that they might be points of entry/exit for contaminated plant products and contribute to the great genetic diversity in these provinces and the dispersion of RSSC strains to other provinces. Interestingly, several phylotype I isolates were present in the highlands, corresponding to the province of Antananarivo (mean elevation = 1268.59 m). This is remarkable as to date, the phylotype I strains have not been reported to be well adapted to cool temperatures (Rasoamanana *et al.*, 2020).

Sixty-seven new haplotypes were characterized during this study. They are distributed among 12 clonal complexes, eight of which were newly discovered. Compared to the previous work (Rasoamanana *et al.*, 2020) a second major Malagasy clonal complex MCC2 that regroups 83 of the strains collected was identified. They were all from the province of Mahajanga and belong to sequevar 18. Moreover, our study of the geographical distribution of the strains revealed that 10 haplotypes belonging to MCC1 were shared between 21 different plots located in the province of Antananarivo (8 plots), Toamasina

(6 plots), and Mahajanga (7 plots). Eleven MCC2 haplotypes were distributed in 9 plots located in Mahajanga province. Two MCC3 haplotypes were distributed in 5 plots located in the province of Toamasina. These observations strongly suggest transmission of these haplotypes from one plot to another via contaminated plant material. However, the genetic differentiation indices F_{st} and R_{st} among collection provinces are high and significant suggesting low genetic flows between them.

Our analysis showed genetic links between the Malagasy and worldwide strains from the small islands of the SWIO and Africa. These are the case for the major malagasy clonal complexes: MCC1 ($n = 270$, $G = 28$, sequevar 18), MCC2 ($n = 83$, $G = 19$, sequevar 18), MCC3 ($n = 18$, $G = 6$, sequevar 46), and the minor malagasy clonal complex MCC12 ($n = 5$, $G = 2$, sequevar 33). The MCC1 regroups strains collected from the provinces of Toamasina, Mahajanga, and Antananarivo. Its founder haplotype, MT004, is shared with a strain from South Africa and MCC1 only differs by two locus from an haplotype sampled in Mauritius (MT059, sequevar 18). The MCC2 regroups strains collected from Mahajanga. It forms a clonal complex with an haplotype from Mayotte (MT080, sequevar 18). In addition, the MCC2 differs only by two locus to a clonal complex from Mayotte ($n = 4$, $G = 4$, sequevar 18) and two singleton from New-Caledonia (MT104 et MT160). The MCC3 consists of haplotypes from Toamasina which share the same clonal complex as an haplotype from Mayotte (MT081) and Ivory Coast (MT074). The MCC12 regroups two haplotypes from the province of Antananarivo. It shares the same clonal complex as three haplotypes from Mauritius including MT025 which is present both in Antananarivo and Mauritius, one haplotype from La Reunion, and one haplotype from Rodrigues. All those genetic links strongly suggest exchanges of infected plant materials between Madagascar, the SWIO islands and Africa (www.tresor.economie.gouv.fr/se/madagascar/, www.lexpress.mu/) (Mathieu and Zaïdani, 2002; Rabemanambola *et al.*, 2009; Fournet-Guérin and Vacchiani-Marcuzzo, 2013).

Phylogenetic assignment based on the partial *egl* gene sequence allowed the identification of four phylotype I sequevars. The sequevar 18 was the most widespread (76.67%), followed by the sequevars 46 (11.39%), 33 (1.39%) and 31 (only two isolates). This result is surprising as the sequevar 31 is the most prevalent in the small SWIO islands (Yahiaoui *et al.*, 2017). The two isolates detected in Madagascar are from the province of Mahajanga, probably introduced via the seaport as it has an intense traffic with the

neighboring islands, mostly the Comoros and Mayotte where the sequevar 31 was reported to be highly prevalent (Yahiaoui *et al.*, 2017); Chesneau *et al.*, 2018).

2. Broad spectrum inhibition activity of the sequevars 18 and 31

Most microorganisms face a constant battle for space and resources (Hibbing *et al.*, 2010). To thrive in their environment, bacteria deploy a wide range of mechanisms that harm and kill their competitors (Granato *et al.*, 2019). One of these mechanisms corresponds to the production of bacteriocins, a ribosomally protein compound with antimicrobial activity (Riley and Chavan, 2007). Direct antagonism of strains representative of the four sequevars identified in Madagascar revealed a broad spectrum inhibition activity of sequevar 18 strains, the most prevalent genetic lineage in Madagascar, and more particularly of a group of strains belonging to sequevar 18 (MCC2, n = 4, and its closely related clonal complex MCC7, n = 2) Partial characterization of these inhibitors showed their sensitivity to high temperature 100°C and proteinase K at a final concentration of 50µg/ml except for the supernatants of sequevar 18 (MCC2, n = 4 and MCC7, n = 2). Although proteinase K is a relatively unspecific proteolytic enzyme (a serine protease) (Saenger, 2013; Farrell, 2017) some bacteriocins characterized in the litterature are also insensitive to proteinase K (Berić *et al.*, 2014; Zhao *et al.*, 2015). Moreover, serial dilution of the supernatants resulted in a clearing zone becoming uniformly more turbid with greater dilution and supernatants were active on RSSC isolates and not active on other bacterial species. Combined together, these results highly suggested an inhibition activity due to bacteriocins (Gross and Vidaver, 1990; Drider and Rebuffat, 2011; Hockett and Baltrus, 2017).

Bacteriocin production and bacteriocin sensitivity vary according to the genetic lineage, and even more finely because some strains of the sequevar 18 inhibit strains from the same sequevar. The inhibitory activity assays allowed the establishment of a bacteriocin typing scheme corresponding to the status of both indicator strains (resistant or sensitive) and bacteriocin producing strains (high or low capacity). Similar bacteriocin schemes have been highlighted previously for RSSC strains (Tsai *et al.*, 1985; Arwiyanto *et al.*, 1993; Frey *et al.*, 1996). Previous studies highlighted the influence of bacteriocins in the competitiveness and prevalence of a genetic lineage over others for several bacterial species (Bradley, 1967; Smidt and Vidaver, 1982; Goel *et al.*, 1999; Tudor-Nelson *et al.*,

2003; Hert *et al.*, 2005). Our results also suggest the potential of bacteriocin production on the geographical distribution and prevalence of RSSC genetic lineages. The less prevalent genetic lineages, sequevars 33 and 46, do not show bacteriocin production. Contrarily, the most prevalent genetic lineage, the sequevar 18, has a broad spectrum inhibition activity and might benefit from this to be more competitive in its environment and be prevalent. The sequevar 31 also produces antimicrobial compounds active against the genetic lineages present in Madagascar (sequevar 18, 33, and 46). Although the sequevar 31 is widely disseminated in the small SWIO islands (Yahiaoui *et al.*, 2017), only two strains were detected in Madagascar. We assume that the strong antibiosis activity of the sequevar 18 against the sequevar 31 could only be a part of the answer as other factors may be involved in bacterial competitions such as the type VI secretion system, whereby bacterial cells inject toxins or other molecules into neighboring adversaries to promote cell lysis (MacIntyre *et al.*, 2010; Basler *et al.*, 2013; Russell *et al.*, 2014; Asolkar and Ramesh, 2020).

Another hypothesis is that sequevar 31 was only recently introduced in Madagascar and had not yet the time to spread in this country. Moreover, plants cultivated in Madagascar might also impact the prevalence of the RSSC genetic lineages. For example, a study revealed the susceptibility of two widely used potato varieties in Madagascar, 720118 (Jaingy) and 800934 (Miova), to several Malagasy sequevar 18 strains and the resistance to a sequevar 31 strain (Ravelomanantsoa, 2016).

This study set out to describe and understand the epidemiological pattern of the *Ralstonia pseudosolanacearum* phylotype I in Madagascar. We revealed that the sequevar 18 was the most prevalent among the four genetic lineages identified. The bacteriocin typing assays revealed a correlation between the inhibitory activity of the producing strains and the observed epidemiology. A natural perspective of this work would be to continue the bacteriocin typing assays to all the genetic lineages belonging to phlotypes I, IIB and III encountered in the country. Moreover, competition assays will be performed between these genetic lineages to see if the most prevalent sequevar 18 is also more competitive *in vitro* and *in planta*. Finally, further research will be undertaken by comparative genomics to look for gene content differences between the Malagasy RRSC genetic lineages, and functional studies on gene candidates that could explain the RSSC epidemiological patterns in Madagascar.

Acknowledgments

The authors wish to thank all the scientific and technical teams from CIRAD-3P (Saint Pierre, Réunion) and FOFIFA (Antananarivo, Madagascar) who took part in the realization of this work. We would also like to thank the Conseil Régional de La Réunion and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) for their financial support. We acknowledge the Plant Protection Platform (3P, IBISA).

References

1. Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R., and Daszak, P. (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends Ecol Evol* **19**: 535–544.
2. Arwiyanto, T., GOTO, M., and TAKIKAWA, Y. (1993) Characterization of Bacteriocins Produced by *Pseudomonas solanacearum*. *Jpn J Phytopathol* **59**: 114–122.
3. Asolkar, T. and Ramesh, R. (2020) The involvement of the Type Six Secretion System (T6SS) in the virulence of *Ralstonia solanacearum* on brinjal. *3 Biotech* **10**: 1–10.
4. Basler, M., Ho, B., and Mekalanos, J. (2013) Tit-for-tat: type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* **152**: 884–894.
5. Berić, T., Stanković, S., Draganić, V., Kojić, M., Lozo, J., and Fira, D. (2014) Novel antilisterial bacteriocin licheniocin 50.2 from *Bacillus licheniformis* VPS 50.2 isolated from soil sample. *J Appl Microbiol* **116**: 502–510.
6. Borg, I. and Groenen, P.J. (2005) Modern multidimensional scaling: theory and applications., Springer-Verlag New York.
7. Bradley, D.E. (1967) Ultrastructure of Bacteriophages and Bacteriocins. *Bacteriol Rev* **31**: 230–314.
8. Cuppels, D.A., Hanson, R.S., and Kelman, A. (1978) Isolation and characterization of a bacteriocin produced by *Pseudomonas solanacearum*. *J Gen Microbiol* **109**: 295–303.
9. Denny, T. (2006) Plant pathogenic *Ralstonia* species. In *Plant-Associated Bacteria*. Gnanamanickam, S.S. (ed). Dordrecht: Springer Netherlands, pp. 573–644.
10. Drider, D. and Rebuffat, S. eds. (2011) Prokaryotic antimicrobial peptides: from genes to applications, New York: Springer.

11. Excoffier, L. and Lischer, H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* **10**: 564–567.
12. Farrell, R.E. (2017) Chapter 2 - Creating a Ribonuclease-Free Environment. In *RNA Methodologies (Fifth Edition)*. Farrell, R.E. (ed). Academic Press, pp. 55–74.
13. Fegan, M. and Prior, P. (2005) How complex is the “*Ralstonia solanacearum* Species Complex.” In *Bacterial wilt disease and the Ralstonia solanacearum species complex*. Minnesota: Allen, C., Prior, P., Hayward, A.C., pp. 449–461.
14. Fournet-Guérin, C. and Vacchiani-Marcuzzo, C. (2013) Vers de nouvelles interactions entre les Suds? *Rev Tiers Monde* 141–160.
15. Francisco, C.S., Ceresini, P.C., Almeida, R.P.P., and Coletta-Filho, H.D. (2017) Spatial genetic structure of coffee associated *Xylella fastidiosa* populations indicates that cross infection does not occur with sympatric citrus orchards. *Phytopathology* **107**: 395–402.
16. Frey, P., Prior, P., Marie, C., Kotoujansky, A., Trigalet-Demery, D., and TRIGALET, A. (1994) Hrp- mutants of *Pseudomonas solanacearum* as potential biocontrol agents of tomato bacterial wilt. *Appl Environ Microbiol* **60**: 7.
17. Frey, P., Smith, J.J., Albar, L., Prior, P., Saddler, G.S., Trigalet-Demery, D., and Trigalet, A. (1996) Bacteriocin typing of *Burkholderia (Pseudomonas) solanacearum* race 1 of the French West Indies and correlation with genomic variation of the pathogen. *Appl Environ Microbiol* **62**: 7.
18. Gironde, S. and Manceau, C. (2012) Housekeeping gene sequencing and Multilocus Variable Number Tandem Repeat Analysis to identify subpopulations within *Pseudomonas syringae* pv. *maculicola* and *Pseudomonas syringae* pv. *tomato* that correlate with host specificity. *Appl Environ Microbiol* **78**: 3266–3279.
19. Goel, A.K., Sindhu, S.S., and Dadarwal, K.R. (1999) Bacteriocin-producing native rhizobia of green gram (*Vigna radiata*) having competitive advantage in nodule occupancy. *Microbiol Res* **154**: 43–48.
20. Granato, E.T., Meiller-Legrand, T.A., and Foster, K.R. (2019) The Evolution and Ecology of Bacterial Warfare. *Curr Biol* **29**: R521–R537.
21. Gross, D.C. and Vidaver, A.K. (1990) Bacteriocins. *Methods Phytobacteriology Akad Kiado Bp Hung* 245–249.
22. He, L.Y., Fan, Y.L., Lie, D.X., and Fang, F. (1988) Intraspecific antagonism and bacteriocin production of *Pseudomonas solanacearum*. *Fifth Int Congr Plant Pathog* 20–27.
23. Hert, A.P., Roberts, P.D., Momol, M.T., Minsavage, G.V., Tudor-Nelson, S.M., and Jones, J.B. (2005) Relative Importance of Bacteriocin-Like Genes in Antagonism of *Xanthomonas*

- perforans Tomato Race 3 to *Xanthomonas euvesicatoria* Tomato Race 1 Strains. *Appl Environ Microbiol* **71**: 3581–3588.
24. Hibbing, M.E., Fuqua, C., Parsek, M.R., and Peterson, S.B. (2010) Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* **8**: 15–25.
 25. Hockett, K.L. and Baltrus, D.A. (2017) Use of the Soft-agar Overlay Technique to Screen for Bacterially Produced Inhibitory Compounds. *J Vis Exp JoVE* **55064**.
 26. Huerta, A.I., Milling, A., and Allen, C. (2015) Tropical strains of *Ralstonia solanacearum* outcompete race 3 biovar 2 strains at lowland tropical temperatures. *Appl Environ Microbiol* **81**: 3542–3551.
 27. Kalinowski, S.T. (2005) hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* **5**: 187–189.
 28. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649.
 29. Li, W., Raoult, D., and Fournier, P.-E. (2009) Bacterial strain typing in the genomic era. *FEMS Microbiol Rev* **33**: 892–916.
 30. MacIntyre, D.L., Miyata, S.T., Kitaoka, M., and Pukatzki, S. (2010) The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc Natl Acad Sci* **107**: 19520–19524.
 31. Mathieu, M. and Zaïdani, M. (2002) Typologie des échanges commerciaux entre Mayotte et sa région entre 1994 et 2001. *Inst D'Emission Dép D'outre-Mer* **5**: 5.
 32. Nakato, G.V., Fuentes Rojas, J.L., Verniere, C., Blondin, L., Coutinho, T., Mahuku, G., and Wicker, E. (2019) A new Multi Locus Variable Number of Tandem Repeat Analysis scheme for epidemiological surveillance of *Xanthomonas vasicola* pv. *musacearum*, the plant pathogen causing bacterial wilt on banana and enset. *PLOS ONE* **14**: e0215090.
 33. Okabe, N. (1954) Studies on *Pseudomonas solanacearum*. V. Antagonism among the strains of *P. solanacearum*. *Rep Fac Of Agric Shizuoka Univ* **4**: 37–40.
 34. Pelé, J., Bécu, J.-M., Abdi, H., and Chabbert, M. (2012) Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling. *BMC Bioinformatics* **13**: 1–7.
 35. Picard, B. (1999) Épidémiologie moléculaire des grandes endémies. 5.
 36. Poulin, L., Grygiel, P., Magne, M., Gagnevin, L., Rodriguez-R, L.M., Forero Serna, N., et al. (2015) New Multilocus Variable Number Tandem Repeat Analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by *Xanthomonas oryzae*. *Appl Environ Microbiol* **81**: 688–698.

37. Pourcel, C. and Vergnaud, G. (2011) Strain typing using Multiple "Variable Number of Tandem Repeat" Analysis and genetic element CRISPR.
38. Prior, P., Ailloud, F., Dalsing, B.L., Remenant, B., Sanchez, B., and Allen, C. (2016) Genomic and proteomic evidence supporting the division of the plant pathogen *Ralstonia solanacearum* into three species. *BMC Genomics* **17**: 90.
39. Pruvost, O., Magne, M., Boyer, K., Leduc, A., Tourterel, C., Drevet, C., et al. (2014) A MLVA genotyping scheme for global surveillance of the Citrus Pathogen *Xanthomonas citri* pv. *citri* suggests a worldwide geographical expansion of a single genetic lineage. *PLoS ONE* **9**: e98129.
40. Rabemanambola, M., Rakotoarisoa, J., and Rieutort, L. (2009) Entre ville et campagne: les adaptations du maraîchage paysan sur les Hautes Terres centrales malgaches. *Cah D'Outre-Mer* **27**.
41. Rapilly, F. (1991) L'épidémiologie en pathologie végétale. Mycoses aériennes., Quae.
42. Rasoamanana, H., Ravelomanantsoa, S., Yahiaoui, N., Dianzinga, N., Rébert, E., Gauche, M.-M., et al. (2020) Contrasting genetic diversity and structure among Malagasy *Ralstonia pseudosolanacearum* phylotype I populations inferred from an optimized Multilocus Variable Number of Tandem Repeat Analysis scheme. *PLOS ONE* **15**: e0242846.
43. Ravelomanantsoa, S. (2016) Biologie des populations du complexe d'espèces *Ralstonia solanacearum* appliquée à l'épidémiologie du flétrissement bactérien de la pomme de terre à Madagascar.
44. Ravelomanantsoa, S., Robène, I., Chiroleu, F., Guérin, F., Poussier, S., Pruvost, O., and Prior, P. (2016) A novel multilocus variable number tandem repeat analysis typing scheme for African phylotype III strains of the *Ralstonia solanacearum* species complex. *PeerJ* **4**: e1949.
45. Ravelomanantsoa, S., Vernière, C., Rieux, A., Costet, L., Chiroleu, F., Arribat, S., et al. (2018) Molecular epidemiology of bacterial wilt in the Madagascar highlands caused by Andean (Phylotype IIB-1) and African (Phylotype III) brown rot strains of the *Ralstonia solanacearum* species complex. *Front Plant Sci* **8**..
46. Riley, M.A. and Chavan, M.A. eds. (2007) Bacteriocins: ecology and evolution, Berlin ; New York: Springer.
47. Russell, A.B., Peterson, S.B., and Mougous, J.D. (2014) Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol* **12**: 137–148.
48. Saenger, W. (2013) Chapter 714 - Proteinase K. In *Handbook of Proteolytic Enzymes (Third Edition)*. Rawlings, N.D. and Salvesen, G. (eds). Academic Press, pp. 3240–3242.
49. Safni, I., Cleenwerck, I., De Vos, P., Fegan, M., Sly, L., and Kappler, U. (2014) Polyphasic taxonomic revision of the *Ralstonia solanacearum* species complex: proposal to emend the

- descriptions of *Ralstonia solanacearum* and *Ralstonia syzygii* and reclassify current *R. syzygii* strains as *Ralstonia syzygii subsp. syzygii subsp. nov.*, *R. solanacearum* phylotype IV strains as *Ralstonia syzygii subsp. indonesiensis subsp. nov.*, banana blood disease bacterium strains as *Ralstonia syzygii subsp. celebesensis subsp. nov.* and *R. solanacearum* phylotype I and III strains as *Ralstonia pseudosolanacearum sp. nov.* *Int J Syst Evol Microbiol* **64**: 3087–3103.
50. Smidt, M.L. and Vidaver, A.K. (1982) Bacteriocin production by *Pseudomonas syringae* PsW-1 in plant tissue. *Can J Microbiol* **28**: 600–604.
 51. Tsai, J.W., Hsu, S.-T., and Chen, L.-C. (1985) Bacteriocin-producing strains of *Pseudomonas solanacearum* and their effect on development of bacterial wilt of tomato.
 52. Tudor-Nelson, S.M., Minsavage, G.V., Stall, R.E., and Jones, J.B. (2003) Bacteriocin-Like Substances from Tomato Race 3 Strains of *Xanthomonas campestris* pv. *vesicatoria*. *Phytopathology* **93**: 1415–1421.
 53. Vernière, C., Bui Thi Ngoc, L., Jarne, P., Ravigné, V., Guérin, F., Gagnevin, L., et al. (2014) Highly polymorphic markers reveal the establishment of an invasive lineage of the citrus bacterial pathogen *Xanthomonas citri* pv. *citri* in its area of origin. *Environ Microbiol* **16**: 2226–2237.
 54. Yahiaoui, N., Chéron, J.-J., Ravelomanantsoa, S., Hamza, A.A., Petrousse, B., Jeetah, R., et al. (2017) Genetic diversity of the *Ralstonia solanacearum* species complex in the South-West Indian Ocean islands. *Front Plant Sci* **8**: 2139.
 55. Zaluga, J., Stragier, P., Van Vaerenbergh, J., Maes, M., and De Vos, P. (2013) Multilocus Variable Number Tandem Repeats Analysis (MLVA) distinguishes a clonal complex of *Clavibacter michiganensis subsp. michiganensis* strains isolated from recent outbreaks of bacterial wilt and canker in Belgium. *BMC Microbiol* **13**: 126.
 56. Zhao, R., Duan, G., Yang, T., Niu, S., and Wang, Y. (2015) Purification, characterization and antibacterial mechanism of bacteriocin from *Lactobacillus acidophilus* XH1. *Trop J Pharm Res* **14**: 989–995.

Chapitre 2

Chapitre 2: Production de bactériocines chez les souches du complexe d'espèces *Ralstonia solanacearum*

Introduction

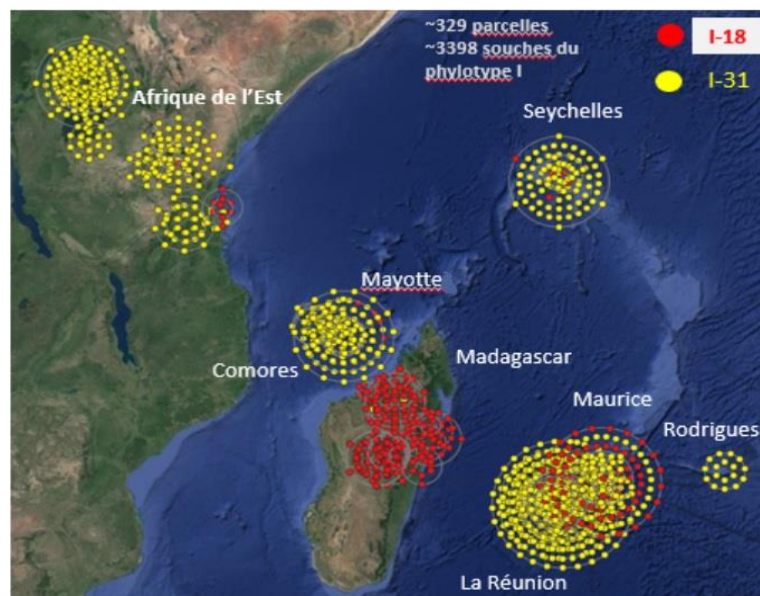
L'étude faite lors de la thèse de Yahiaoui (2015 – 2018) a révélé que le phylotype I est majoritaire dans les petites îles du SOOI (La Réunion, Maurice, Seychelles, Mayotte, Comores, Rodrigues), car il représente 87% des 1704 isolats collectés [203]. De plus, elle a mis en évidence une forte prévalence du sequevar I-31 dans la région, qui représente 68 % des souches du phylotype I isolées [203] (Figure 1).

À Madagascar, à l'exception des missions de collectes réalisées en 2013 dans cinq régions des hautes terres centrales, le phylotype I est largement majoritaire. En effet, 85.54% des 401 souches collectées en 2018 et 2019 dans 9 régions situées en basses, moyennes et hautes altitudes appartiennent au phylotype I. Un fait marquant de la situation épidémiologique à Madagascar est la prévalence du sequevar I-18 qui représente 76.67% des souches du phylotype I (n = 276), souches issues des collectes de 2006 - 2013 - 2018 - 2019) tandis que le sequevar I-46 ne représente que 11.39% des souches collectées (n = 41), le sequevar I-33, 1.39% (n = 5), et le sequevar I-31 n'est représenté que par deux souches (Figure 1).

Cette situation épidémiologique nous amène à penser que les souches fortement prévalentes pourraient être dotées de trait(s) biologique(s) impliqué(s) dans la compétition entre microorganismes : tels que la production de toxines, d'antibiotiques, et/ou de bactériocines. La capacité à mobiliser des bactériocines a particulièrement attiré l'attention de l'équipe, car elle a le potentiel d'expliquer un phénomène biologique de cette importance entre des groupes phylogénétiques apparentés. En effet, quelques études ont mis en évidence l'avantage compétitif que procure la production de bactériocines envers des souches sensibles. C'est le cas, par exemple, du *Pseudomonas*

chlororaphis 30-84 qui grâce à la production de R-tailocine est plus compétitif que *Pseudomonas putida* KT2440 dans la rhizosphère et en biofilm [139]. En outre, une étude faite par Kommineni et al. [145] a montré que la production de bactériocine 21 par la bactérie commensale *Enterococcus faecalis* portant le plasmide pPD1 favorise sa compétitivité vis-à-vis d'autres souches d'*enterococci* dans le tractus gastro-intestinal de la souris. De plus, la disponibilité du génome de quelques souches représentatives de la diversité génétique du ceRs dans les petites îles du SOOI et le séquençage de 8 souches représentatives des 4 lignées génétiques présentes à Madagascar nous a incités à réaliser une comparaison du contenu en gène des lignées génétiques présentes dans la zone. Trois principales questions de recherche sont alors posées au cours de cette étude :

- Quelle est l'activité bactériocine des souches du ceRs à Madagascar, dans les petites îles du SOOI et les souches représentatives de la diversité mondiale du ceRs ?
- Quelle est la diversité des clusters de gènes codant potentiellement des bactériocines chez les souches du ceRs ?
- Quels sont les gènes candidats pouvant être impliqués dans le succès épidémiologique des lignées I-18 à Madagascar et I-31 dans les petites îles du SOOI?



Figure_ 1: Répartition géographique des sequevars I-18 et I-31 à Madagascar, dans les petites îles du SOOI et en Afrique de l'Est

Matériels et méthodes

1. Souches bactériennes étudiées

Un total de 99 souches appartenant à différents groupes phylogénétiques du ceRs ont été utilisées pour l'étude de l'activité bactériocine du ceRs. Ces souches proviennent de Madagascar (n = 40), des petites îles du SOOI (n = 20), et d'autres parties du monde (n = 39). Cette sélection s'est basée sur les critères suivants : existence ou non de liens génétiques avec les souches malgaches, prévalence ou non de l'haplotype d'appartenance des souches, et disponibilité de leurs données génomiques (Tableau 1).

De plus, les souches RUN305, RUN3223 qui n'a pas été utilisée pour l'étude de l'activité bactériocine des souches du ceRs a été rattachée à la liste des souches bactériennes étudiées, car son génome a été utilisé lors de l'analyse pangénomique des souches du phylotype I Malgache et des petites îles du SOOI.

Tableau_ 1: Information sur les souches bactériennes étudiées

En gras, les souches dont les génomes sont disponibles.

En gras et colorée en **vert**, les souches utilisées pour l'analyse pangénomique.

RUN.ID	Souche ID	Hôte	Pays	Année de collecte	Phylotype-sequevar	Haplotype	Complexe clonal	Prévalence
RUN36	CFBP2957	<i>Solanum lycopersicum</i>	Martinique	1987	IIA-36			
RUN44	ACH92	<i>Zingiber officinale</i>	Australia	1966	I-16	MT032		Non
RUN47	CIP365	<i>Solanum tuberosum</i>	Philippines	1989	I-45			Non
RUN54	GMI1000	<i>Solanum lycopersicum</i>	French Guiana	1978	I-18	MT034		Non
RUN62	R229	<i>Banana</i>	Indonesia	1988	IV-10			
RUN68	M71	<i>Soil</i>	Malaysia	1998	I-18			
RUN83	PSI07	<i>Solanum lycopersicum</i>	Indonesia		IV-10			
RUN100	UW308	<i>Capsicum annuum</i>	Kenya	1974	I-31	MT168		Non
RUN104	UW395	<i>Solanum lycopersicum</i>	South Africa	1986	I-31	MT035		Oui
RUN133	CMR15	<i>Solanum lycopersicum</i>	Cameroon	2005	III-29			
RUN155	PSS366	<i>Solanum lycopersicum</i>	Taiwan	2003	I-15			
RUN157	PSS4	<i>Solanum lycopersicum</i>	Taiwan	1988	I-15	MT040		Non
RUN159	PSS358	<i>Solanum lycopersicum</i>	Taiwan	2003	I-15			
RUN184	CMR97	<i>Solanum scabrum</i>	Cameroon	2005	I-13			
RUN257	PSS219	<i>Solanum lycopersicum</i>	Taiwan	1996	I-34	MT045		Non
RUN305	MAD-002	<i>Solanum lycopersicum</i>	Madagascar	2006	I-18	MT002		Non
RUN309	MAD-006	<i>Capsicum annuum</i>	Madagascar	2006	I-18	MT005	Singleton	Non
RUN312	MAD-009	<i>Solanum melongena</i>	Madagascar	2006	I-18	MT004	MCC1	Oui
RUN314	MAD-011	<i>Solanum melongena</i>	Madagascar	2006	I-18	MT007	MCC1	Oui
RUN320	MAD-017	<i>Capsicum annuum</i>	Madagascar	2006	I-46	MT010	MCC3	Non
RUN323	MAD-020	<i>Capsicum annuum</i>	Madagascar	2006	I-46	MT011	MCC3	Non
RUN332	MAD-029	<i>Solanum tuberosum</i>	Madagascar	2006	III-19			
RUN343	M02	<i>Morus alba</i>	China		I-48			

RUN.ID	Souche ID	Hôte	Pays	Année de collecte	Phylotype-sequevar	Haplotype	Complexe clonal	Prévalence
RUN449	UW551	<i>Pelargonium asperum</i>	Kenya	2003	IIB-1			Oui
RUN586	UW163	<i>Banana plantain</i>	Peru	1967	IIB-4			
RUN619	RS41	<i>Solanum lycopersicum</i>	China	2007	I-17			Non
RUN825	02-167	<i>Solanum lycopersicum</i>	Martinique	2002	I-18			
RUN892	K60	<i>Solanum lycopersicum</i>	Usa	1953	IIA-7			
RUN939	JQ1118	<i>Solanum tuberosum</i>	Reunion	1988	IIB-1			Oui
RUN969	TO10	<i>Solanum lycopersicum</i>	Indonesia	2003	I-47			Non
RUN1002	RS67	<i>Solanum melongena</i>	India	1998	I-33			Non
RUN1012	RS52	<i>Solanum lycopersicum</i>	India	1997	I-48			
RUN1234	RS35	<i>Solanum melongena</i>	India	1997	I-31			
RUN1332	UQRS524	<i>Syzygium aromaticum</i>	Indonesia		IV-9A			
RUN1488	TNT 82A	<i>Solanum lycopersicum</i>	Trinidad	2008	I-14			
RUN1744	CIV23	<i>Solanum melongena</i>	Ivory Coast	2010	I-31			Oui
RUN1894	CIV138	<i>Solanum lycopersicum</i>	Ivory Coast	2011	I-18			
RUN1906	CIV150	<i>Solanum lycopersicum</i>	Ivory Coast	2011	I-46	MT074	Partage le même CC que MCC3	Non
RUN1985	E16A2	<i>Solanum melongena</i>	French Guiana	2011	I-17			
RUN2127	YT16.163	<i>Capsicum annuum</i>	Mayotte	2012	I-46	MT081	Partage le même CC que MCC3	Non
RUN2138	YT19.179	<i>Solanum melongena</i>	Mayotte	2012	I-15			
RUN2161	YT24.210	<i>Solanum melongena</i>	Mayotte	2012	I-18			
RUN2180	E'4T2	<i>Solanum lycopersicum</i>	French Guiana	2012	I-47			
RUN2279	MG9	<i>Solanum tuberosum</i>	Madagascar	2013	III-60			
RUN2280	MG11	<i>Solanum tuberosum</i>	Madagascar	2013	III-59			
RUN2340	MG78	<i>Solanum tuberosum</i>	Madagascar	2013	III-58			
RUN3012	TD13.01-C1	<i>Solanum melongena</i>	Reunion	2012	I-31			Oui

RUN.ID	Souche ID	Hôte	Pays	Année de collecte	Phylotype-sequevar	Haplotype	Complexe clonal	Prévalence
RUN3035	21	<i>Solanum lycopersicum</i>	Madagascar	2013	II			Oui
RUN3075	79	<i>Solanum lycopersicum</i>	Madagascar	2013	I-18	MT004	MCC1	Oui
RUN3082	83	<i>Solanum lycopersicum</i>	Madagascar	2013	I-18	MT015	MCC5	Oui
RUN3130	140	<i>Solanum lycopersicum</i>	Madagascar	2013	I-18	MT015	MCC5	Non
RUN3149	168	<i>Solanum tuberosum</i>	Madagascar	2013	I-33	MT021	MCC12	Non
RUN3159	186	<i>Solanum lycopersicum</i>	Madagascar	2013	I-18	MT022	MCC1	Non
RUN3210	250	<i>Solanum lycopersicum</i>	Madagascar	2013	I-18	MT004	MCC1	Oui
RUN3211	251	<i>Solanum lycopersicum</i>	Madagascar	2013	I-33	MT025	MCC12	Oui
RUN3216	257	<i>Solanum lycopersicum</i>	Madagascar	2013	I-33	MT025	MCC12	Non
RUN3223	264	<i>Solanum lycopersicum</i>	Madagascar	2013	I-33	MT026	MCC1	Non
RUN3245	295	<i>Solanum gilo</i>	Madagascar	2013	I-18	MT007	MCC1	Oui
RUN3277	368	<i>Solanum gilo</i>	Madagascar	2013	I-18	MT007	MCC1	Oui
RUN3287	380	<i>Solanum tuberosum</i>	Madagascar	2013	I-18	MT007	MCC1	Oui
RUN3430	584	<i>Phaseolus vulgaris</i>	Madagascar	2013	III-19			
RUN3431	614	<i>Solanum tuberosum</i>	Madagascar	2013	IIB-1			Oui
RUN3459	303	<i>Solanum gilo</i>	Madagascar	2013	I-18	MT030	MCC1	Non
RUN3553	Rs-PH-EGP-1	<i>Solanum melongena</i>	Philippines	2003	I-15			
RUN3568	Rs-TH-TOM-1	<i>Solanum lycopersicum</i>	Thailand	2010	I-13			Oui
RUN3570	BMa-A13	<i>Solanum lycopersicum</i>	Reunion	2014	I-33			Oui
RUN3634	43897	<i>Solanum macrocarpon</i>	Bénin	2014	I-31			Oui
RUN3812	S1T1	<i>Solanum lycopersicum</i>	Reunion	2015	IIA-36			
RUN4306	130	<i>Capsicum annuum</i>	Seychelles	2015	I-18			
RUN4326	241	<i>Solanum lycopersicum</i>	Seychelles	2015	I-31	MT035		Oui
RUN4407	90	<i>Solanum lycopersicum</i>	Mauritius	2015	I-15	MT092		Oui
RUN4509	Tu2P3	<i>Solanum tuberosum</i>	Reunion	2015	IIB-1			Oui
RUN4584	1.3	<i>Solanum tuberosum</i>	Mauritius	2015	I-18	MT093		
RUN4593	1.23	<i>Solanum tuberosum</i>	Mauritius	2015	I-33			Non
RUN4606	2.4	<i>Solanum tuberosum</i>	Mauritius	2015	IV-11			

RUN.ID	Souche ID	Hôte	Pays	Année de collecte	Phylotype-sequevar	Haplotype	Complexe clonal	Prévalence
RUN4630	3.19	<i>Solanum tuberosum</i>	Mauritius	2015	I-15			Oui
RUN5163	19.3	<i>Solanum tuberosum</i>	Mauritius	2015	I-31			Oui
RUN5274	2.32	<i>Solanum tuberosum</i>	Reunion	2015	III-19			
RUN5418	M3T10.126	<i>Solanum lycopersicum</i>	Rodrigues	2016	I-33			Oui
RUN5419	M3T10.127	<i>Solanum lycopersicum</i>	Rodrigues	2016	I-31			Oui
RUN5501	NDRB2T6	<i>Solanum lycopersicum</i>	Comoros	2016	I-31			Oui
RUN5538	OH2-1-2 (C212)	<i>Solanum lycopersicum</i>	Burkina Faso	2017	I-46			Oui
RUN5564	0	<i>Solanum lycopersicum</i>	New-Caledonia	2017	I-18	MT101		
RUN5598	BFR6-3-1	<i>Solanum aethiopicum</i>	Burkina Faso	2017	I-14			
RUN5935	170.1	<i>Capsicum annuum</i>	Madagascar	2018	I-46	MT014	MCC3	Oui
RUN5945	205.1	<i>Helianthus</i>	Madagascar	2018	I-46	MT014	MCC3	Oui
RUN5957	148.2	<i>Solanum americanum</i>	Madagascar	2018	I-46	MT329	MCC3	Non
RUN6185	T435	<i>Solanum lycopersicum</i>	Tanzania	2019	I-46			Non
RUN6253	7	<i>Solanum melongena</i>	Madagascar	2019	I-18	MT237	MCC2	Non
RUN6272	14	<i>Solanum nigrum</i>	Madagascar	2019	I	MT249	MCC2	Oui
RUN6275	17	<i>Solanum nigrum</i>	Madagascar	2019	I-18	MT241	MCC7	Non
RUN6279	19	<i>Capsicum annuum</i>	Madagascar	2019	I-18	MT242	MCC7	Non
RUN6285	26	<i>Solanuma aethiopicum</i>	Madagascar	2019	I-31	MT244	MCC11	Non
RUN6291	29	<i>Solanum lycopersicum</i>	Madagascar	2019	I-18	MT246	MCC1	Non
RUN6296	29	<i>Solanum lycopersicum</i>	Madagascar	2019	I	MT007	MCC1	Oui
RUN6340	32	<i>Solanum lycopersicum</i>	Madagascar	2019	I-31	MT363	MCC11	Non
RUN6351	32	<i>Solanum lycopersicum</i>	Madagascar	2019	I-18	MT253	MCC1	Non
RUN6361	33	<i>Physalis sp</i>	Madagascar	2019	I-18	MT255	MCC2	Non
RUN6414	38	<i>Solanuma aethiopicum</i>	Madagascar	2019	I	MT004	MCC1	Oui
RUN6499	45	<i>Solanum aethiopicum</i>	Madagascar	2019	I	MT249	MCC2	Oui
RUN6534	209	<i>Capsicum annuum</i>	Madagascar	2019	I	MT004	MCC1	Oui

2. Protocole d'expérimentation de typage bactériocine

À partir d'une colonie bactérienne, un repiquage est effectué sur du milieu Kelman et les boîtes de Petri ont été placées à l'étuve (28°C) pendant 48h pour permettre le développement bactérien et l'obtention d'un pâté bactérien.

À partir du pâté bactérien, l'expérimentation s'est faite en deux étapes (Figure 2): (1) la production de bactériocines; (2) la préparation du tapis bactérien et le dépôt des spots de bactériocine. Le protocole d'expérimentation est adapté des protocoles publiés dans les articles de Gross, Vivader [209] et Frey et al. en 1994 [210]. Pour vérifier la répétabilité des résultats observés, l'expérimentation a été répétée deux fois.

2.1. Production de bactériocines

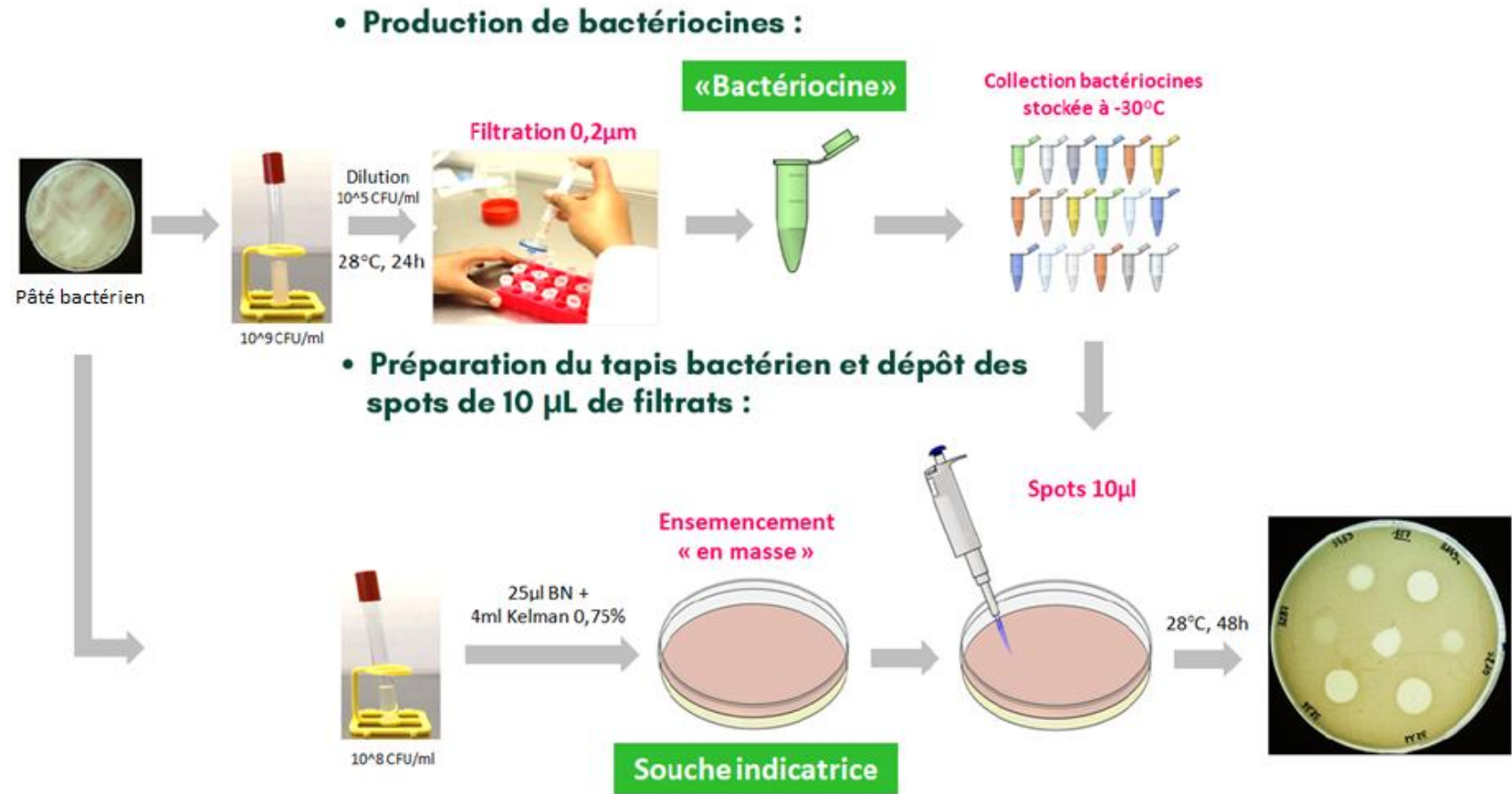
Un total de 37 souches représentatives de la diversité phylogénétique du ceRs (4 phylotypes, 13 sequevars) et provenant de Madagascar (n = 21), des petites îles du SOOI (n = 8) et d'autres régions du monde (n = 8) ont été sélectionnées pour tester leurs capacités d'inhibition de croissance (Tableau 1).

Des tubes de BN ayant chacun un volume de 4.5 ml ont étéensemencés avec les pâtés bactériens des 37 souches étudiées pour avoir une concentration finale de 10^5 CFU/ml. Ensuite, ils ont été placés à l'étuve (28°C) pendant 24h. Au bout de 24h, les BN ont été filtrés avec un filtre 0.2µm pour enlever toutes bactéries, et les filtrats récupérés qui contiendraient les bactériocines ont été stockés dans des tubes eppendorfs 2 ml à -30°C. Il est à noter qu'un test permettant de vérifier la stérilité des filtrats a été fait. Pour cela, 10 µl de chacun des 37 filtrats récupérés ont été déposés sur du milieu Kelman et placés à l'étuve (28°C) pendant 48h pour observer la croissance ou non des bactéries. De plus, un test a été préalablement effectué pour vérifier l'absence d'altération des activités d'inhibition des bactériocines après plusieurs congélations et décongélations.

2.2. Préparation des tapis bactériens et dépôts des spots de bactériocine

Un total de 99 souches représentatives de la diversité phylogénétique du ceRs (4 phylotypes, 23 sequevars) et provenant de Madagascar (n = 40), des petites îles du SOOI (n = 20), et d'autres régions du monde (n = 39) ont été sélectionnées pour servir de souches indicatrices de l'activité d'inhibition.

Des tubes de BN ayant chacun un volume de 4.5 ml ont étéensemencés avec les patés bactériens des 99 souches indicatrices pour avoir une concentration finale de 10^8 CFU/ml. Ensuite, 4 ml de milieu Kelman semi-solide (avec 0.75% d’agar) ont étéensemencés avec 25 μ l de culture bactérienne et coulés dans des boıtes de Petri contenant du milieu Kelman solide. Apres sechage et solidification, 10 μ l de bactériocine ont été déposés en spot et les boıtes de Petri ont été placées à l’etuve (28°C) pendant 48h pour permettre la croissance du tapis bactérien et l’observation des halos d’inhibition.



Figure_2: Schéma récapitulatif du protocole d'expérimentation de typage bactériocine

3. Contrôle de l'activité des bactériocines

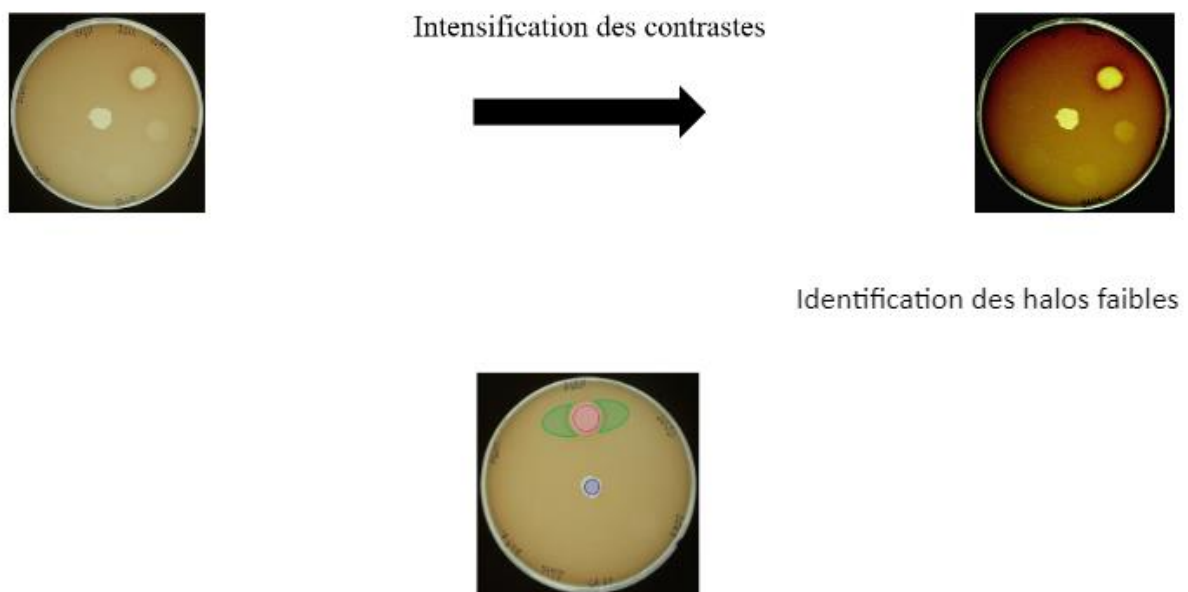
Afin de vérifier que les halos d'inhibition étaient bien dus à l'activité bactériocine, quatre tests basés sur les caractéristiques des bactériocines ont été effectués :

- 1er test : Vérification de l'absence de halo d'inhibition après traitement à la chaleur étant donné que la grande majorité des bactériocines sont thermolabiles. Pour cela, 100 µL du filtrat ont été placés dans un thermocycleur pendant 10 min à 100°C.
- 2ème test : Vérification de l'absence de halo d'inhibition après traitement à la protéinase K. Pour cela, 50µl d'extraits de bactériocine ont été mélangés à 50 µl de protéinase K à une concentration de 50 µg/ml puis placés au thermocycleur à 37°C pendant 2h. Les bactériocines sont de nature protéique et devraient être dégradées par les protéases.
- 3ème test : Vérification de l'absence de halo d'inhibition sur 19 souches d'espèces plus ou moins éloignées phylogénétiquement du ceRs : *Ralstonia eutropha*, *Ralstonia pickettii*, *Pseudomonas savastanoi*, *Pseudomonas fluorescens*, *Pseudomonas savastanoi pv. savastanoi*, *Pseudomonas cichorii*, *Pseudomonas corrugata*, *Pseudomonas syringae pv. tomato*, *Pseudomonas putida*, *Erwinia chrysanthemi*, *Burkholderia andropogonis*, *Xanthomonas axonopodis pv. dieffenbachiae*, *Xanthomonas vesicatoria*, *Xanthomonas citri pv. citri*, *Xanthomonas gardneri*, *Xanthomonas euvesicatoria*, *Xanthomonas perforans*, *Clavibacter michiganensis* (GRAM +), *Escherichia coli*. L'activité bactériocine des souches du ceRs ne devrait concerner que des souches apparentées phylogénétiquement.
- 4ème test : Vérification de l'atténuation de l'intensité du halo d'inhibition après dilution en série du filtrat. Ce test permet de distinguer les inhibitions dues aux phages des inhibitions dues aux bactériocines. En effet, pour le cas des inhibitions dues aux phages, on observe au fur et à mesure des dilutions en série l'apparition de plaque de lyse tandis que pour les inhibitions dues aux bactériocines, on observe une atténuation de l'intensité du halo d'inhibition voir sa disparition lorsque le filtrat est très dilué.

4. Analyse des données de typage bactériocine

Pour servir de témoin positif dans le calcul de l'intensité de l'activité bactériocine, le tapis bactérien au centre de la boîte de Petri a été enlevé manuellement. Les données de typage bactériocine ont été obtenues après photographie des halos d'inhibition. En déposant les boîtes de Petri sur une source lumineuse, elles ont été prises en photo avec un appareil Canon EOS 6D Mark II paramétré de la manière suivante : focale F/11 de longueur 35mm, 1/160 secondes de temps d'exposition, vitesse ISO : ISO-1600.

Les photos des boîtes de Petri ont été analysées avec le logiciel imageJ (<https://imagej.nih.gov/>). Tout d'abord, nous avons procédé à l'identification des halos faibles en ajustant le contraste et la luminosité des photos (Figure 3). Une fois les halos faibles identifiés, la photo a été réinitialisée à son paramètre initial puis l'intensité lumineuse traversant les halos ont été mesurés et comparés à l'intensité lumineuse traversant le tapis bactérien et le témoin positif correspondant à l'absence de tapis bactérien (Figure 3).



Figure_ 3: Identification des halos d'inhibition

La formule suivante a été utilisée pour calculer l'activité bactériocine, soit la perte en tapis bactérien:

$$\text{Activité bactériocine} = \frac{\text{Intensité lumineuse traversant le halo} - \text{Intensité lumineuse traversant le tapis bactérien}}{\text{Intensité lumineuse traversant le témoin positif} - \text{Intensité lumineuse traversant le tapis bactérien}}$$

Deux jeux de données ont été obtenus à partir des deux répétitions biologiques de l'expérimentation de typage bactériocine. Ainsi, pour vérifier la répétabilité des résultats obtenus, elles ont été combinées et analysées ensemble. Un dendrogramme a été construit sur R v.4.1.3 pour visualiser le regroupement des données similaires en se basant sur la distance de Minkowski qui est une généralisation à la fois de la distance euclidienne et de la distance de Manhattan. Ensuite un heatmap basé sur l'intensité de l'inhibition des souches étudiées a été créé sur iTOL v6 (<https://itol.embl.de>) en y important le dendrogramme et les jeux de données combinés.

5. Protocole d'extraction d'ADN génomique à haut poids moléculaire

L'ADN génomique de 8 souches (RUN320, RUN3149, RUN3223, RUN5935, RUN6285, RUN6340, RUN6499, RUN6534) (Tableau 1) représentatives de la diversité génétique du phylotype I malgache a été extrait selon un protocole adapté de Wilson [211]. Du bouillon nutritif (volume = 4.5 ml) a été inoculé et incubé pendant une nuit sous agitation à 180 rpm à 28 °C. Ensuite, 4.5 ml de culture ont été centrifugés pendant 5 minutes à 4000 g et le surnageant a été éliminé. Le culot bactérien a été remis en suspension dans 1 ml de NaCl (0,5 M) et centrifugé pendant 5 minutes à 4000 g. Pour la lyse cellulaire, le culot a été remis en suspension avec 567 µl de Tris-EDTA et 30 µl de SDS à 10%, 3 µl de protéinase K à 20 mg/ml ont été ajoutés. Le mélange a été incubé pendant 1 heure à 37°C. Puis 100 µl de NaCl (5M) ont été ajoutés, mélangés par pipetages répétés et incubés 10 min à 65°C. Un volume égal de chloroforme : alcool isoamylique 24:1 (généralement compris entre 0,7 mL et 0,8 mL) a été ajouté au lysat, mélangé par pipetages répétés et ce mélange a été centrifugé 10 minutes à 10000 g. La phase supérieure aqueuse a été transférée dans un nouvel eppendorf de 2 ml, en prenant soin de ne pas pipeter l'interface. Deux microlitres de RNase A ont été ajoutés au surnageant et incubés pendant 30 min à 37°C. Un volume égal de phénol : chloroforme : alcool isoamylique 25:24:1 (généralement compris entre 0,5 mL et 0,6 mL) a ensuite été ajouté au surnageant et mélangé par pipetages répétés. Ce mélange a été centrifugé pendant 10 min à 10000g. La phase supérieure aqueuse a été transférée dans un nouvel eppendorf de 1,5 ml puis 0,6 volume d'isopropanol a été ajouté pour précipiter les acides nucléiques. Le tube a été agité par inversion jusqu'à ce qu'un précipité d'ADN blanc et filamenteux devienne clairement visible. L'ADN a été culotté par centrifugation à 10000 g pendant 2 minutes et le surnageant a été éliminé. À ce stade, 1 ml d'éthanol à 70 % a été ajouté pour rincer l'ADN. Une nouvelle centrifugation a été

réalisée pendant 1 min à 10000 g puis le surnageant a été soigneusement jeté et l'ADN a été resuspendu dans 100 µl de tampon TE.

6. Séquençage, assemblage et annotation des génomes

Les 8 échantillons d'ADN ont été multiplexés et séquencés par la plateforme de séquençage du Laboratoire Genewiz-Azenta (Leipzig, Allemagne) avec la technologie Pacific Bioscience (PacBio) Sequel I à l'aide d'une SMRT Cell. Les données démultiplexées des subreads au format fastq compressé (subreads.fastq.gz) ont été utilisées pour réaliser les assemblages des génomes avec le programme Canu v2.1.1 [212]. L'intégrité du génome a été évaluée en utilisant Busco v.5.2.2 [213] qui a recherché dans les génomes assemblés un set de 124 gènes conservés parmi les espèces bactériennes. De plus, Quast v.5.0.0rc1 [214] a été utilisé pour évaluer la qualité de l'assemblage en donnant des informations sur le niveau de fragmentation du génome (nombre de contigs, N50⁵, etc...). Les génomes assemblés ont également été alignés avec le génome de référence du ceRs (GMI1000) en utilisant l'outil D-GENIES [215] afin de vérifier la colinéarité des assemblages avec le génome de référence GMI1000. L'annotation du génome a été faite à l'aide du logiciel Prokka v1.14.5 [216] qui est un outil bioinformatique spécialisé dans l'annotation du génome des procaryotes.

7. Recherche *in silico* des clusters de gènes bactériocine

L'outil d'annotation des bactériocines BAGEL 4 [217] a été utilisé pour la recherche *in silico* des clusters de gènes codant potentiellement des bactériocines contenues dans 51 génomes du ceRs (Tableau 1). Le résultat a été affiché sur un arbre phylogénétique en soumettant à IQTREE v2.1.3 [218] le résultat de l'alignement des SNPs présents dans le core gène⁶ des 51 génomes. L'identification des SNPs a été faite par SNIPPY v4.6.0 [219] en utilisant la fonction multi qui permet de réaliser un appel de variants à partir de séquences de génomes assemblés.

⁵ Le N50 est une mesure de la fragmentation du génome assemblé en contigs. Lorsque les contigs sont triés par taille et que l'on somme les contigs un par un du plus grand au plus petit, le N50 est la longueur du dernier contig obtenu lorsque tous les contigs déjà parcourus couvrent la moitié de la taille du génome. Plus le N50 est élevé, moins le génome est fragmenté.

⁶ Le core gène est composé des gènes communs aux génomes analysés.

De plus, le gène SinB qui a été suggéré comme étant impliqué dans la production de bactériocines chez le ceRs lors de la thèse d'Huerta [220] a été recherché dans nos 51 génomes via l'outil bioinformatique ncbi-blast-2.10.1+.

8. Analyse pangénomique des souches du ceRs

Pour rechercher d'autres gènes candidats pouvant expliquer le succès épidémiologique de certaines lignées génétiques du ceRs dans le SOOI, nous avons réalisé une analyse pangénomique⁷ à partir de tous les génomes disponibles du phylotype I dans le SOOI et du génome de référence du ceRs (RUN54 ou GMI1000). Au total, 20 génomes ont été retenus (Tableau 1): *i*) RUN305, RUN3223, RUN3277, RUN6499, et RUN6534 provenant de Madagascar et RUN54 provenant de la Guyane Française représentent les souches I-18; *ii*) RUN4630 et RUN4407 représentent les souches mauriciennes I-15; *iii*) RUN3012 (La Réunion), RUN4326 (Seychelles), RUN5163 (Maurice), RUN5419 (Rodrigues), RUN6285 et RUN6340 (Madagascar) représentent la lignée génétique I-31, *iv*) RUN3149 et RUN3216 représentent les souches malgaches I-33, *v*) RUN320 (issue des séquençages PacBio et Illumina), RUN2127, RUN5935 représentent les souches malgaches I-46.

Tout d'abord, un appel de variants a été réalisé avec SNIPPY v4.6.0 [219] à partir des 20 génomes des souches du SOOI et en utilisant comme génome de référence celui de la souche malgache fortement prévalente RUN6435. Cet appel de variants a permis d'identifier 1.184.023 SNPs. Parallèlement à cela, le programme Roary v1.14.5 [221] a été utilisé pour identifier le pangéome⁸ des 20 souches étudiées. Le résultat a été illustré en construisant un diagramme de Venn via la plateforme web InteractiVenn [222]. De plus, les effecteurs de type III qui sont impliqués dans la virulence des souches du ceRs ont été identifiés en soumettant les 20 génomes à la plateforme web <http://iant.toulouse.inra.fr/bacteria/annotation/site/prj/T3Ev3>.

⁷ Le pangéome représente l'ensemble des gènes d'une espèce.

⁸ Ici, le pangéome est constitué par l'ensemble des gènes présents au moins une fois dans un des 20 génomes.

Résultats

1. Activités bactériocines des souches du ceRs

D'un point de vue général, les activités d'inhibition des souches productrices et la sensibilité ou non des souches indicatrices se structurent suivant leurs sequevars (Figure 4).

On constate remarquablement que deux souches appartenant au sequevar IIA-36 (RUN 36 et RUN 3812) présentent une activité antimicrobienne forte contre tous les sequevars testés sauf la souche RUN 1332 du sequevar IV-9. La souche RUN4606 appartenant au sequevar IV-10, collectée à Maurice, présente également un profil d'inhibition intéressant, car elle a une activité antimicrobienne forte contre tous les sequevars testés sauf I-13, IIA-36, et IV-9. La souche RUN 892 (IIA-7) quoique présentant une activité antimicrobienne à intensité variable, a aussi un large spectre d'action inhibant les souches autres que celles des sequevars IIA-36 et IV-9.

Par opposition, les 37 souches productrices testées n'ont pas d'activité antimicrobienne contre les souches RUN 3812 (IIA-36) et RUN 1332 (IV-9).

Certaines souches sont particulièrement actives contre un sequevar précis. C'est le cas des souches I-30 d'Inde ($n = 1$) et III-19 de Madagascar ($n = 1$) qui ont une activité antimicrobienne forte contre les I-18; la souche I-48 de Chine contre les souches du sequevar I-15 et I-46; et la souche I-15 de Philippines contre les I-31.

Les souches du sequevar I-18 ($n = 14$) et I-31 ($n = 5$) provenant de Madagascar et des petites îles du SOOI sont actives contre les lignées génétiques présentes dans la zone (I-15, I-18, I-31, I-33, I-46, IIB-1, III-19, IV-11) [12,203,223]. Or les situations épidémiologiques qu'on y observe nous révèlent une forte prévalence de ces deux lignées génétiques, le sequevar I-18 à Madagascar et le sequevar I-31 dans les petites îles du SOOI. Étant donné l'avantage que pourrait donner la production de substances antimicrobiennes dans la compétition pour l'espace et les ressources, ces résultats nous montrent une cohérence entre les activités d'inhibition et les situations épidémiologiques décrites dans la zone.

Selon leurs profils d'inhibition, les souches productrices du sequevar I-18 peuvent être regroupées en trois clusters. Le premier cluster est composé de 6 souches provenant de Madagascar, le second plus grand complexe clonal MCC2 (n = 3) et un complexe clonal mineur proche génétiquement MCC7 (n = 2), et d'une souche de Nouvelle-Calédonie. Les souches composant ce cluster ont une activité antimicrobienne à large spectre, attaquant des souches appartenant à 19 sequevars I-14, I-15, I-16, I-17, I-30, I-31, I-33, I-34, I-45 (halo très faible), I-46, I-47, IIA-7, IIA-36, IIB-1, IIB-4, III-19, III-29, IV-10, IV-11 collectés à Madagascar, dans les petites îles du SOOI et d'autres régions du monde. Remarquablement deux souches malgaches du MCC2 (RUN 6275, RUN 6279) et la souche néo-calédonienne (RUN 5564) sont mêmes faiblement actives contre des souches du sequevar I-18 malgache MCC1 (RUN 312, RUN 6534, RUN 6414), MCC2 (RUN 6499) et ivoirienne (RUN 1894).

Le second cluster est composé de deux souches appartenant au plus grand complexe clonal malgache MCC1. Les souches composant ce cluster ont une activité antimicrobienne à large spectre contre des souches appartenant à 17 sequevars I-14, I-15, I-16, I-17, I-30 (halo très faible), I-31, I-33, I-34, I-46, I-47, I-48, IIA-7, IIB-1, IIB-4, III-19, IV-10, IV-11 présents à Madagascar, dans les petites îles du SOOI et d'autres régions du monde. Remarquablement, ces deux souches ont également une activité antimicrobienne forte contre des souches du sequevar 18 provenant de Madagascar: MCC1 (RUN 312, RUN 3075, RUN 3159, RUN 3277, RUN 6291, RUN 6534), MCC2 (RUN 6253), MCC5 (RUN 3082, RUN 3130); et d'une souche seychelloise (RUN 4306), et mauricienne (RUN 4584).

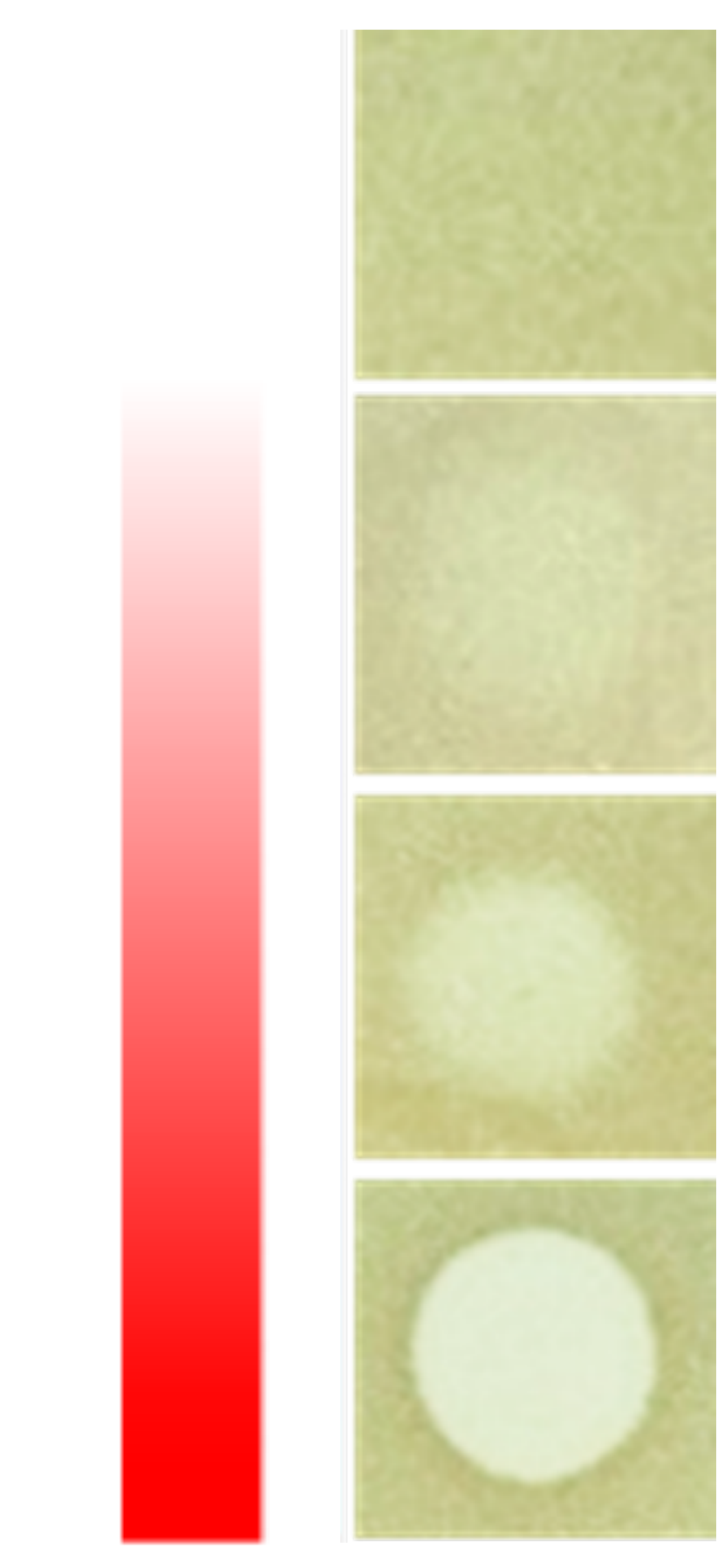
Le troisième cluster est composé de sept souches malgaches issues du MCC1 (n = 6) et d'une souche issue du MCC2 (n = 1, RUN 3130). Contrairement aux deux précédents clusters, les souches composant celui-ci présentent une activité antimicrobienne moins intense et elles n'ont pas d'activité antimicrobienne contre les souches du sequevar I-18. Néanmoins, elles ont un large spectre d'action, inhibant la croissance de 15 sequevars: I-14 (halo très faible), I-15, I-16, I-17, I-30 (halo très faible), I-31, I-33, I-34, I-47, I-48, IIA-7, IIB-1, IIB-4, III-19, IV-10, IV-11 collectés à Madagascar, dans les petites îles du SOOI et d'autres régions du monde.

Un résultat notable est l'activité antimicrobienne forte des souches du sequevar I-18 malgaches et néo-calédoniennes contre les souches du sequevar I-31 provenant de Madagascar (n = 2), La Réunion (n = 1, une représentante de l'haplotype le plus prévalent dans les petites îles du SOOI), Seychelles (n = 1), Maurice (n = 1), Rodrigues (n = 1), Kenya (n = 1), Afrique du Sud (n = 1), Côte d'Ivoire (n = 1), Bénin (n = 1), Inde (n = 1), et Thailand (n = 1).

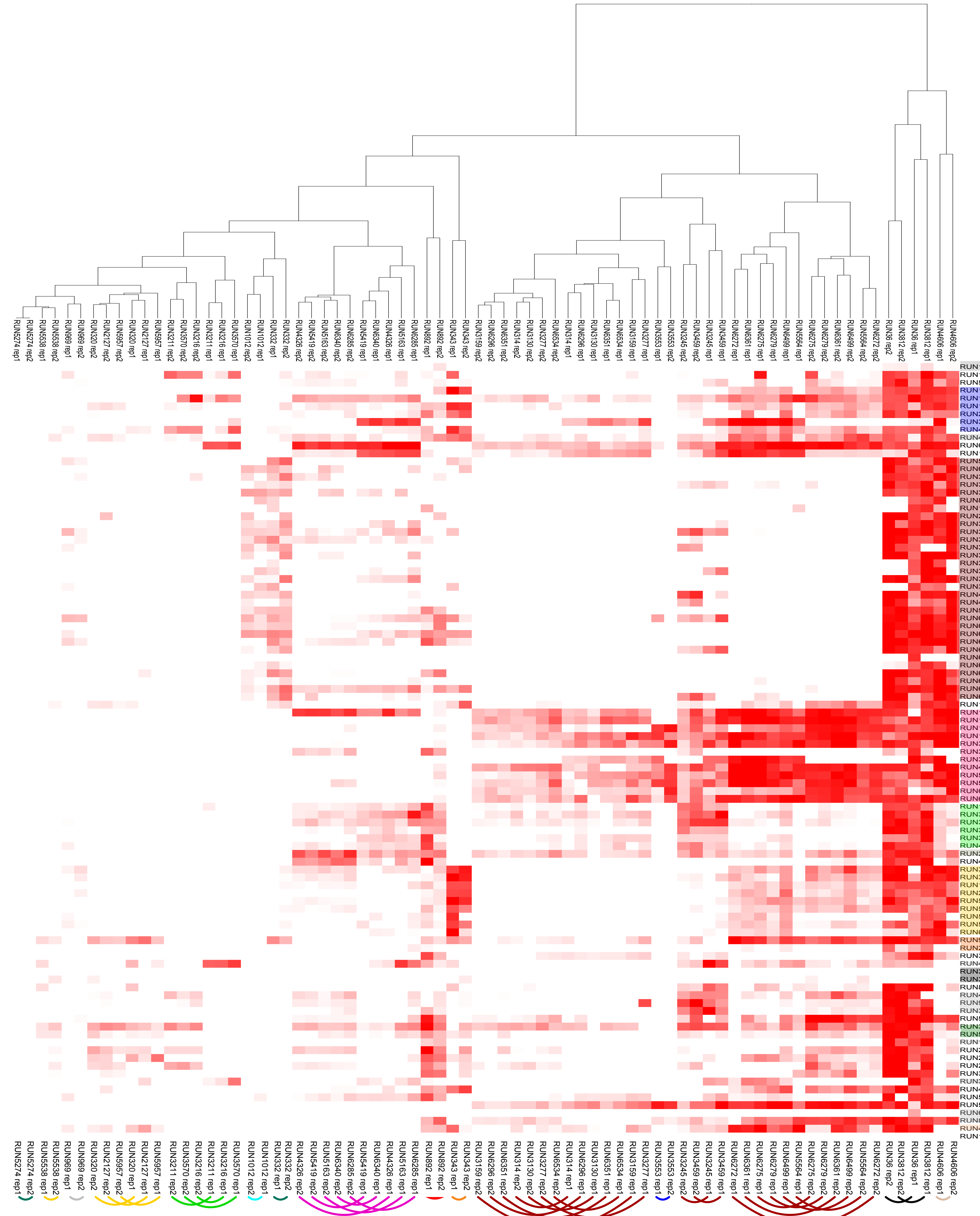
Par ailleurs, la caractérisation partielle des surnageants de culture a révélé que leurs activités inhibitrices sont sensibles à la chaleur (100°C, 10 mn), et à la protéinase K sauf ceux de 10 souches appartenant aux sequevars I-15, I-18 (MCC2), IIA-36, IIA-7, IV-11. Les surnageants n'ont pas d'activités inhibitrices sur les souches autres que le ceRs. La dilution en série des surnageants a montré une atténuation de l'intensité du halo d'inhibition voir sa disparition lorsque le filtrat est très dilué et l'apparition de plaque de lyse pour le cas de 8 souches appartenant aux sequevars I-15 (RUN 3553), I-18 (RUN 3130, et les RUN 3245, RUN 3459 qui compose le cluster 2 selon le profil d'inhibition des souches I-18), I-33 (RUN3211, RUN 3216, RUN 3570), I-47 (RUN 969). Combinés ensemble, ces résultats suggèrent une inhibition bactérienne due aux bactériocines pour le cas de 29 souches productrices et une inhibition due aux phages pour le cas des 8 souches citées précédemment.

Tree scale: 100

Growth inhibition activity



- I-15
- I-18
- I-30
- I-31
- I-33
- I-46
- I-47
- I-48
- IIA-7
- IIA-36
- III-19
- IV-11



RUN4606 rep2
RUN4606 rep1
RUN3812 rep1
RUN3812 rep2
RUN36 rep2
RUN36 rep1
RUN606 rep1
RUN606 rep2
RUN1488
RUN5598
RUN155
RUN157
RUN2138
RUN3553
RUN4407
RUN184
RUN619
RUN1985
RUN54
RUN68
RUN309
RUN312
RUN314
RUN825
RUN1894
RUN2161
RUN3075
RUN3082
RUN3130
RUN3159
RUN3210
RUN3245
RUN3277
RUN3287
RUN3459
RUN4306
RUN4584
RUN5564
RUN6253
RUN6272
RUN6275
RUN6279
RUN6291
RUN6296
RUN6351
RUN6361
RUN6414
RUN6499
RUN6534
RUN1012
RUN100
RUN104
RUN1234
RUN1744
RUN3012
RUN3568
RUN3634
RUN4326
RUN5163
RUN5419
RUN6285
RUN6340
RUN1002
RUN3149
RUN3211
RUN3216
RUN3570
RUN4593
RUN257
RUN47
RUN320
RUN323
RUN1906
RUN2127
RUN5538
RUN5935
RUN5945
RUN5957
RUN6185
RUN969
RUN2180
RUN343
RUN4509
RUN3812
RUN36
RUN892
RUN449
RUN939
RUN3431
RUN586
RUN332
RUN5274
RUN133
RUN2279
RUN2280
RUN2340
RUN3430
RUN3035
RUN4630
RUN5418
RUN5501
RUN62
RUN83
RUN4606
RUN1332
RUN4606 rep2
RUN4606 rep1
RUN3812 rep1
RUN3812 rep2
RUN36 rep2
RUN36 rep1
RUN606 rep1
RUN606 rep2
RUN1488
RUN5598
RUN155
RUN157
RUN2138
RUN3553
RUN4407
RUN184
RUN619
RUN1985
RUN54
RUN68
RUN309
RUN312
RUN314
RUN825
RUN1894
RUN2161
RUN3075
RUN3082
RUN3130
RUN3159
RUN3210
RUN3245
RUN3277
RUN3287
RUN3459
RUN4306
RUN4584
RUN5564
RUN6253
RUN6272
RUN6275
RUN6279
RUN6291
RUN6296
RUN6351
RUN6361
RUN6414
RUN6499
RUN6534
RUN1012
RUN100
RUN104
RUN1234
RUN1744
RUN3012
RUN3568
RUN3634
RUN4326
RUN5163
RUN5419
RUN6285
RUN6340
RUN1002
RUN3149
RUN3211
RUN3216
RUN3570
RUN4593
RUN257
RUN47
RUN320
RUN323
RUN1906
RUN2127
RUN5538
RUN5935
RUN5945
RUN5957
RUN6185
RUN969
RUN2180
RUN343
RUN4509
RUN3812
RUN36
RUN892
RUN449
RUN939
RUN3431
RUN586
RUN332
RUN5274
RUN133
RUN2279
RUN2280
RUN2340
RUN3430
RUN3035
RUN4630
RUN5418
RUN5501
RUN62
RUN83
RUN4606
RUN1332

Figure_ 4: Activité inhibitrice de 37 souches productrices de bactériocine (horizontal) sur 99 souches indicatrices (vertical)

Le dendrogramme a été construit sur la base de la similarité/dissimilarité des profils d'inhibition des 37 souches productrices. La couleur des liens entre les souches productrices est dépendante de leurs sequevars. Cette représentation en heatmap est basée sur la sensibilité des souches indicatrices aux composés antimicrobiens. Les souches indicatrices (vertical) ont été regroupées en fonction de leur sequevar.

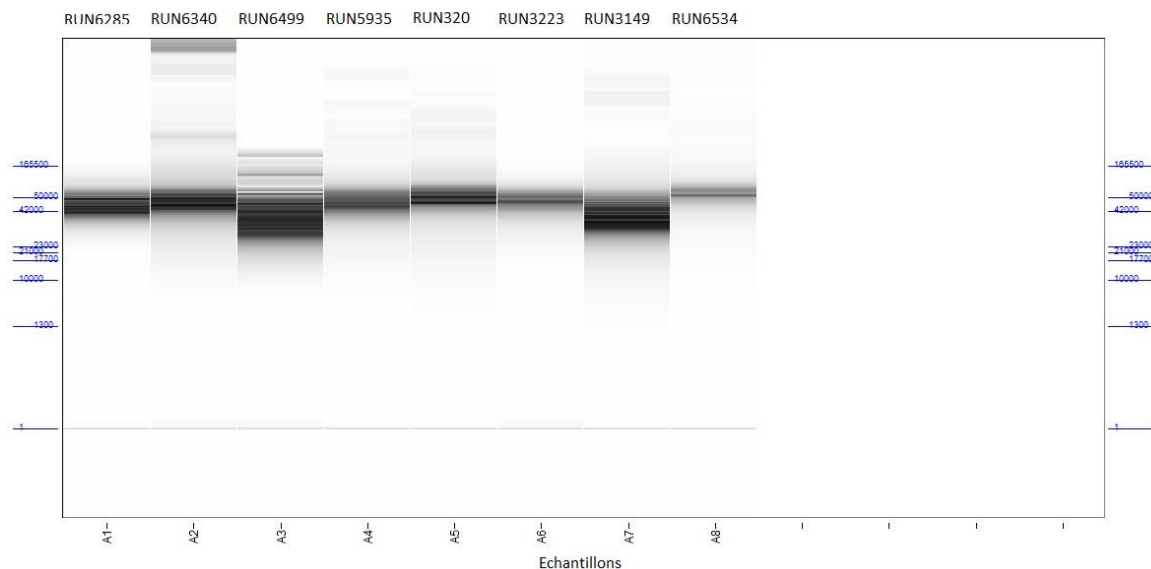
2. Qualité et quantité des ADN génomiques extraits

Cette thèse s'est également attelée à produire les ressources génomiques des souches représentatives de la diversité génétique du ceRs à Madagascar. L'ADN génomiques de ces souches ont alors été extraits avant de procéder à leurs séquençages. L'évaluation de la qualité et de la quantité des extraits d'ADN génomiques a été réalisée par le laboratoire Genewiz-Azenta NGS (Tableau 2).

Tableau_ 2: Qualité et quantité d'ADN génomiques extraits

Sample ID	AZENTA ID	Sample Type	Sample Vol.(ul)	Nanodrop 2000				DNA Qubit	
				Actual Nucleic Acid Conc. (ng/ul)	Total Amount (ng)	A260/A280	A260/A230	Actual Nucleic Acid Conc. (ng/ul)	Total Amount (ng)
RUN6285	SEB06822	HMW gDNA	77	169.60	13059.20	1.85	1.80	169.00	13013.00
RUN6340	SEB06823	HMW gDNA	77	321.50	24755.50	1.90	2.04	246.00	18942.00
RUN6499	SEB06824	HMW gDNA	79	290.90	22981.10	1.90	2.03	296.00	23384.00
RUN5935	SEB06825	HMW gDNA	80	524.90	41992.00	1.90	2.08	362.00	28960.00
RUN320	SEB06826	HMW gDNA	78	176.30	13751.40	1.85	1.87	197.00	15366.00
RUN3223	SEB06827	HMW gDNA	78	194.20	15147.60	1.88	1.80	67.00	5226.00
RUN3149	SEB06828	HMW gDNA	77	188.00	14476.00	1.89	1.87	232.00	17864.00
RUN6534	SEB06829	HMW gDNA	79	228.50	18051.50	1.88	1.92	180.00	14220.00

Le ratio A260/A280 est compris entre 1.85 et 1.9 et le ratio A260/A230 est compris entre 1.8 et 2.08. Ces résultats indiquent que nos solutions d'ADN sont exemptes de protéines absorbantes à 280 nm et des polysaccharides, carbohydrates et autres contaminants absorbants à 230 nm. Les concentrations d'ADN des échantillons varient entre 169.60 et 524.90 ng/μl au nanodrop et entre 67 et 362 ng/μl au Qubit. L'analyse au Fragment Analyzer des échantillons d'ADN génomique révèle leurs bonnes intégrités, car les pics de distribution de taille des molécules varient entre ~42 kb et ~50 kb (Figure 5).



Figure_ 5: Gel digital de migration des ADN génomiques extraits

A droite et à gauche du gel sont positionnés les marqueurs de taille.

3. Qualité de l'assemblage des génomes

Les résultats Busco (Tableau 3) nous montrent que la totalité des 124 gènes conservés parmi les espèces bactériennes y ont été identifiés, et 0.8% de ces gènes sont dupliqués.

Tableau_ 3: Résultats de l'analyse Busco présentant les pourcentages d'intégrité des génomes, les pourcentages de genes dupliqués, fragmentés et manquants parmi les 124 gènes recherchés

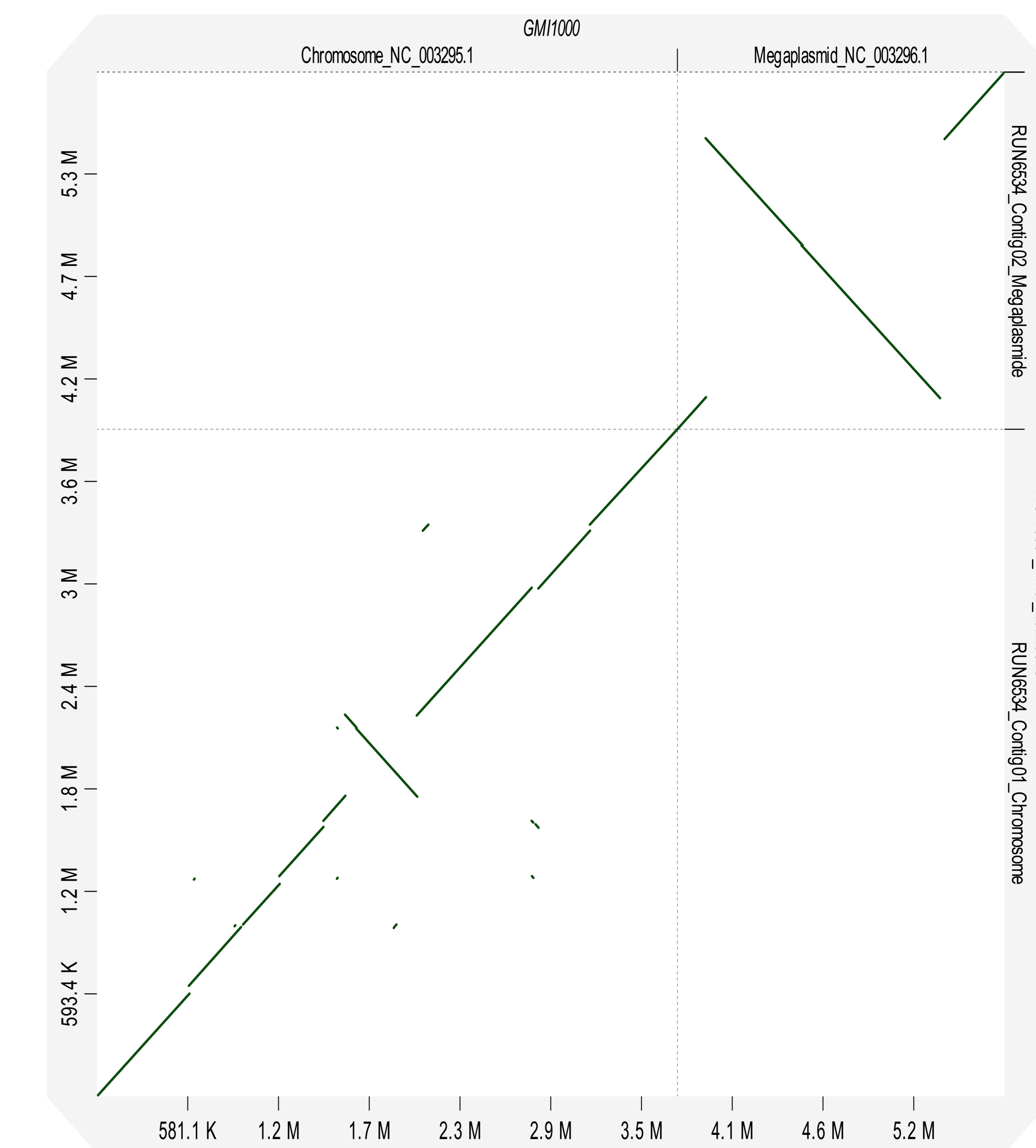
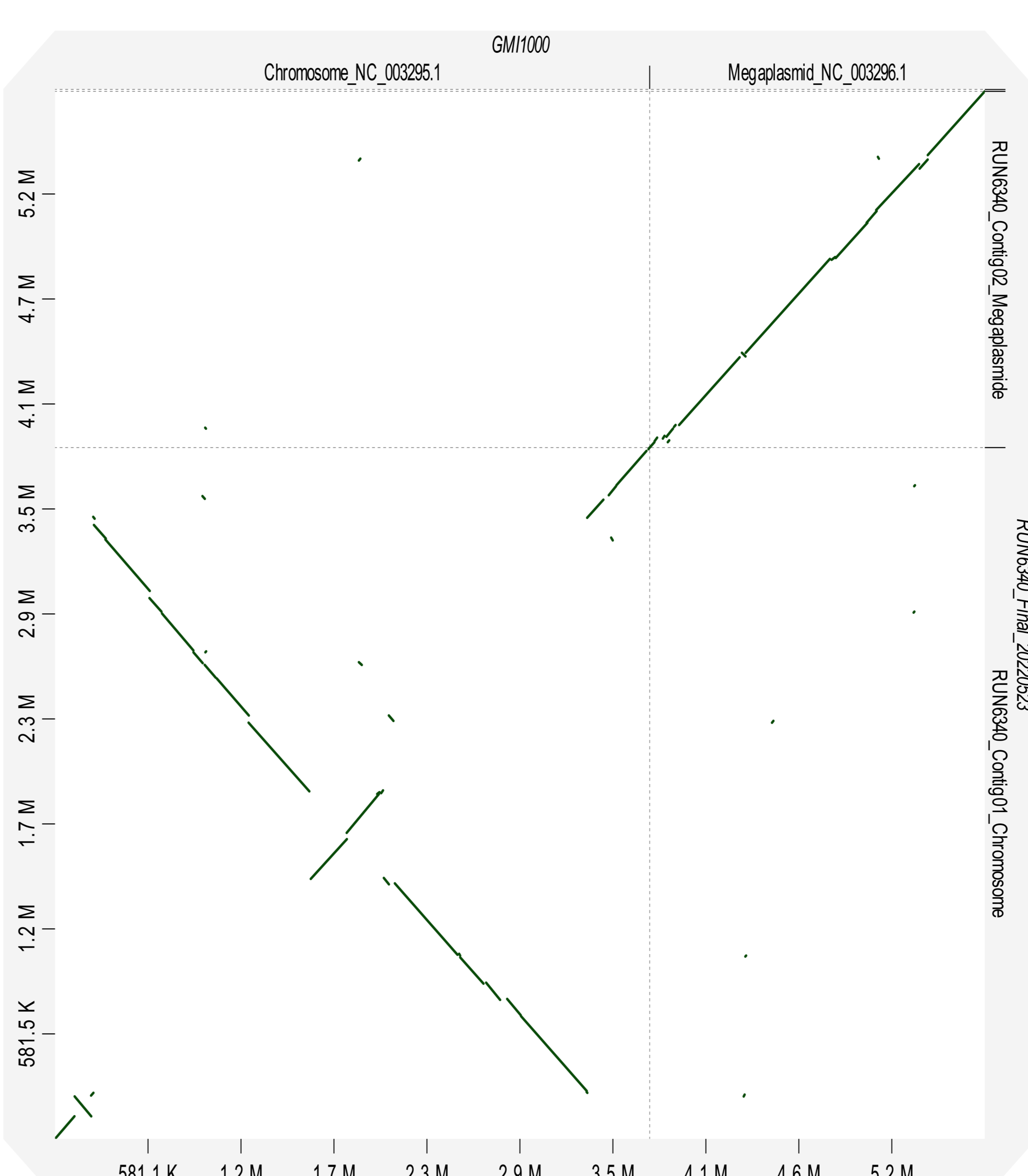
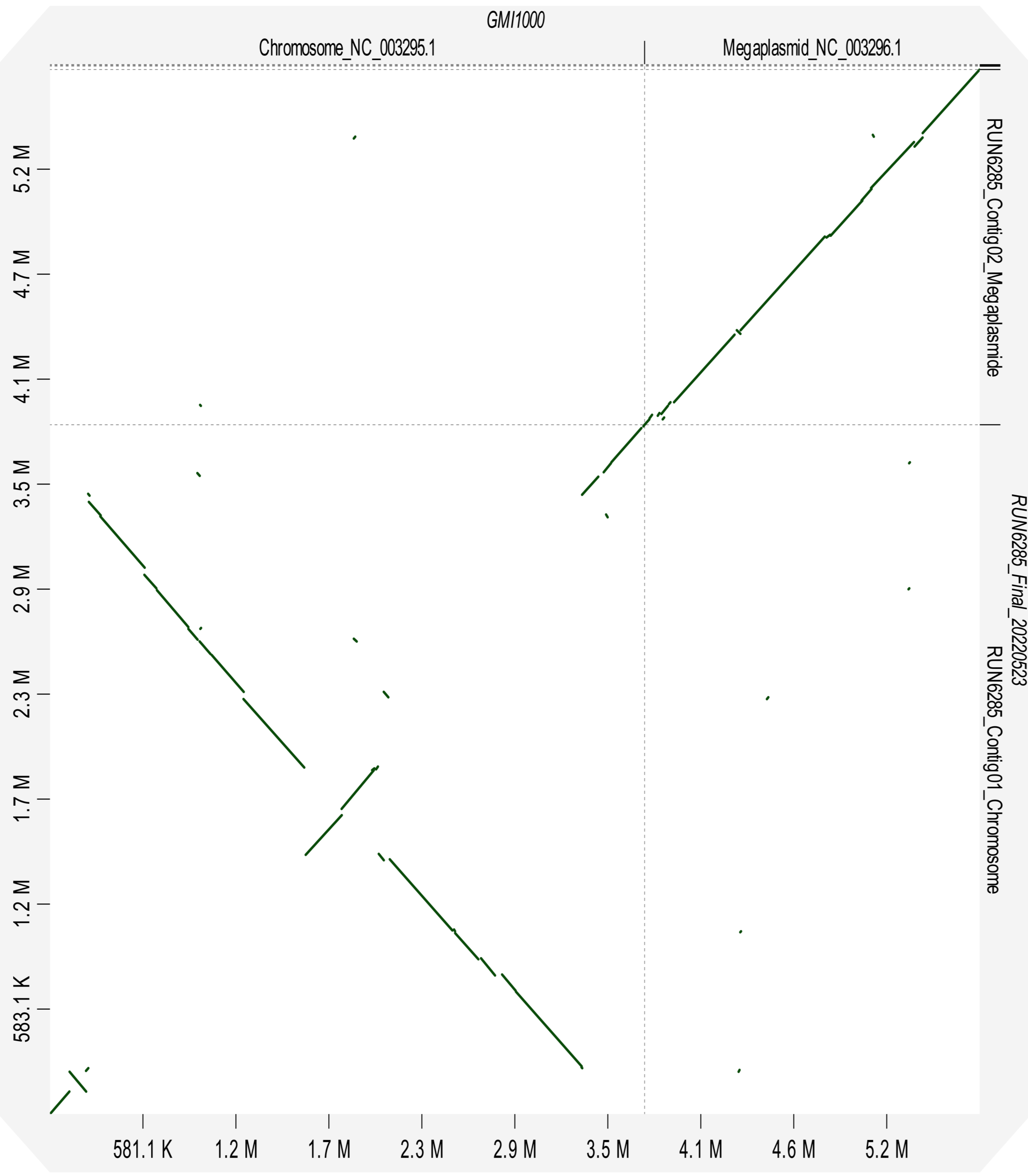
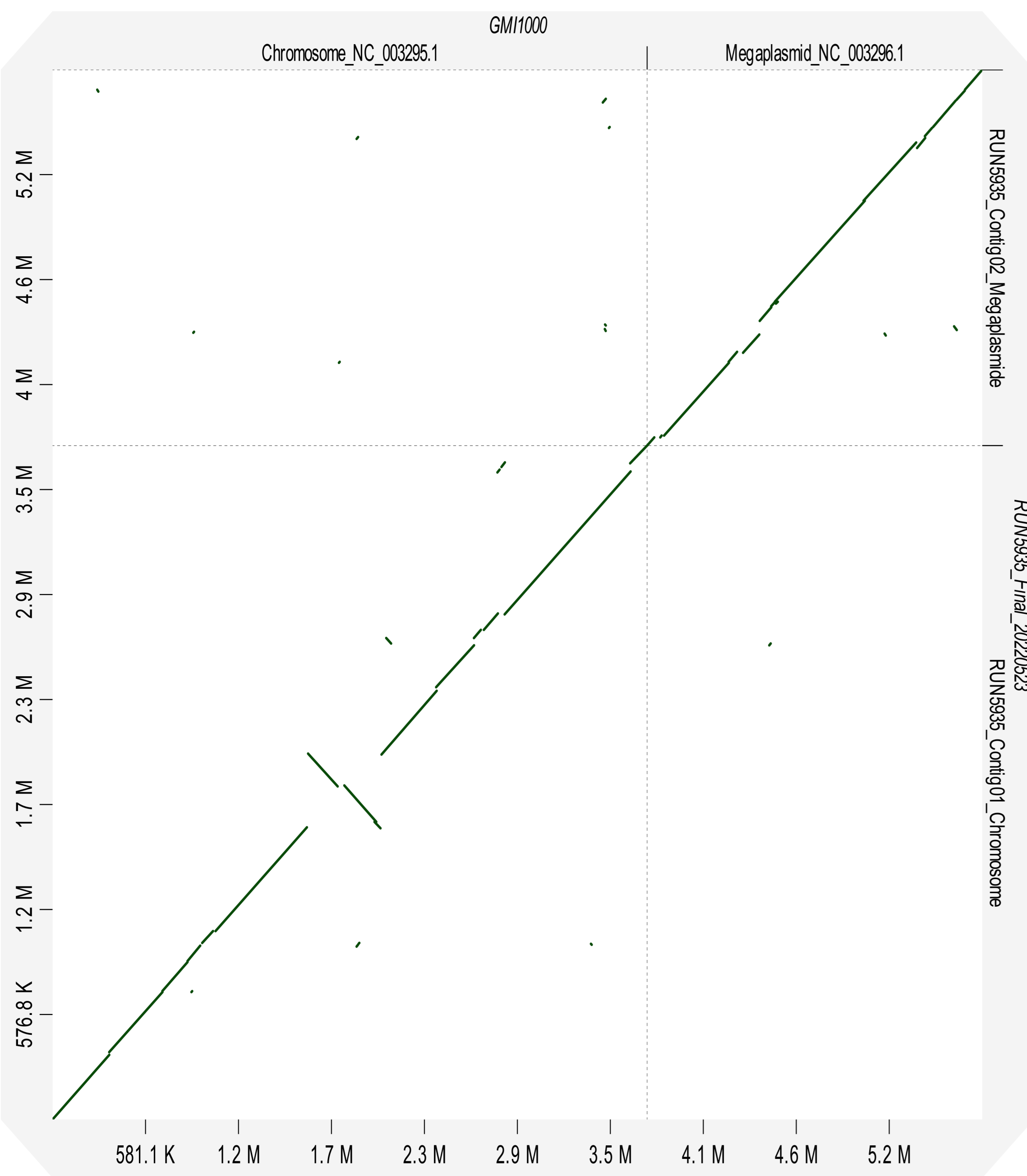
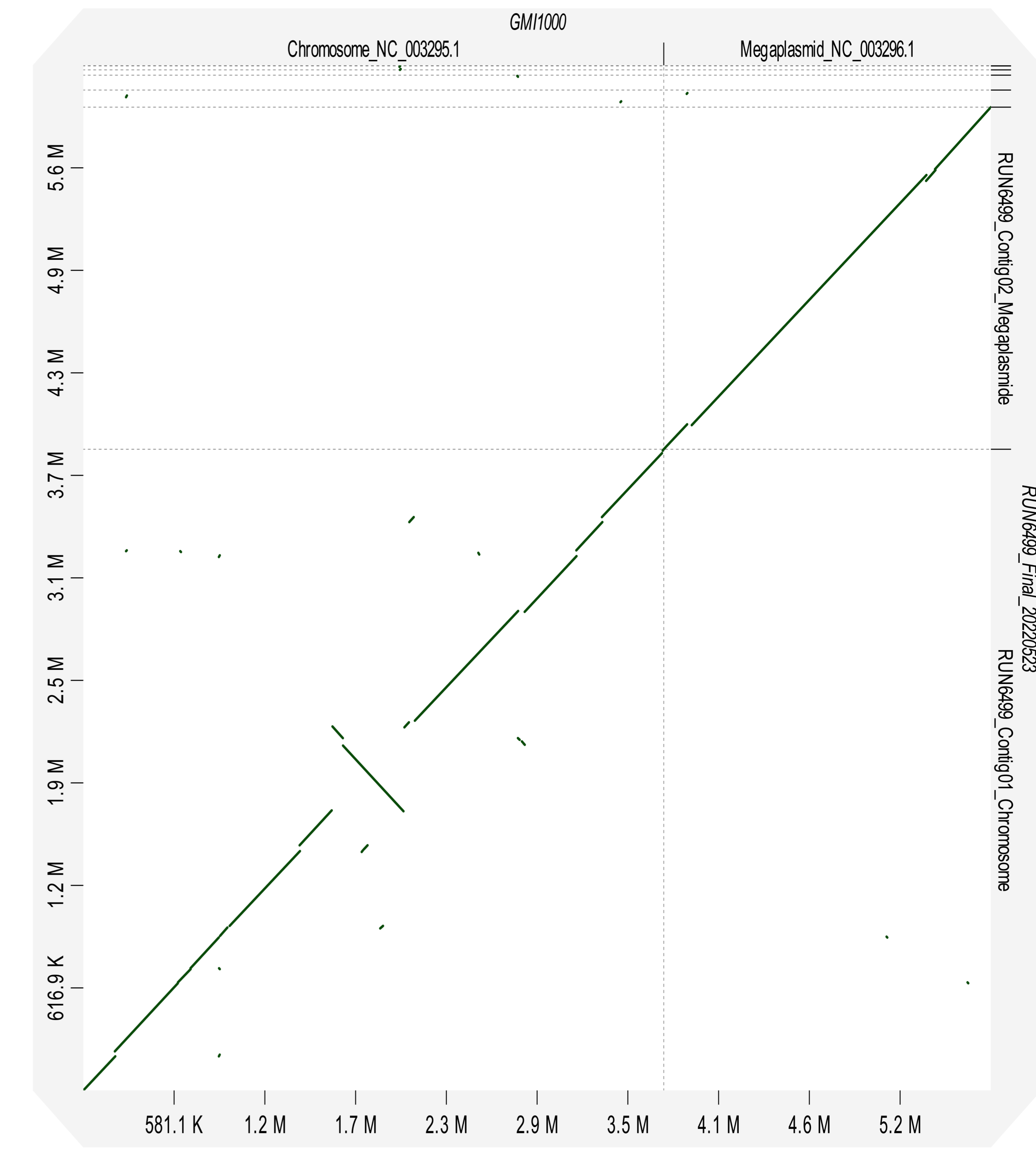
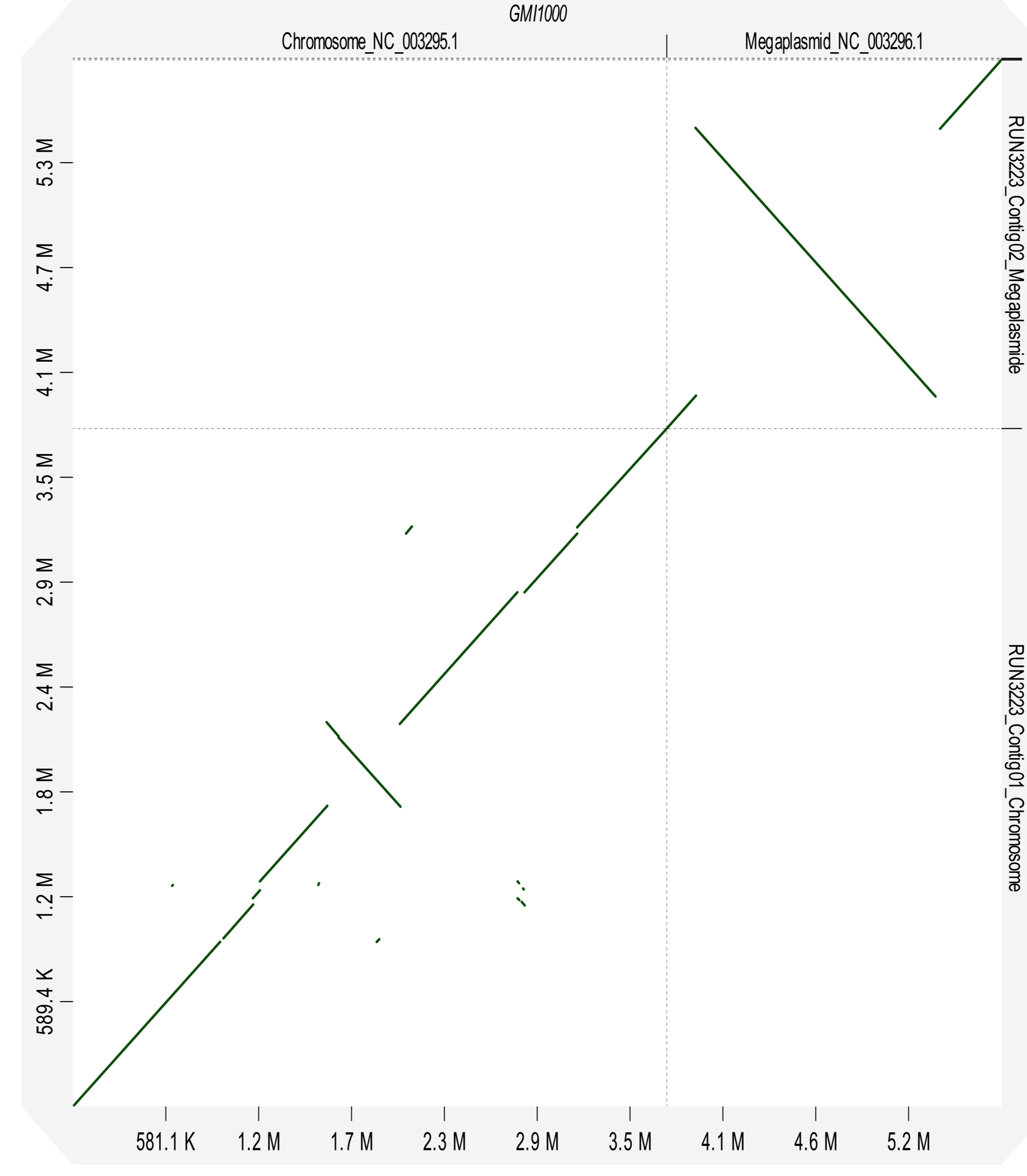
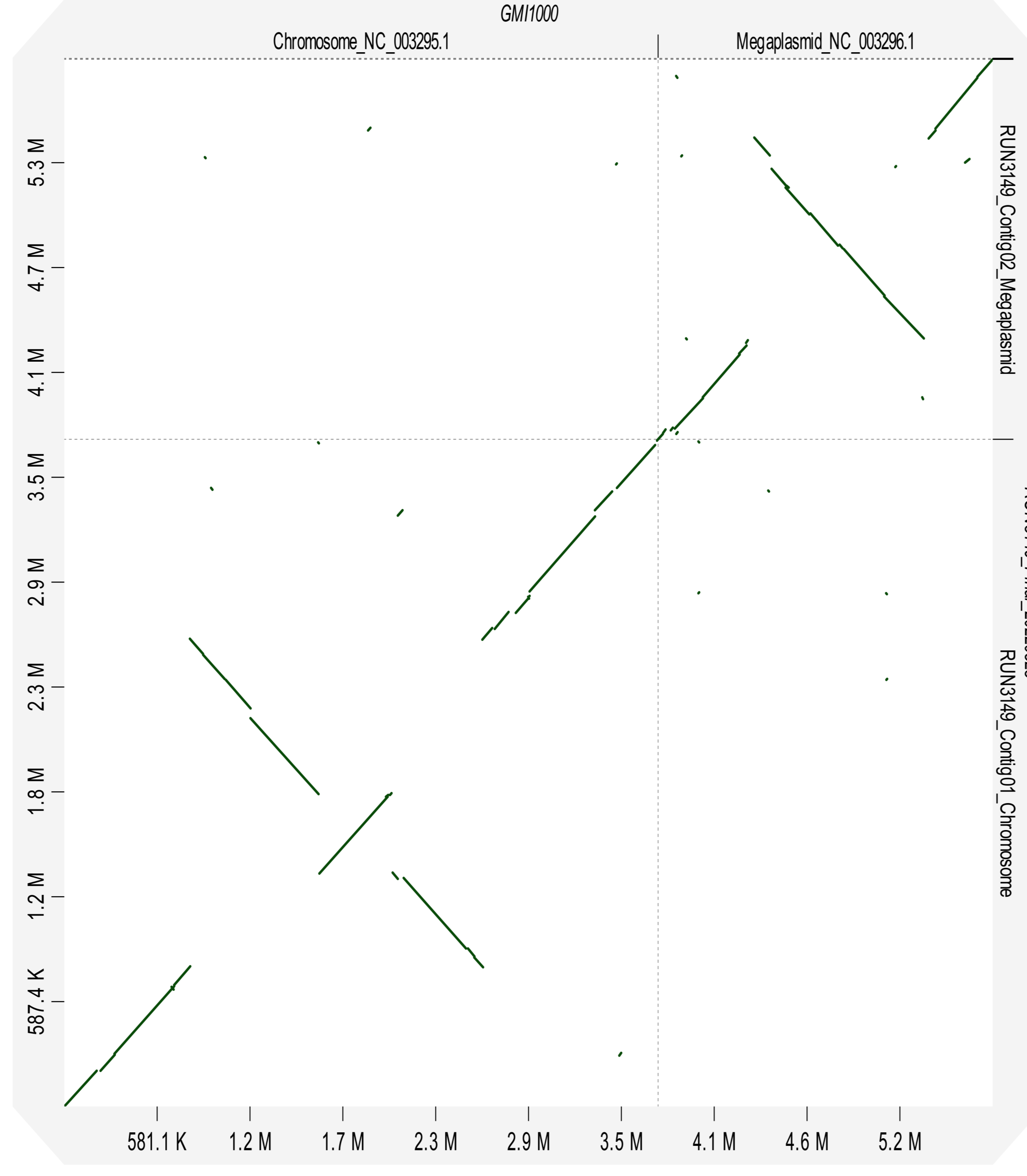
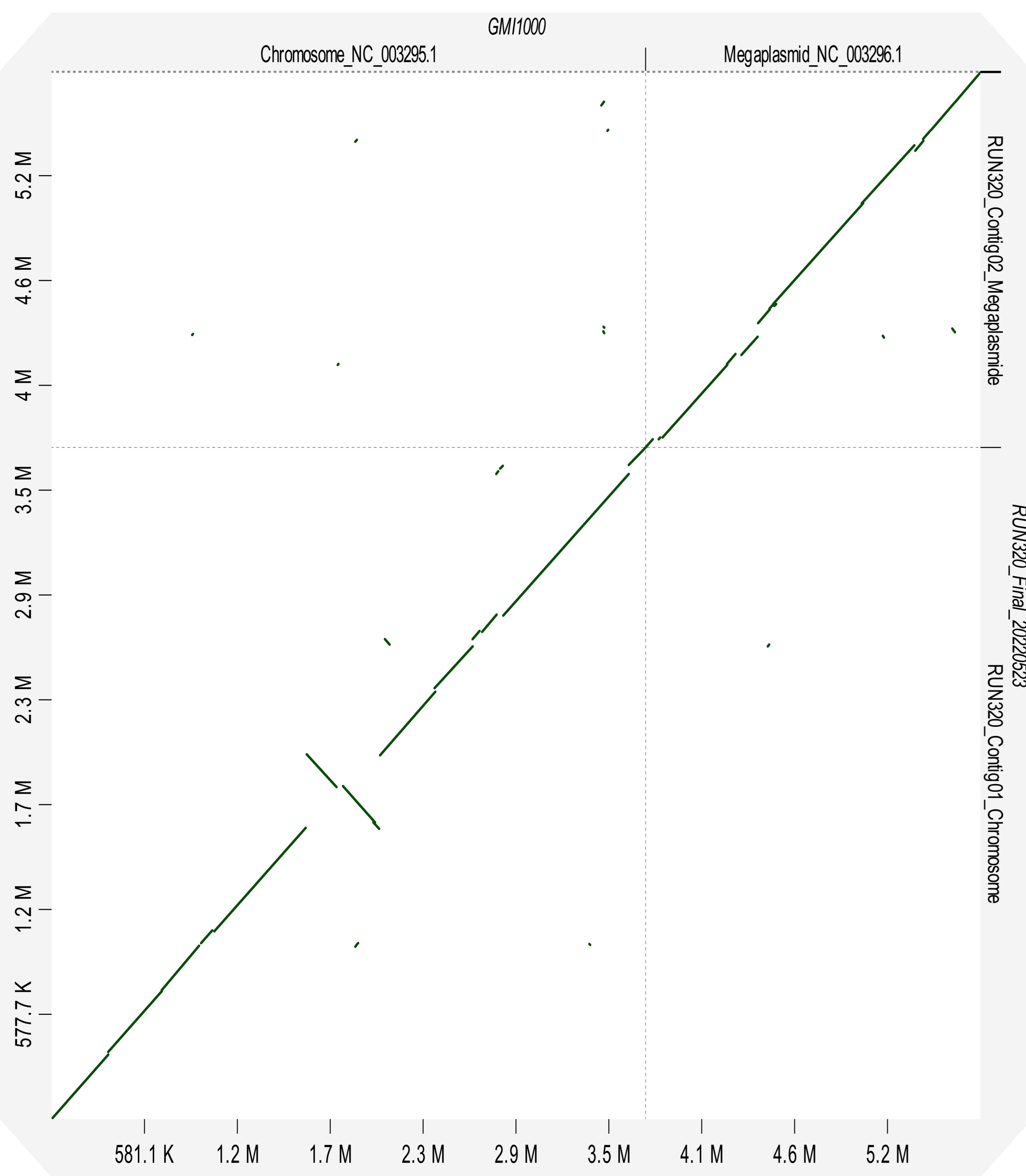
RUN.ID	COMPLETE	SINGLE	DUPLICATED	FRAGMENTED	MISSING
RUN3223	100	99.2	0.8	0	0
RUN6499	100	99.2	0.8	0	0
RUN3149	100	99.2	0.8	0	0
RUN5935	100	99.2	0.8	0	0
RUN6340	100	99.2	0.8	0	0
RUN6285	100	99.2	0.8	0	0
RUN6534	100	99.2	0.8	0	0
RUN320	100	99.2	0.8	0	0

De plus, les résultats Quast (Tableau 4) révèlent un N50 élevé >3701373 pb. Les deux contigs correspondant aux chromosome et mégaplasme sont systématiquement identifiés. Le nombre de contigs varie entre 3 et 8 correspondant respectivement aux nombres de contigs chez les génomes de RUN5935 et RUN6499.

Tableau_ 4: Résultats de l'analyse Quast

	RUN6534	RUN6499	RUN6340	RUN6285	RUN5935	RUN3149	RUN3223	RUN320
# CONTIGS (>= 0 BP)	4	8	4	7	3	5	5	5
# CONTIGS (>= 1000 BP)	4	8	4	7	3	5	5	5
# CONTIGS (>= 5000 BP)	2	6	3	3	2	2	3	3
# CONTIGS (>= 10000 BP)	2	6	3	3	2	2	2	2
# CONTIGS (>= 25000 BP)	2	5	2	2	2	2	2	2
# CONTIGS (>= 50000 BP)	2	4	2	2	2	2	2	2
TOTAL LENGTH (>= 0 BP)	5933548	6168801	5814798	5831399	5768014	5874279	5894145	5776698
TOTAL LENGTH (>= 1000 BP)	5933548	6168801	5814798	5831399	5768014	5874279	5894145	5776698
TOTAL LENGTH (>= 5000 BP)	5929020	6165101	5812640	5820736	5766768	5865677	5888929	5772325
TOTAL LENGTH (>= 10000 BP)	5929020	6165101	5812640	5820736	5766768	5865677	5881954	5767270
TOTAL LENGTH (>= 25000 BP)	5929020	6141441	5801819	5801817	5766768	5865677	5881954	5767270
TOTAL LENGTH (>= 50000 BP)	5929020	6109666	5801819	5801817	5766768	5865677	5881954	5767270
# CONTIGS	4	8	4	7	3	5	5	5
LARGEST CONTIG	3862242	3858437	3828785	3828784	3702362	3737374	3810476	3701373
TOTAL LENGTH	5933548	6168801	5814798	5831399	5768014	5874279	5894145	5776698
GC (%)	66.84	66.65	66.93	66.92	66.96	66.93	66.9	66.95
N50	3862242	3858437	3828785	3828784	3702362	3737374	3810476	3701373
# N'S PER 100 KBP	0	0	0	0	0	0	0	0

La soumission des séquences génomiques nouvellement assemblées à D-genies nous a permis de comparer visuellement nos séquences génomiques avec la séquence de GMI1000 et de montrer qu'ils présentent dans l'ensemble une forte colinéarité avec cette souche de référence du ceRs (Figure 6). Plusieurs variations structurales ont toutefois été observées. Ainsi, une inversion caractéristique de GMI1000 par rapport à nos génomes a été identifiée au niveau du chromosome entre ~ 1.7 Mb et ~ 2.3 Mb. De plus, chez la souche RUN3149 (I-33), on observe une inversion par rapport à GMI1000 de plusieurs milliers de bases au niveau du chromosome entre ~ 1.2 Mb et ~ 2.9 Mb, et au niveau du mégaplasme entre ~ 4.1 Mb et ~ 5.2 Mb. Chez les souches I-18 (RUN3223 et RUN6534), une inversion par rapport à GMI1000 est observée au niveau du mégaplasme entre ~ 4.1 Mb et ~ 5.2 Mb. En outre, les souches I-31 (RUN6285 et RUN6340) présentent au niveau du chromosome une inversion par rapport à GMI1000 entre ~ 581.1 Kb et ~ 3.5 Mb.



Figure_ 6: Alignement des séquences génomiques nouvellement assemblées à la séquence de référence du ceRs GMI1000

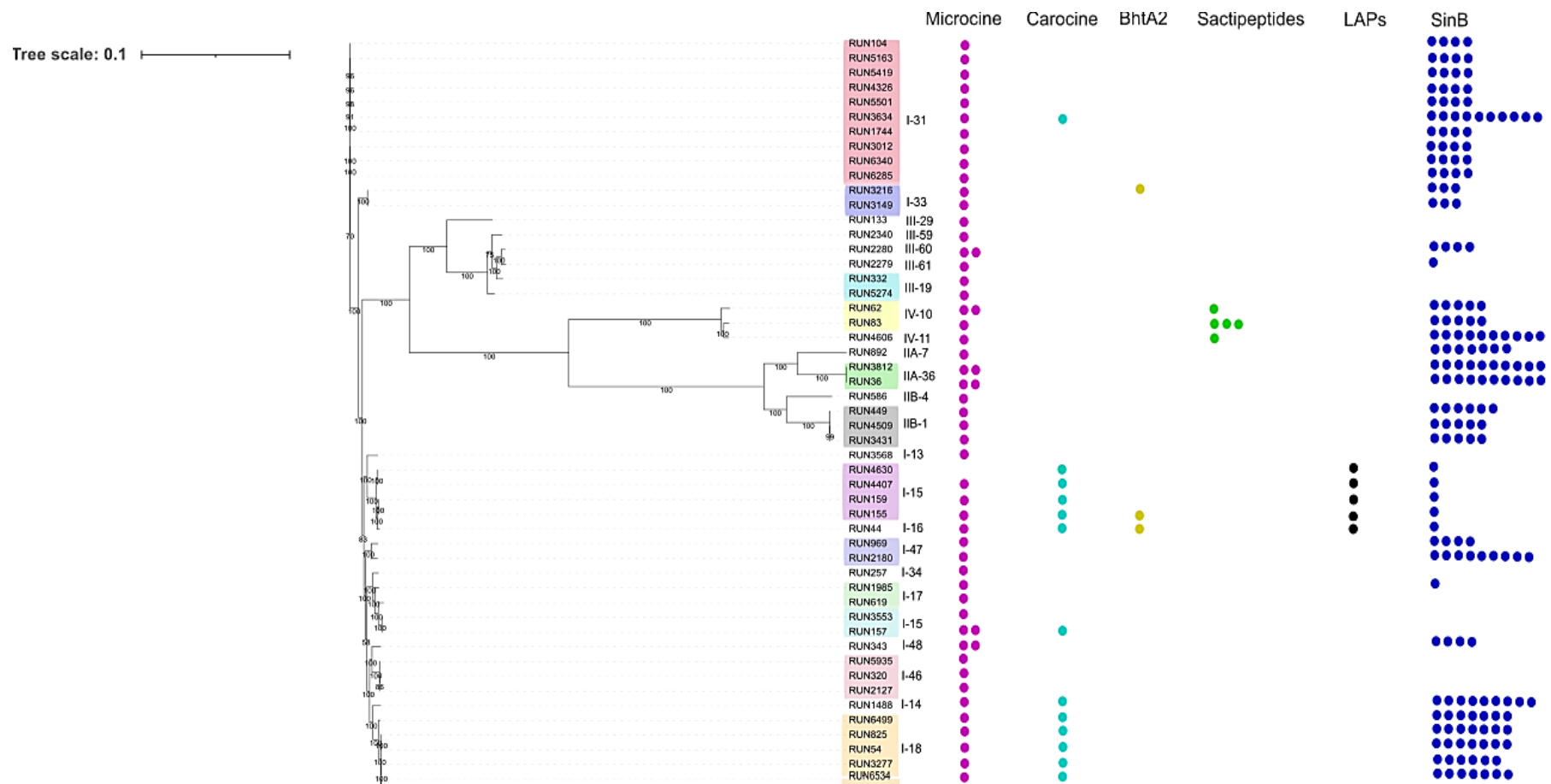
Sur l'axe des abscisses et ordonnées sont affichés en paire de base des échelles permettant de positionner approximativement les séquences respectives de GMI1000 et de nos génomes assemblés.

4. Caractérisation génomique des bactériocines chez les souches du ceRs

Cinq types de bactériocines seraient présents dans les génomes étudiés (Figure 7). La microcine, la carocine, BhtA2, les sactipeptides et les LAPs (Linear Azole containing Peptides) ont été identifiés par BAGEL4 et le gène SinB, suggéré par Huerta [220] comme étant un gène de production de bactériocines chez le ceRs, a été identifié par ncbi-blast-2.10.1+.

Remarquablement, la microcine est présente chez toutes les souches sauf RUN4630 (I-15). De plus, le gène SinB est largement distribué, car seul 10 sur les 51 souches étudiées n'en possèdent pas. Ces 10 souches sont RUN133 (III-29), RUN2340 (III-59), RUN332 et RUN5274 (III-19), RUN586 (II-4), RUN3568 (I-13), RUN257 (I-34), RUN320 - RUN5935 et RUN2127 (I-46).

En outre, on n'observe pas de corrélation spécifique entre les types de bactériocines identifiés dans les génomes et la lignée génétique des souches. En effet, la carocine est par exemple présente chez les souches RUN3634 (I-31), RUN155 - RUN157 - RUN159 - RUN3553 - RUN4407 et RUN 4630 appartenant au sequevar I-46, RUN157 (I-15), RUN1488 (I-14), RUN6499 - RUN6534 - RUN54 - RUN825 - RUN3277 appartenant au sequevar I-18. Le BhtA2 est présent chez les souches RUN3216 (I-33), RUN155 (I-15) et RUN44 (I-16). Les sactipeptides sont présents chez les souches RUN62 et RUN83 (IV-10), et la souche RUN4606 (IV-11). Les LAPs sont présents chez les souches RUN155, RUN159, RUN4407, RUN4630 qui appartiennent au sequevar I-15 et la souche RUN44 qui appartient au sequevar I-16. Par ailleurs, on n'a pas observé de corrélation évidente entre l'intensité de l'activité bactériocine et les types de bactériocines identifiés dans les génomes des souches.



Figure_ 7: Diversité des bactériocines chez le ceRs

La phylogénie des souches représentatives du ceRs a été construite via IQTree sur la base de l'alignement des SNPs présents dans le core gène⁹ des 51 génomes. Les valeurs de bootstrap de chacun des nœuds sont précisées.

⁹ Le core gène est composé des gènes communs aux génomes analysés.

5. Vers la recherche d'autres gènes candidats pouvant être impliqués dans le succès épidémiologique du I-18 à Madagascar et du I-31 dans les petites îles du SOOI

Grâce à la disponibilité des ressources génomiques: 8 génomes représentatifs de la diversité génétique du ceRs malgache, 11 génomes représentatifs des lignées génétiques présentes dans les petites îles du SOOI et le génome de référence GMI1000, nous avons initié une comparaison de leurs contenus en gène afin d'identifier *in fine* des gènes pouvant expliquer la prévalence de certaines lignées génétiques par rapport aux autres dans le SOOI.

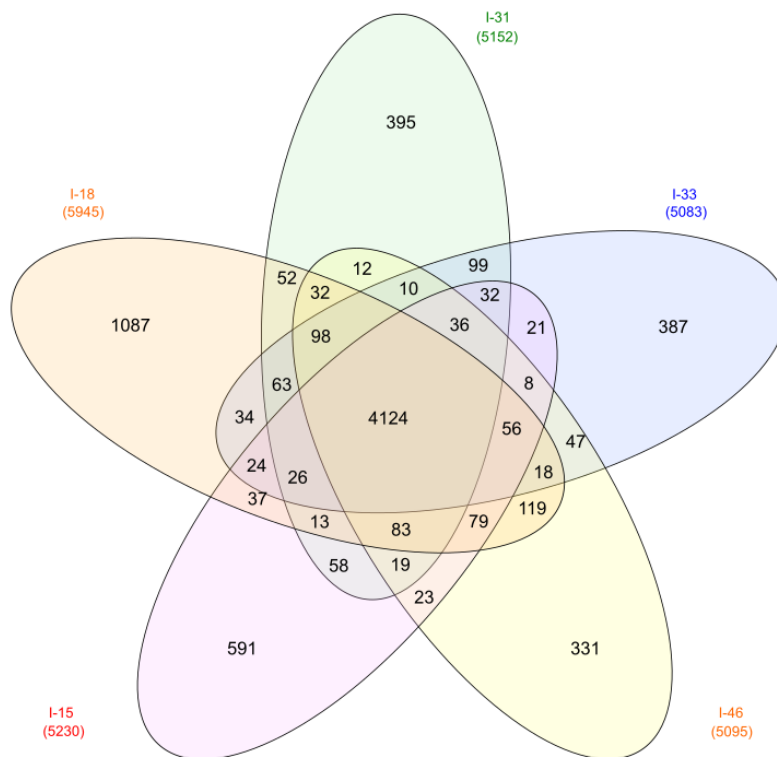
Le tableau 5 compare le nombre de SNPs entre les 20 génomes (Tableau 1). Globalement, les souches du sequevar I-18 sont proches génétiquement des souches I-46 et les souches du sequevar I-31 sont proches génétiquement des souches I-33. Au sein du sequevar I-18, la souche RUN6499 qui appartient à un complexe clonal (MCC2) différent des autres souches malgaches, se montre plus polymorphe avec > 5000 SNPs. Les 2 seules souches malgaches du sequevar I-31 identifiées jusqu'à présent sont très proches génétiquement avec 4 SNPs seulement de différence entre elles.

Tableau_ 5: Comparaison deux à deux du nombre de SNPs entre les 20 génomes du ceRs

La couleur des cellules est en fonction de l'importance du nombre de SNPs entre les génomes.

		I-15		I-18					I-31					I-33		I-46					
				MCC1			MCC2					MCC11		MCC12		MCC3					
		RUN4407	RUN4630	GMI1000	RUN0305	RUN3223	RUN3277	RUN6534	RUN6499	RUN3012	RUN4326	RUN5163	RUN5419	RUN6285	RUN6340	RUN3149	RUN3216	RUN0320	RUN320	RUN5935	RUN2127
I-15	RUN4407	0	258	26879	26834	26834	26831	26834	26915	27625	27609	27605	27608	27683	27679	28015	28017	26078	26080	26088	26306
	RUN4630	258	0	26918	26877	26877	26874	26877	26949	27658	27642	27636	27641	27716	27712	28034	28036	26109	26111	26119	26328
I-18	GMI1000	26879	26918	0	222	222	219	222	5948	26551	26536	26530	26535	26601	26597	27447	27445	19701	19699	19718	20112
	RUN0305	26834	26877	222	0	10	7	10	5826	26516	26501	26495	26500	26566	26562	27403	27401	19546	19544	19565	20093
	RUN3223	26834	26877	222	10	0	5	8	5826	26516	26501	26495	26500	26566	26562	27403	27401	19546	19544	19565	20093
	RUN3277	26831	26874	219	7	5	0	5	5823	26513	26498	26492	26497	26563	26559	27400	27398	19543	19541	19562	20090
	RUN6534	26834	26877	222	10	8	5	0	5826	26516	26501	26495	26500	26566	26562	27403	27401	19546	19544	19565	20093
	RUN6499	26915	26949	5948	5826	5826	5823	5826	0	26629	26633	26627	26632	26679	26675	27533	27531	19432	19430	19437	19642
I-31	RUN3012	27625	27658	26551	26516	26516	26513	26516	26629	0	92	88	92	84	88	20542	20546	25956	25956	25967	26187
	RUN4326	27609	27642	26536	26501	26501	26498	26501	26633	92	0	26	30	150	154	20526	20530	25929	25929	25940	26180
	RUN5163	27605	27636	26530	26495	26495	26492	26495	26627	88	26	0	26	146	150	20520	20522	25923	25923	25934	26174
	RUN5419	27608	27641	26535	26500	26500	26497	26500	26632	92	30	26	0	150	154	20526	20530	25931	25931	25942	26182
	RUN6285	27683	27716	26601	26566	26566	26563	26566	26679	84	150	146	150	0	4	20592	20596	26006	26006	26017	26237
	RUN6340	27679	27712	26597	26562	26562	26559	26562	26675	88	154	150	154	4	0	20588	20592	26002	26002	26013	26233
I-33	RUN3149	28015	28034	27447	27403	27403	27400	27403	27533	20542	20526	20520	20526	20592	20588	0	10	26578	26580	26591	26776
	RUN3216	28017	28036	27445	27401	27401	27398	27401	27531	20546	20530	20522	20530	20596	20592	10	0	26578	26578	26589	26776
I-46	RUN0320	26078	26109	19701	19546	19546	19543	19546	19432	25956	25929	25923	25931	26006	26002	26578	26578	0	4	42	565
	RUN320	26080	26111	19699	19544	19544	19541	19544	19430	25956	25929	25923	25931	26006	26002	26580	26578	4	0	42	567
	RUN5935	26088	26119	19718	19565	19565	19562	19565	19437	25967	25940	25934	25942	26017	26013	26591	26589	42	42	0	547
	RUN2127	26306	26328	20112	20093	20093	20090	20093	19642	26187	26180	26174	26182	26237	26233	26776	26776	565	567	547	0

L'analyse réalisée avec Roary montre que le pangénome de ces 20 souches est composé de 8014 familles de gènes. Parmi ces gènes, 4080 sont communs aux cinq lignées génétiques et constituent le core génome. Un diagramme de Venn a alors été construit pour illustrer les résultats de la présence et absence de ces 8014 familles de gènes identifiés chez les 20 génomes (Figure 8). Les souches de la lignée I-18 ont le plus de gènes spécifiques avec 1087 gènes soit 18.28% des gènes identifiés chez cette lignée. Par opposition, les souches de la lignée I-46 présentent le moins de gènes spécifiques avec 331 gènes soit 6.49% des gènes identifiés chez cette lignée.



Figure_ 8: Diagramme de Venn représentant la distribution des 8014 familles de gènes du pangénome entre les différentes lignées génétiques

Par ailleurs, nous avons comparé les effecteurs de type III présents chez les 20 génomes étant donné que c'est un facteur de virulence important chez le ceRs (Tableau 6). Un total de 88 effecteurs a été identifié dans au moins un des génomes étudiés. Certains effecteurs semblent être spécifiques de certaines lignées génétiques. En effet, les effecteurs RipBJ, RipAX1 et RipBO sont présents uniquement chez les souches I-18 et les effecteurs RipAL, RipBE, RS_T3E_Hyp18, et RipBQ sont présents uniquement chez les souches I-15. Cependant, chez les lignées I-31, I-33 et I-46 aucun effecteur ne semble être spécifique.

	RUN4407	RUN4630	RUN305	RUN3223	RUN3277	RUN6499	RUN6534	GMI1000	RUN3012	RUN4326	RUN5163	RUN5419	RUN6285	RUN6340	RUN3149	RUN3216	RUN320	RUN320 PacBio	RUN2127	RUN5935	
	I-15		I-18					I-31						I-33		I-46					
RipTAL	[Redacted]																				
RipX	[Redacted]																				
RipZ	[Redacted]																				
RipAK	[Redacted]																				
RipAO	[Redacted]																				
RipAU	[Redacted]																				
RipAV	[Redacted]																				
RipM	[Redacted]																				
RipW	[Redacted]																				
RipO1	[Redacted]																				
RipAS	[Redacted]																				
RipV1	[Redacted]																				
RipE1	[Redacted]																				
RipR	[Redacted]																				
RipI	[Redacted]																				
RipAQ	[Redacted]																				
RipAM	[Redacted]																				
RipAN	[Redacted]																				
RipQ	[Redacted]																				
RipG2	[Redacted]																				
RipD	[Redacted]																				
RipH1	[Redacted]																				
RipTPS	[Redacted]																				
RipS3	[Redacted]																				
RipAG	[Redacted]																				
RipU	[Redacted]																				

	RUN4407	RUN4630	RUN305	RUN3223	RUN3277	RUN6499	RUN6534	GMI1000	RUN3012	RUN4326	RUN5163	RUN5419	RUN6285	RUN6340	RUN3149	RUN3216	RUN320	RUN320 PacBio	RUN2127	RUN5935	
	I-15		I-18					I-31						I-33		I-46					
RipY																					
RipH3																					
RipAR																					
RipS6																					
RipAZ1																					
RipP1																					
RipG6																					
RipP3																					
RipA5																					
RipAX2																					
RipH2																					
RipA2																					
RipS5																					
RipP2																					
RipS2																					
RipG3																					
RS_T3E_Hyp7																					
RipA4																					
RipS1																					
RipS8																					
RipS4																					
RS_T3E_Hyp6																					
RipG4																					
RipA3																					
RipG1																					
RipE2																					

	RUN4407	RUN4630	RUN305	RUN3223	RUN3277	RUN6499	RUN6534	GMI1000	RUN3012	RUN4326	RUN5163	RUN5419	RUN6285	RUN6340	RUN3149	RUN3216	RUN320	RUN320 PacBio	RUN2127	RUN5935
	I-15		I-18					I-31							I-33		I-46			
RipG5																				
RipBD																				
RS_T3E_Hyp14																				
RipBN																				
RipBJ																				
RipAX1																				
RipAH																				
RipBO																				
RipC2																				
RipAL																				
RipBE																				
RipG7																				
RS_T3E_Hyp18																				
RipBQ																				
RS_T3E_Hyp4																				
RipBI																				
RS_T3E_Hyp8																				
RipBG																				
RipF2																				
RS_T3E_Hyp17																				
RipAF2																				
RipAZ2																				
RipBB																				
RipBH																				
RS_T3E_Hyp10																				
RipV2																				

RUN4407	RUN4630	RUN305	RUN3223	RUN3277	RUN6499	RUN6534	GMI1000	RUN3012	RUN4326	RUN5163	RUN5419	RUN6285	RUN6340	RUN3149	RUN3216	RUN320	RUN320 PacBio	RUN2127	RUN5935
I-15		I-18					I-31					I-33		I-46					

RS_T3E_Hyp1	
RipBC	
RS_T3E_Hyp2	
RS_T3E_Hyp9	
RS_T3E_Hyp5	
RipO2	
RipH4	
RipG8	
RS_T3E_Hyp11	
RipBF	
RipBL	
RipBP	
RipBK	
RS_T3E_Hyp13	
RipS7	
RS_T3E_Hyp12	
RipK	
RS_T3E_Hyp3	

Discussion

1. Mise au point d'un protocole d'extraction d'ADN génomique à haut poids moléculaire chez le ceRs

Les progrès technologiques ont entraîné un changement rapide des stratégies de séquençage du génome [224]. De plus en plus, les technologies de séquençage deuxième génération de l'ADN, comme Illumina et Ion Torrent, sont supplantées par les technologies de troisième génération, comme Pacific Biosciences (PacBio), qui fournissent des fragments d'ADN plus long (supérieurs à 10 kb) [224]. En effet, ces long fragments d'ADN sont plus faciles à assembler, car ils peuvent couvrir les régions répétées. Or, ces technologies de troisième génération ont besoin en input de séquencer de l'ADN génomique pur et peu fragmenté, soit de haut poids moléculaire. Des kits commerciaux pour l'extraction d'ADN à haut poids moléculaire sont alors disponibles, mais ils sont coûteux [225]. De plus, ils ne sont pas toujours adaptés au ceRs qui produit des exopolysaccharides bouchant les colonnes utilisées pour l'extraction d'ADN. D'où l'intérêt de mettre au point, au cours de cette étude, un protocole d'extraction d'ADN génomique à haut poids moléculaire, peu coûteux et adapté aux souches du ceRs. Pour cela, nous nous sommes basés sur un protocole d'extraction développé par Wilson [211]. Les tests de la qualité et de la quantité d'ADN génomiques extraits nous montrent une bonne intégrité de l'ADN des 8 souches étudiées dont la taille varie entre 42 à 50 kb. De plus, la solution d'ADN obtenue est pure avec des ratios A260/A280 compris entre 1.85 et 1.9 et A260/A230 compris entre 1.8 et 2.08. Ce qui répond aux exigences demandées pour un séquençage Pacific Biosciences (PacBio). En outre, grâce à notre protocole d'extraction, nous avons pu extraire entre 18.59 à 41.2 µg d'ADN.

2. Qu'est-ce qui pourrait expliquer le succès épidémiologique des lignées génétiques I-18 à Madagascar et I-31 dans les petites îles du SOOI?

Au cours de cette étude, l'activité bactériocine de 37 souches productrices a été testée sur 99 souches indicatrices. Ces souches sont choisies pour représenter la diversité génétique du ceRs à Madagascar, dans les petites îles du SOOI et dans les autres régions du monde. Les résultats de typage bactériocine montrent une activité bactériocine et une immunité dépendante du sequevar des souches, voir plus finement encore, car certaines souches de I-18 arrivent à inhiber d'autres souches du même sequevar. Les tests d'activité inhibitrice

ont permis d'établir un schéma de typage bactériocines des souches du ceRs rejoignant ceux qui ont déjà été mis en évidence dans les études précédentes [96,207,226]. De plus, la recherche *in silico* des clusters de gènes codant potentiellement les bactériocines a montré aucune corrélation évidente entre l'intensité de l'activité bactériocine et les types de bactériocines identifiés dans les génomes des souches. Cette constatation pourrait indiquer qu'un nouveau type de bactériocine non encore décrite dans la littérature est à l'origine des activités inhibitrices, d'où la nécessité de mener une caractérisation biochimique. Cependant une autre explication pourrait résider dans la régulation de l'expression des gènes codant ces bactériocines, suggérant le besoin d'effectuer une analyse transcriptomique de l'expression de ces gènes.

La description des situations épidémiologiques à Madagascar et dans les petites îles du SOOI ouvre sur de nombreuses questions : pourquoi est-ce que le sequevar I-18 est fortement prévalent à Madagascar et non dans les petites îles du SOOI ? Pourquoi est-ce que seulement deux souches du sequevar I-31 y ont été détectées malgré sa large dissémination dans les îles voisines (SOOI) et au-delà en Afrique ? La littérature met en évidence l'avantage que procure la production de bactériocine dans la compétition pour l'espace et les ressources entre des lignées génétiques apparentées [138,144,76,145,139]. Ces études renforcent l'intérêt de tester l'activité bactériocine des souches du ceRs collectées à Madagascar et dans les petites îles du SOOI. Nos résultats ont montré une cohérence entre leurs activités bactériocines et la situation épidémiologique dans la zone. En effet, les souches du sequevar I-18 et I-31 respectivement prévalent à Madagascar et dans les petites îles du SOOI ont une activité antimicrobienne contre les lignées génétiques présentes dans la zone, ce qui devrait leur fournir un avantage compétitif. De plus, les souches du sequevar I-18 ont une activité antimicrobienne forte contre les souches du sequevar I-31, ce qui pourrait apporter un élément de réponse dans la non prévalence du sequevar I-31 à Madagascar. *À contrario*, les souches du sequevar I-18 ne se sont pas largement disséminées dans les petites îles du SOOI et il reste étonnant que seules deux souches du sequevar I-31 soient présentes à Madagascar malgré une large prospection. Ces faits nous font ainsi penser que la production de bactériocine seule ne pourrait expliquer la situation épidémiologique observée dans le SOOI. Soit ces deux lignées génétiques n'ont pas encore eu le temps de se disséminer, le I-18 dans les petites îles du SOOI et le I-31 à Madagascar, soit d'autres facteurs encore à déterminer entrent en

jeu. La disponibilité des séquences génomiques de 20 souches représentatives de la diversité génétique du ceRs à Madagascar, dans le SOOI et de la souche de référence GMI1000 nous a permis d'initier une exploration de leur contenu en gènes. Un grand nombre de SNPs a été identifié en comparant deux à deux les génomes. Ce résultat pourra servir de base pour la recherche de marqueurs moléculaires permettant de différencier les souches appartenant aux différentes lignées génétiques. De plus, grâce à la liste des gènes présents et absents dans les 20 génomes fournis par l'analyse pangénomique, une analyse de génétique fonctionnelle des gènes qui seraient présents chez les lignées génétiques prévalentes, mais absents chez les lignées non prévalentes pourrait aboutir à l'identification de gènes à l'origine du succès épidémiologique des lignées génétiques les plus répandues. En outre, l'hypothèse d'une meilleure adaptation des souches de la lignée I-18 malgache aux variétés cultivées et aux systèmes de culture dans ce pays devra également être explorée.

Conclusion générale et perspectives

Depuis l'aube de l'agriculture, les ennemis des cultures ont été un fléau pour l'humanité [227]. La FAO estime qu'ils sont responsables chaque année de la perte de près de 40% des cultures alimentaires mondiales et des pertes commerciales en produits agricoles dépassant 220 milliards de dollars par an [228]. Dans un monde devant faire face à des défis de sécurité alimentaire, de lutte contre la pauvreté, et de changement climatique, les sujets liés à l'épidémiosurveillance et au développement de stratégies de prévention et de lutte contre les ennemis des cultures sont plus que jamais d'actualité. En pathologie végétale, une épidémie correspond à une maladie infectieuse provoquée par un agent pathogène s'exprimant de manière soudaine et inhabituelle sur une population végétale, et se propageant rapidement dans une zone géographique donnée [13]. Pour éviter ou gérer ces épidémies, la mise en place d'un réseau d'épidémiosurveillance s'avère nécessaire. Dans un tel réseau, les différents partenaires vont interagir pour la collecte, l'évaluation et la dissémination des informations afin de prévenir l'arrivée et/ou le déclenchement d'une maladie et de gérer le mieux possible une épidémie déjà présente dans un territoire donné. Les méthodes de typage permettant de différencier les isolats bactériens de la même espèce et de tracer leurs origines et propagations sont essentielles pour la prévention, le suivi et le contrôle des infections [229]. En effet, il est désormais bien connu qu'au niveau des espèces bactériennes, la diversité génétique est suffisante pour permettre l'identification de différents clones à l'origine des épidémies [230].

Le développement de graves épidémies de flétrissement bactérien dans les bassins de production de pomme de terre, situées dans les Hautes Terres Centrales malgaches (1000 - 2000 m d'altitude), a attiré l'attention sur la nécessité d'y mener une étude épidémiologique entre 2013 et 2016. Cette étude a mis en évidence une large dissémination du phylotype IIB-1 (représentant 72% des 1224 souches collectées) probablement due aux tubercules de pomme de terre infectées; suivie du phylotype III (18%); et du phylotype I (10 %) qui n'a pas pu être analysé lors de ces travaux à cause de l'absence de schéma de génotypage spécifique à ces souches [12]. La première démarche de cette thèse a alors été de mettre au point une méthode de génotypage spécifique aux souches du phylotype I (**chapitre 1**). Ensuite, nous avons décrit la situation

épidémiologique du ceRs à Madagascar en étendant les zones de collectes sur 5 régions situées dans le Nord-Ouest et l'Est (**chapitre 1**). Enfin, nous avons exploré en particulier l'influence des bactériocines, des substances antimicrobiennes, dans la situation épidémiologique observée à Madagascar et dans les petites îles du SOOI (**chapitre 1-2**) et globalement nous avons voulu étudier la capacité de production de bactériocines des souches du ceRs (**chapitre 2**).

1. Pertinence du schéma RS1-MLVA14 pour l'étude épidémiologique du phylotype I à Madagascar

Pour mettre en place un réseau d'épidémiosurveillance du ceRs, le génotypage des souches est nécessaire, car il permet l'identification précise et la discrimination des populations bactériennes. Diverses méthodes de typage moléculaire sont disponibles pour différencier entre eux les isolats bactériens du phylotype I, mais nous avons choisi de développer la méthode MLVA. Celle-ci se base sur les différences entre le nombre de répétitions des "tandem repeat" (TR) qui sont de petites séquences nucléotidiques répétées en tandem au niveau d'un à une dizaine de loci. Une simple recherche des mots clés "MLVA + typing" sur google scholar a montré 11800 résultats correspondant à des articles et publications scientifiques en lien à ces mots clés dont 8200 sont sorties ces 10 dernières années. Ce qui montre un intérêt de la communauté scientifique dans cette méthode de génotypage.

Le pouvoir de discrimination (Indice de discrimination de Hunter et Gaston, HGDI) du RS1-MLVA14 a été comparé avec celui du schéma MLST-7, qui est un outil moléculaire de typage basé sur le séquençage de 6 gènes de ménage (*gdhA*, *mutS*, *adK*, *leuS*, *rplB*, *gyrB*) et d'un gène associé à la virulence (*egl*) [231]. RS1-MLVA14 qui a une valeur HGDI de 0.793 est plus résolutif que le MLST-7 (HGDI = 0,574) révélant ainsi une plus grande diversité génétique et haplotypique. Ce qui souligne l'intérêt de ce nouveau schéma MLVA. Selon les liens génétiques entre les haplotypes, des complexes clonaux (CC) ont pu être définis en regroupant les haplotypes différents au niveau d'un seul locus. Au centre de chaque complexe clonal se trouve l'haplotype fondateur dont est issue les haplotypes du CC [232]. Les CC ont permis d'identifier les profils épidémiologiques du phylotype I malgache, les types de souches impliquées dans l'épidémie et les liens génétiques avec les souches collectées dans les petites îles du SOOI et le reste du monde. De plus, le schéma RS1-

MLVA14 a montré des liens de parenté intraspécifique entre les souches, car on a observé une cohérence entre les CC et leurs sequevars. Cependant, 9.33% des souches n'ont pas amplifié sur un VNTR et 4.66% des souches ont même été retirées de l'analyse, car elles n'ont pas amplifié sur deux VNTR et plus. Ces problèmes de non amplification des souches utilisées dans cette étude pourraient s'expliquer par la présence des mutations ponctuelles au niveau des sites d'hybridation des amorces [233,234]. Ce nouveau schéma de génotypage du phylotype I, RS1-MLVA14, complète les deux autres schémas de génotypage MLVA, RS2-MLVA9 pour les souches du phylotype II et RS3-MLVA16 pour les souches du phylotype III [12].

2. Situation épidémiologique du ceRs à Madagascar et liens génétiques avec les souches internationales

À Madagascar, l'analyse phylogénétique des souches issues des collectes de 2006 - 2013 - 2018 - 2019 révèle une forte prévalence du sequevar I-18 (76.67%) suivie du sequevar I-46 (11.39%), du sequevar I-33 (1.39%), et du sequevar I-31 qui n'est représenté que par deux souches. Nos analyses de diversité génétique ont renforcé l'observation faite par Rasoamanana et al., 2020 [223] à savoir une plus grande diversité dans la province de Toamasina ($H_{nb} = 0.48$) et une possible influence des ports maritimes dans l'entrée de matériel végétal contaminé.

La construction des réseaux d'haplotype a permis de les rassembler selon leurs proximités génétiques en complexe clonal (CC). Au total, 12 CC malgaches ont été définis dont trois majeurs, car rassemblant 54.64% des haplotypes. Les complexes clonaux majeurs MCC1 ($n = 270$, 28 haplotypes, sequevar I-18) et MCC2 ($n = 83$, 19 haplotypes, sequevar I-18) présentent une structure en étoile (avec plusieurs liens de SLV (Single Locus Variant) qui divergent de l'haplotype présumé fondateur) caractéristique d'une épidémie à expansion clonale. De plus, les complexes clonaux majeurs MCC1, MCC2, MCC3 ($n = 18$, 6 haplotypes, sequevar I-46) et le complexe clonal malgache mineur MCC12 ($n = 2$, 2 haplotypes, sequevar I-33) forment des CC avec des haplotypes internationaux issues des petites îles du SOOI (La Réunion, Mayotte, Maurice, Rodrigues) et d'Afrique (Côte d'Ivoire et Afrique du Sud). Ces liens génétiques entre les CC malgaches et les haplotypes internationaux sont probablement dus aux échanges de matériel végétale infecté entre Madagascar et ces pays

(www.tresor.economie.gouv.fr/se/madagascar/, www.lexpress.mu/), [235–237].

Les CC mineurs MCC4 (sequevar I-46), MCC5 (sequevar I-18), MCC6 (sequevar I-18 et I-46), MCC7 (sequevar I-18), MCC8 (sequevar I-18), MCC9 (sequevar I-46), MCC10 (sequevar I-46), MCC11 (sequevar I-31) quant à eux ne partagent pas de CC avec des haplotypes internationaux. Cependant, le MCC11 ne diffère que de deux locus du CC le plus prévalent dans les petites îles du SOOI (sequevar I-31) soulignant un lien génétique entre les I-31 malgaches et ceux des petites îles du SOOI [238].

L'étude faite au cours de cette thèse vient en complément de celle réalisée par Ravelomanantsoa sur les phylotypes IIB-1 et III dans les Hautes Terres Centrales entre 2013 - 2016. Les souches malgaches de IIB-1 ont été génotypées avec le schéma RS2-MLVA9 [51] et a révélé une faible diversité génétique ($H_{nb} = 0.20$) reflétant un caractère épidémique des populations et la transmission par tubercules a été suggérée, car la dispersion des souches IIB-1 dans les zones agroécologiques suit le flux de distribution de semences de tubercules [12]. Le génotypage des souches malgaches du phylotype III, avec le schéma RS3-MLVA16 développé lors de son étude, a révélé une forte diversité ($H_{nb} = 0.33$) reflétant un caractère endémique des populations et l'absence de transmission par tubercules a été suggéré car les haplotypes n'étaient pas partagés entre les zones agroécologiques [12]. Pour le cas du phylotype I étudié lors de cette thèse, l'analyse de la répartition géographique des haplotypes montre que 10 haplotypes du MCC1 sont partagés dans 21 parcelles différentes localisées dans les provinces d'Antananarivo (8 parcelles), Toamasina (6 parcelles), et Mahajanga (7 parcelles). Onze haplotypes du MCC2 sont répartis dans 9 parcelles localisées dans la province de Mahajanga. Deux haplotypes du MCC3 sont distribués dans 5 parcelles localisées dans la province de Toamasina. Ces observations suggèrent fortement une transmission de ces haplotypes à partir de matériel végétal contaminé. En effet, l'accès aux semences certifiées est difficile à Madagascar entraînant une autoproduction ou l'achat des semences contaminées entre producteurs. Les échanges de produits agricoles entre les provinces de Madagascar pourraient également favoriser la dispersion des haplotypes d'une province à l'autre.

Pour une étude exhaustive de l'épidémiologie du ceRs à Madagascar, un effort particulier a été fait pour couvrir les principaux bassins de production maraîchers où sévissent le flétrissement bactérien. Vu la localisation géographique des parcelles de collecte (au

centre, dans le Nord-Ouest et à l'Est), le lecteur peut se poser des questions sur l'absence d'échantillonnage provenant du sud de Madagascar. Des missions de collectes ont bien été réalisées dans cette zone aux mois de juillet et août 2021 permettant la collecte de 790 échantillons de tiges flétries. Après la collecte, ces échantillons ont été conservés sous forme de macérats de tige et stockés à 4°C. Un délai de 1 à 2 mois entre la collecte et le traitement de ces échantillons à La Réunion a cependant entraîné soit la mort des bactéries soit leurs changements d'état en bactéries viables, mais non cultivables, car aucune colonie caractéristique du ceRs ne s'est développée sur le milieu sélectif Sequeira. Cette constatation nous a conduit à réaliser malgré tout des PCR multiplex afin de déterminer la présence ou non du ceRs dans les macérats et leurs phylotypes [188]. De plus, nous avons utilisé deux outils LAMP (amplification isotherme médiée par les boucles) développés spécifiquement pour détecter des souches du ceRs et du phylotype IIB-1 afin de confirmer les résultats obtenus par PCR multiplex. Au total, 98 échantillons ont été positifs dont 10 appartenant au phylotype I (2 parcelles), 50 souches au phylotype II (5 parcelles), 38 souches au phylotype III (4 parcelles). Ces résultats renforcent l'intérêt de reprogrammer des missions de collecte dans le sud de Madagascar, afin d'élucider la situation épidémiologique du ceRs dans cette partie de l'île. De plus, pour pallier aux problèmes de mort cellulaire ou de changement d'état des bactéries en bactéries viables, mais non cultivables, il faudrait éviter de stocker longtemps les macérats de tige, mais les isoler rapidement à Madagascar et envoyer à La Réunion des suspensions bactériennes pour l'analyse phylogénétique et le géotypage des souches.

3. Activité antibactérienne entre les souches du ceRs

Une description fine de la situation épidémiologique est essentielle, mais insuffisante pour gérer une épidémie. En effet, la compréhension des facteurs et mécanismes la favorisant est également importante pour établir des stratégies de contrôle efficaces. Dans le cadre de cette thèse, nous avons testé l'activité antibactérienne entre les souches du ceRs et observé une cohérence ou non entre ces activités antibactériennes et la situation épidémiologique à Madagascar et dans les petites îles du SOOI.

L'activité inhibitrice des surnageants de 37 souches dites productrices ont été testées contre 99 souches indicatrices. Ces souches sont représentatives de la diversité phylogénétique du ceRs et elles proviennent de Madagascar, des petites îles du SOOI et du

reste du monde. La caractérisation partielle des surnageants a montré qu'ils sont thermolabiles et dégradés par la protéinase K sauf pour ceux de 10 souches appartenant aux sequevars I-15, I-18 (MCC2), IIA-36, IIA-7, et IV-11. Bien que la protéinase K soit une enzyme protéolytique relativement peu spécifique (une sérine protéase) [239,240], certaines bactériocines caractérisées dans la littérature en sont également insensibles [241,242]. Les surnageants n'ont pas d'activités inhibitrices sur des souches autres que le ceRs. La dilution en série des surnageants a montré une atténuation de l'intensité du halo d'inhibition voir sa disparition lorsque le filtrat est très dilué pour le cas de 29 surnageants. Combinés ensemble, ces résultats suggèrent une inhibition bactérienne due aux bactériocines pour le cas de 29 souches indicatrices.

Les résultats ont montré une structuration des activités d'inhibition selon le sequevar des souches. Deux résultats marquants sont à retenir de ces tests. Le premier est le large spectre d'activités des bactériocines produites par les I-18 et I-31 qui sont notamment actives contre des lignées génétiques présentes à Madagascar et dans les petites îles du SOOI (I-15, I-18, I-31, I-33, I-46, IIB-1, III-19, IV-11) [12,203,223]. De plus, les I-18 malgaches produisent des bactériocines très actives contre les I-31. Ce qui pourrait apporter un élément de réponse concernant le fait qu'ils ne se soient pas disséminés à Madagascar. Pourtant les questions suivantes persistent: pourquoi est-ce que le sequevar I-18 est fortement prévalent à Madagascar et non dans les petites îles du SOOI? Pourquoi est-ce que seulement deux souches du sequevar I-31 y ont été détectées malgré sa large dissémination dans les îles voisines (SOOI) et au-delà en Afrique?

Dans la littérature, quelques études ont mis en évidence une plus forte compétitivité des souches produisant des bactériocines en conditions de co-inoculation [77,139,144,145]. Nous avons alors travaillé sur la mise au point d'un protocole de compétition *in vitro* entre les souches du ceRs. Ces travaux de mise au point devront être poursuivis afin de connaître si *in vitro* la lignée génétique malgache I-18 est effectivement plus compétitive que les lignées génétiques malgaches I-46, I-33, et I-31. De plus, afin de comprendre la non dissémination des I-18 dans les petites îles du SOOI, des tests de compétition *in vitro* entre la lignée génétique malgache I-18 et des souches représentatives des lignées génétiques identifiés dans les petites îles du SOOI devront également être menés. Les résultats permettront de connaître laquelle des lignées génétiques est la plus compétitive et ils apporteront un élément de réponse dans la compréhension de la situation

épidémiologique contrastée entre Madagascar et les petites îles du SOOI. Par la suite, des tests de compétition *in planta*, plus proche des conditions réelles, devront également être effectués pour conforter ou non les résultats observés *in vitro*.

Le deuxième résultat marquant de ces tests de l'activité bactériocine des souches du ceRs est l'identification des RUN36 (Martinique) et RUN3812 (Réunion) d'une même lignée génétique (sequevar IIA-36) qui produisent des bactériocines très actives contre un large spectre de souches du ceRs soulignant ainsi leur intérêt pour une utilisation en biocontrôle du flétrissement bactérien.

4. Potentiel des bactériocines dans le biocontrôle du ceRs

Les substances antimicrobiennes telles que les bactériocines ont été proposées comme étant la prochaine génération d'agents de lutte contre les maladies bactériennes, car elles sont capables de cibler spécifiquement l'agent pathogène concerné avec un impact minimal sur la communauté microbienne et l'environnement [243]. En effet, face aux dangers que représentent les pesticides et l'apparition des résistances aux antibiotiques, de plus en plus d'utilisateurs et de consommateurs recherchent des produits "sûrs" pour la santé humaine et l'environnement. Étant donné que les bactériocines sont naturellement produites par les bactéries dans l'environnement, il est considéré qu'elles ont une toxicité limitée envers les humains, les animaux ou les espèces bactériennes environnementales bénignes et certaines bactériocines sont déjà considérées comme généralement sûres pour une utilisation dans la conservation des aliments [244]. Quelques études illustrent bien l'utilisation de la bactériocine dans la lutte contre les agents phytopathogènes. Le succès de l'agrocine 84 comme stratégie de contrôle de la maladie de la galle du collet a entraîné le développement d'une souche non phytopathogène *Rhizobium rhizogenes* (anciennement appelé *Agrobacterium radiobacter*) productrice d'agrocine 84 qui a été commercialisée avec succès sous le nom de Nogall par Bio-care Technology [172] (<https://bio-caretechnology.com/nogall/>). La commercialisation sous le nom "Biokeeper" de souches avirulentes de *Pectobacterium* productrice de tailocines afin de lutter contre la pourriture molle de la pomme de terre au Japon [245]. Des expérimentations sous serre ont démontré l'efficacité du traitement par pulvérisation de tailocines produites par *Pseudomonas fluorescens* SF4c dans la lutte contre la galle bactérienne de la tomate [246]. Récemment, Rooney *et al.*, [176] ont mis en

évidence qu'une bactériocine, la putidacine L1 (PL1), peut être exprimée *in planta* sous une forme active et à des niveaux élevés dans *Arabidopsis* et *Nicotiana benthamiana* pour fournir une résistance efficace contre divers pathovars de *Pseudomonas*.

Ces exemples renforcent l'intérêt d'explorer le potentiel des bactériocines dans la lutte contre les souches du ceRs. Depuis le début de l'année 2021, des travaux de caractérisation génétique et biochimique de(s) bactériocine(s) produites par la souche réunionnaise RUN3812 (IIA-36) sont menés au sein de notre équipe. L'objectif principal étant l'identification de(s) bactériocines à l'origine de l'activité d'inhibition exceptionnelle de cette souche. De plus, un test préliminaire a été mené pour vérifier l'efficacité d'une solution de bactériocine produite par RUN3812 contre la souche réunionnaise RUN3012 (sequevar I-31, fortement prévalent dans les petites îles du SOOI) [247]. Pour cela, nous avons trempé pendant 15 mn dans la solution de bactériocine les racines des plants de tomate avant leur repotage dans de la terre infectée par la souche RUN3012 à une concentration finale de 10^5 CFU/ml/g de terre [247]. Immédiatement après, les plantes étaient encore arrosées par la solution de bactériocine. Au bout de 12 jours, le taux de survie des plantes traitées aux bactériocines (66,67%) est nettement supérieur à celles traitées à l'eau (témoin négatif) (16,67%) [247]. Ce résultat encourageant devra être poursuivi en réalisant des tests de traitement prophylactique (avant le repotage comme indiqué ci-dessus) et curatif par arrosage des plantes avec de la bactériocine quelques jours après le repotage dans de la terre infectée.

Résumé

Le flétrissement bactérien, causé par les bactéries du complexe d'espèces *Ralstonia solanacearum* (ceRs), est considéré comme l'une des maladies les plus dévastatrices sur les cultures de solanacées à Madagascar et dans les petites îles du Sud-Ouest de l'Océan Indien (SOOI) (Réunion, Maurice, Rodrigues, Seychelles, Comores, Mayotte). Afin de lutter contre cette maladie et pour gérer les épidémies rencontrées dans la zone, une mise en lumière et la compréhension de la situation épidémiologique du ceRs à Madagascar et dans les petites îles du Sud-ouest de l'océan indien (SOOI) s'avèrent nécessaires. Notre première démarche a alors été de mettre au point un schéma de génotypage spécifique aux souches du phylotype I. Ensuite, nous avons analysé la diversité et la structure génétique du ceRs malgache. La situation épidémiologique observée à Madagascar est contrastée par rapport à celle observée dans les petites îles du SOOI où l'on rencontre une forte prévalence du phylotype I sequevar 31. En effet, seules deux souches du sequevar I-31 y ont été identifiées tandis que le phylotype I sequevar 18 représente la grande majorité des échantillons malgaches typés. Ce résultat surprenant nous a conduit à explorer en particulier le rôle de molécules antimicrobiennes, les bactériocines, dans la prévalence des deux lignées génétiques à Madagascar, dans les petites îles du SOOI et en général à étudier la capacité de production de bactériocines des souches du ceRs. Nos résultats ont montré que les sequevars I-18 et I-31 produisent des bactériocines actives contre un large spectre de souches rencontrées à Madagascar et dans les petites îles du SOOI. De plus, les souches I-18 malgaches produisent des bactériocines très actives contre les souches du sequevar I-31. Néanmoins, les résultats obtenus ne permettent pas d'expliquer entièrement les situations épidémiologiques observées dans la région et d'autres pistes devront encore être explorées. Un autre résultat majeur de cette thèse a été l'identification de deux souches appartenant à la lignée génétique IIA-36 qui produisent des bactériocines fortement actives contre la très grande majorité de souches testées du ceRs. Vu l'importance économique des pertes causées par le flétrissement bactérien qui est considéré comme étant responsable de la deuxième phytobactériose la plus nuisible au monde et le manque de moyen de lutte efficace contre cette bactérie, une stratégie de biocontrôle via les bactériocines va maintenant être explorée.

Mots clés: Epidémiologie moléculaire - Complexe d'espèces *Ralstonia solanacearum* - Flétrissement bactérien - MLVA - Bactériocines

Abstract

Bacterial wilt (BW), caused by the bacteria *Ralstonia solanacearum* species complex (RSSC), is considered as one of the most devastating diseases on solanaceous crops in Madagascar and the small islands of the South-West Indian Ocean (SOOI) (Réunion, Mauritius, Rodrigues, Seychelles, Comoros, and Mayotte). To manage BW, it is essential to analyze the epidemiological pattern of RSSC in this region. Firstly, we developed a genotyping scheme specific to phylotype I strains. Then, we analyzed the diversity and genetic structure of the Malagasy strains. The epidemiological pattern of RSSC in Madagascar is contrasted compared to that observed in the small islands of the SWIO (high prevalence of phylotype I sequevar 31). Indeed, only two strains of the sequevar I-31 have been identified while phylotype I sequevar 18 represented the vast majority of the Malagasy isolates. This surprising result prompted us to explore the influence of antimicrobial molecules, the bacteriocins, in the prevalence of the genetic lineages I-18 in Madagascar and I-31 in the small islands of the SWIO. We also studied the bacteriocin activity of various RSSC strains. Sequevars I-18 and I-31 produce bacteriocins active against a broad spectrum of strains encountered in Madagascar and the small islands of the SWIO. In addition, I-18 Malagasy strains produce bacteriocins highly active against sequevar I-31. Nevertheless, the results obtained do not fully explain the epidemiology of RSSC in the region and other mechanisms and factors should be explored. Another major result of this PhD thesis was the identification of two strains belonging to IIA-36 that produce bacteriocins highly active against the vast majority of RSSC strains tested. Clearly, in view of the need to control this devastating plant pathogen which is considered as the second most damaging phyto-bacteria in the world and the lack of effective control methods, a biocontrol strategy based on bacteriocins will now be explored.

Keywords: molecular epidemiology – *Ralstonia solanacearum* species complex – Bacterial wilt – MLVA - Bacteriocins

Bibliographie

1. Stukenbrock EH, McDonald BA. The Origins of Plant Pathogens in Agro-Ecosystems. *Annu Rev Phytopathol.* 2008;46: 75–100. doi:10.1146/annurev.phyto.010708.154114
2. Balter M. Seeking agriculture's ancient roots. *Science.* 2007;316: 1830–1835.
3. Diamond J. *Guns, germs, and steel: the fates of human societies.* New York: Norton. 1997.
4. Foëx E. Histoire de l'introduction de quelques nouvelles Maladies des plantes. *Rev Bot Appliquée Agric Colon.* 1924;4: 561–568. doi:10.3406/jatba.1924.4214
5. Diller JD. Chestnut Blight. USDA Forest Service. Washington, DC; 1965.
6. Saponari M, Giampetruzzi A, Loconsole G, Boscia D, Saldarelli P. *Xylella fastidiosa* in Olive in Apulia: Where We Stand. *Phytopathology®.* 2019;109: 175–186. doi:10.1094/PHYTO-08-18-0319-FI
7. Keim PS. Genotyping (Subtyping) Bacteria. *Brenner's Encyclopedia of Genetics.* Elsevier; 2013. pp. 321–323. doi:10.1016/B978-0-12-374984-0.00646-X
8. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, et al. Top 10 plant pathogenic bacteria in molecular plant pathology: Top 10 plant pathogenic bacteria. *Mol Plant Pathol.* 2012;13: 614–629. doi:10.1111/j.1364-3703.2012.00804.x
9. Denny T. Plant pathogenic *Ralstonia* species. In: Gnanamanickam SS, editor. *Plant-Associated Bacteria.* Dordrecht: Springer Netherlands; 2006. pp. 573–644. doi:10.1007/978-1-4020-4538-7_16
10. Hayward AC. Biology and Epidemiology of Bacterial Wilt Caused by *Pseudomonas solanacearum*. *Annu Rev Phytopathol.* 1991;29: 65–87.
11. Ravelomanantsoa S. Biologie des populations du complexe d'espèces *Ralstonia solanacearum* appliquée à l'épidémiologie du flétrissement bactérien de la pomme de terre à Madagascar. Université de La Réunion; 2016.
12. Ravelomanantsoa S, Vernière C, Rieux A, Costet L, Chiroleu F, Arribat S, et al. Molecular epidemiology of bacterial wilt in the Madagascar highlands caused by Andean (Phylotype IIB-1) and African (Phylotype III) brown rot strains of the *Ralstonia solanacearum* species complex. *Front Plant Sci.* 2018;8. doi:10.3389/fpls.2017.02258
13. Rapilly F. L'épidémiologie en pathologie végétale. *Mycoses aériennes.* Quae. 1991.
14. Zadoks JC, Schein RD. Epidemiology and plant-disease management, the known and the needed. *Comparative epidemiology A tool for better disease management.* Wageningen, Neetherlands: Centre for Agricultural Publishing and Documentation Wageningen; 1980. pp. 1–17.
15. Picard B. Épidémiologie moléculaire des grandes endémies. 1999; 5.
16. Roughgarden J. *Theory of population genetics and evolutionary ecology: an introduction.* New York: Macmillan Publishing Company; 1979.
17. McDonald BA, Linde C. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu Rev Phytopathol.* 2002;40: 349–379. doi:10.1146/annurev.phyto.40.120501.101443
18. Schork NJ, Fallin D, Lanchbury JS. Single nucleotide polymorphisms and the future of genetic epidemiology: SNPs and genetic epidemiology. *Clin Genet.* 2000;58: 250–264. doi:10.1034/j.1399-0004.2000.580402.x
19. Brynildsrud O, Snipen L-G, Bohlin J. CNOGpro: detection and quantification of CNVs

- in prokaryotic whole-genome sequencing data. *Bioinformatics*. 2015;31: 1708–1715. doi:10.1093/bioinformatics/btv070
20. Bi X, Liu LF. A Replicational Model for DNA Recombination between Direct Repeats. *J Mol Biol*. 1996;256: 849–858. doi:10.1006/jmbi.1996.0131
 21. Richard G-F, Pâques F. Mini- and microsatellite expansions: the recombination connection. *EMBO Rep*. 2000;1: 122–126. doi:10.1093/embo-reports/kvd031
 22. Tidjani A-R. Evolution génomique au sein d'une population naturelle de *Streptomyces*. Université de Lorraine; 2019.
 23. Furuya EY, Lowy FD. Antimicrobial-resistant bacteria in the community setting. *Nat Rev Microbiol*. 2006;4: 36–45. doi:10.1038/nrmicro1325
 24. Duminil J. Etudes comparatives de la structure génétique des Plantes. Université Henri Poincaré - Nancy 1; 2006.
 25. Wright S. The interpretation of population structure by *f*-statistics with special regard to systems of mating. *Evolution*. 1965;19: 395–420. doi:10.1111/j.1558-5646.1965.tb01731.x
 26. Kimura M, Crow JF. The number of alleles that can be maintained in a finite population. *Genetics*. 1964;49: 725–738.
 27. Balloux F, Lugon-Moulin N. The estimation of population differentiation with microsatellite markers. *Mol Ecol*. 2002;11: 155–165.
 28. Slatkin M. A Measure of Population Subdivision Based on Microsatellite Allele Frequencies. *Genetics*. 1995;139: 457–462.
 29. Kimura M, Ohta T. Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proc Natl Acad Sci*. 1978;75: 2868–2872. doi:10.1073/pnas.75.6.2868
 30. Nei M. *F*-statistics and analysis of gene diversity in subdivided populations. *Ann Hum Genet*. 1977;41: 225–233. doi:10.1111/j.1469-1809.1977.tb01918.x
 31. Roberts D, Ayre D. Panmictic population structure in the migratory marine sparid *Acanthopagrus australis* despite its close association with estuaries. *Mar Ecol Prog Ser*. 2010;412: 223–230. doi:10.3354/meps08676
 32. D'Alelio D, Salmaso N, Gandolfi A. Frequent recombination shapes the epidemic population structure of *Planktothrix* (Cyanoprokaryota) in Italian subalpine lakes. Valentin K, editor. *J Phycol*. 2013;49: 1107–1117. doi:10.1111/jpy.12116
 33. Li W, Raoult D, Fournier P-E. Bacterial strain typing in the genomic era. *FEMS Microbiol Rev*. 2009;33: 892–916.
 34. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol*. 1975;98: 503–517.
 35. Vergnaud G, Pourcel C. Multiple Locus VNTR (Variable Number of Tandem Repeat) Analysis. In: Stackebrandt E, editor. *Molecular Identification, Systematics, and Population Structure of Prokaryotes*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006. pp. 83–104. doi:10.1007/978-3-540-31292-5_4
 36. Vergnaud G, Pourcel C. Multiple Locus Variable Number of Tandem Repeats Analysis. In: Caugant DA, editor. *Molecular epidemiology of microorganisms*. Totowa, NJ: Humana Press; 2009. pp. 141–158. doi:10.1007/978-1-60327-999-4_12
 37. Pourcel C, Vergnaud G. Strain typing using Multiple "Variable Number of Tandem Repeat" Analysis and genetic element CRISPR. 2011.
 38. Van Belkum A, Scherer S, Van Leeuwen W, Willemse D, Van Alphen L, Verbrugh H. Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. *Infect Immun*. 1997;65: 5017–5027.
 39. Van Belkum A. Tracing isolates of bacterial species by Multilocus Variable number of

- tandem repeat Analysis (MLVA). FEMS Immunol Med Microbiol. 2007;49: 22–27. doi:10.1111/j.1574-695X.2006.00173.x
40. Coletta-Filho HD, Takita MA, de Souza AA, Aguilar-Vildoso CI, Machado MA. Differentiation of strains of *Xylella fastidiosa* by a variable number of tandem repeat analysis. Appl Environ Microbiol. 2001;67: 4091–4095.
 41. Zaluga J, Stragier P, Van Vaerenbergh J, Maes M, De Vos P. Multilocus Variable Number Tandem Repeats Analysis (MLVA) distinguishes a clonal complex of *Clavibacter michiganensis subsp. michiganensis* strains isolated from recent outbreaks of bacterial wilt and canker in Belgium. BMC Microbiol. 2013;13: 126. doi:10.1186/1471-2180-13-126
 42. Gironde S, Manceau C. Housekeeping gene sequencing and Multilocus Variable Number Tandem Repeat Analysis to identify subpopulations within *Pseudomonas syringae pv. maculicola* and *Pseudomonas syringae pv. tomato* that correlate with host specificity. Appl Environ Microbiol. 2012;78: 3266–3279. doi:10.1128/AEM.06655-11
 43. Francisco CS, Ceresini PC, Almeida RPP, Coletta-Filho HD. Spatial genetic structure of coffee associated *Xylella fastidiosa* populations indicates that cross infection does not occur with sympatric citrus orchards. Phytopathology. 2017;107: 395–402. doi:10.1094/PHYTO-08-16-0300-R
 44. Bui Thi Ngoc L, Verniere C, Jarne P, Brisse S, Guerin F, Boutry S, et al. From local surveys to global surveillance: three high-throughput genotyping methods for epidemiological monitoring of *Xanthomonas citri pv. citri* pathotypes. Appl Environ Microbiol. 2009;75: 1173–1184. doi:10.1128/AEM.02245-08
 45. Pruvost O, Magne M, Boyer K, Leduc A, Tourterel C, Drevet C, et al. A MLVA genotyping scheme for global surveillance of the Citrus Pathogen *Xanthomonas citri pv. citri* suggests a worldwide geographical expansion of a single genetic lineage. Vinatzer BA, editor. PLoS ONE. 2014;9: e98129. doi:10.1371/journal.pone.0098129
 46. Poulin L, Grygiel P, Magne M, Gagnevin L, Rodriguez-R LM, Forero Serna N, et al. New Multilocus Variable Number Tandem Repeat Analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by *Xanthomonas oryzae*. Goodrich-Blair H, editor. Appl Environ Microbiol. 2015;81: 688–698. doi:10.1128/AEM.02768-14
 47. López-Soriano P, Boyer K, Cesbron S, Morente MC, Peñalver J, Palacio-Bielsa A, et al. Multilocus Variable Number of Tandem Repeat Analysis reveals multiple introductions in Spain of *Xanthomonas arboricola pv. pruni*, the causal agent of bacterial spot disease of stone fruits and almond. Vinatzer BA, editor. PLoS ONE. 2016;11: e0163729. doi:10.1371/journal.pone.0163729
 48. Rache L, Blondin L, Flores C, Trujillo C, Szurek B, Restrepo S, et al. An Optimized Microsatellite Scheme for Assessing Populations of *Xanthomonas phaseoli pv. manihotis*. Phytopathology. 2019;109: 859–869. doi:10.1094/PHYTO-06-18-0210-R
 49. Nakato GV, Fuentes Rojas JL, Verniere C, Blondin L, Coutinho T, Mahuku G, et al. A new Multi Locus Variable Number of Tandem Repeat Analysis scheme for epidemiological surveillance of *Xanthomonas vasicola pv. musacearum*, the plant pathogen causing bacterial wilt on banana and enset. Chiang T-Y, editor. PLoS ONE. 2019;14: e0215090. doi:10.1371/journal.pone.0215090
 50. Ferreira MASV, Bonneau S, Briand M, Cesbron S, Portier P, Darrasse A, et al. *Xanthomonas citri pv. viticola* affecting grapevine in Brazil: emergence of a successful monomorphic pathogen. Front Plant Sci. 2019;10: 489. doi:10.3389/fpls.2019.00489

51. Ravelomanantsoa S, Robène I, Chiroleu F, Guérin F, Poussier S, Pruvost O, et al. A novel multilocus variable number tandem repeat analysis typing scheme for African phylotype III strains of the *Ralstonia solanacearum* species complex. PeerJ. 2016;4: e1949. doi:10.7717/peerj.1949
52. N'Guessan CA, Brisse S, Le Roux-Nio A-C, Poussier S, Koné D, Wicker E. Development of variable number of tandem repeats typing schemes for *Ralstonia solanacearum*, the agent of bacterial wilt, banana Moko disease and potato brown rot. J Microbiol Methods. 2013;92: 366–374. doi:10.1016/j.mimet.2013.01.012
53. Parkinson N, Bryant R, Bew J, Conyers C, Stones R, Alcock M, et al. Application of Variable-Number Tandem-Repeat Typing to discriminate *Ralstonia solanacearum* strains associated with English watercourses and disease outbreaks. Appl Environ Microbiol. 2013;79: 6016–6022. doi:10.1128/AEM.01219-13
54. Guinard J, Latreille A, Guérin F, Poussier S, Wicker E. New Multilocus Variable Number Tandem Repeat Analysis (MLVA) scheme for fine scale monitoring and microevolution related study of *Ralstonia pseudosolanacearum* phylotype I populations. Drake HL, editor. Appl Environ Microbiol. 2017;83. doi:10.1128/AEM.03095-16
55. Rasschaert G, Koen De Reu, Heyndrickx M. un outil épidémiologique de valeur ? Labinfo. 2013;9: 4.
56. Vergnaud G, Denoeud F. Minisatellites: mutability and genome architecture. Genome Res. 2000;10: 899–907.
57. Levinson G, Gutman GA. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol. 1987 [cited 10 May 2019]. doi:10.1093/oxfordjournals.molbev.a040442
58. Leduc A. Apport de l'épidémiologie moléculaire et des approches inférentielles dans l'analyse de l'émergence et des routes d'invasion de *Xanthomonas citri* pv. *citri* en Afrique, bactérie responsable du Chancre asiatique des agrumes. 2015; 332.
59. Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB. Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci. 1994;91: 3166–3170. doi:10.1073/pnas.91.8.3166
60. Jarne P, Lagoda PJ. Microsatellites, from molecules to populations and back. Trends Ecol Evol. 1996;11: 424–429.
61. Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol Evol. 2004;19: 535–544. doi:10.1016/j.tree.2004.07.021
62. Morse SS. Factors in the Emergence of Infectious Diseases. Emerg Infect Dis. 1995;1: 9.
63. Kuyek D. Blast, Biotech and Big Business: Implications of Corporate Strategies on Rice Research in Asia. Genetic Resources Action International; 2000. Available: http://www.grain.org/briefings_files/blast.pdf
64. Fletcher JT, Butler D. Strain changes in populations of tobacco mosaic virus from tomato crops. Ann Appl Biol. 1975;81: 409–412.
65. Irey M, Gottwald TR, Graham JH, Riley TD, Carlton G. Post-hurricane analysis of citrus canker spread and progress towards the development of a predictive model to estimate disease spread due to catastrophic weather events. Plant Health Prog. 2006;7: 16.
66. Peterson JW. Bacterial pathogenesis. Med Microbiol 4th Ed. 1996.
67. Salmond GP. Secretion of extracellular virulence factors by plant pathogenic bacteria. Annu Rev Phytopathol. 1994;32: 181–200.

68. Morris CE, Bardin M, Kinkel LL, Moury B, Nicot PC, Sands DC. Expanding the Paradigms of Plant Pathogen Life History and Evolution of Parasitic Fitness beyond Agricultural Boundaries. Rall GF, editor. PLoS Pathog. 2009;5: e1000693. doi:10.1371/journal.ppat.1000693
69. Escalon. Evolution et spécialisation du pouvoir pathogène de *Xanthomonas citri* pv. *citri*: rôle des effecteurs de type 3. Université de La Réunion; 2013.
70. Wicker E, Grassart L, Coranson-Beaudu R, Mian D, Prior P. Epidemiological evidence for the emergence of a new pathogenic variant of *Ralstonia solanacearum* in Martinique (French West Indies). Plant Pathol. 2009;58: 853–861. doi:10.1111/j.1365-3059.2009.02098.x
71. Hibbing ME, Fuqua C, Parsek MR, Peterson SB. Bacterial competition: surviving and thriving in the microbial jungle. Nat Rev Microbiol. 2010;8: 15–25. doi:10.1038/nrmicro2259
72. Granato ET, Meiller-Legrand TA, Foster KR. The Evolution and Ecology of Bacterial Warfare. Curr Biol. 2019;29: R521–R537. doi:10.1016/j.cub.2019.04.024
73. Russell AB, Peterson SB, Mougous JD. Type VI secretion system effectors: poisons with a purpose. Nat Rev Microbiol. 2014;12: 137–148. doi:10.1038/nrmicro3185
74. Riley MA, Chavan MA, editors. Bacteriocins: ecology and evolution. Berlin ; New York: Springer; 2007.
75. Drider D, Rebuffat S, editors. Prokaryotic antimicrobial peptides: from genes to applications. New York: Springer; 2011.
76. Goel AK, Sindhu SS, Dadarwal KR. Bacteriocin-producing native rhizobia of green gram (*Vigna radiata*) having competitive advantage in nodule occupancy. Microbiol Res. 1999;154: 43–48. doi:10.1016/S0944-5013(99)80033-0
77. Tudor-Nelson SM, Minsavage GV, Stall RE, Jones JB. Bacteriocin-Like Substances from Tomato Race 3 Strains of *Xanthomonas campestris* pv. *vesicatoria*. Phytopathology®. 2003;93: 1415–1421. doi:10.1094/PHYTO.2003.93.11.1415
78. Hert AP, Roberts PD, Momol MT, Minsavage GV, Tudor-Nelson SM, Jones JB. Relative Importance of Bacteriocin-Like Genes in Antagonism of *Xanthomonas perforans* Tomato Race 3 to *Xanthomonas euvesicatoria* Tomato Race 1 Strains. Appl Environ Microbiol. 2005;71: 3581–3588. doi:10.1128/AEM.71.7.3581-3588.2005
79. Zacharof MP, Lovitt RW. Bacteriocins Produced by Lactic Acid Bacteria a Review Article. APCBEE Procedia. 2012;2: 50–56. doi:10.1016/j.apcbee.2012.06.010
80. De Vuyst L, Vandamme EJ. Influence of the carbon source on nisin production in *Lactococcus lactis* subsp. *lactis* batch fermentations. Microbiology. 1992;138: 571–578.
81. Parente E, Ricciardi A. Influence of pH on the production of enterocin 1146 during batch fermentation. Lett Appl Microbiol. 1994;19: 12–15.
82. Mortvedt-Abildgaa CI, Nissen-Meyer J, Jelle B, Grenov B, Skaugen M, Nes IF. Production and pH-dependent bactericidal activity of lactocin S, a lantibiotic from *Lactobacillus sake* L45. Appl Environ Microbiol. 1995;61: 175–179.
83. Kozak W, Bardowski J, Dobrzański WT. Lactostrepcins-acid bacteriocins produced by lactic streptococci. J Dairy Res. 1978;45: 247–257.
84. Biswas SR, Ray P, Johnson MC, Ray B. Influence of growth conditions on the production of a bacteriocin, pediocin ACh, by *Pediococcus acidilactici* H. Appl Environ Microbiol. 1991;57: 1265–1267.
85. Jiménez-Díaz R, Rios-Sanchez RM, Desmazeaud M, Ruiz-Barba JL, Piard J-C. Plantaricins S and T, two new bacteriocins produced by *Lactobacillus plantarum* LPCO10 isolated from a green olive fermentation. Appl Environ Microbiol. 1993;59:

- 1416–1424.
86. West SA, Diggle SP, Buckling A, Gardner A, Griffin AS. The Social Lives of Microbes. *Annu Rev Ecol Syst.* 2007;38: 53–77. doi:10.1146/annurev.ecolsys.38.091206.095740
 87. Cascales E, Buchanan SK, Duché D, Kleantous C, Llobès R, Postle K, et al. Colicin Biology. *Microbiol Mol Biol Rev.* 2007;71: 158–229. doi:10.1128/MMBR.00036-06
 88. McAuliffe O, O’Keeffe T, Hill C, Ross RP. Regulation of immunity to the two-component lantibiotic, lactacin 3147, by the transcriptional repressor LtnR. *Mol Microbiol.* 2001;39: 982–993. doi:10.1046/j.1365-2958.2001.02290.x
 89. Franz CMAP, Worobo RW, Quadri LEN, Schillinger U, Holzappel WH, Vederas JC, et al. Atypical Genetic Locus Associated with Constitutive Production of Enterocin B by *Enterococcus faecium* BFE 900. *Appl Environ Microbiol.* 1999;65: 2170.
 90. Kleerebezem M. Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis. *M Altstein.* 2004;25: 1405–1414. doi:10.1016/j.peptides.2003.10.021
 91. Eijsink VGH, Axelsson L, Diep DB, Håvarstein LS, Holo H, Nes IF. Production of class II bacteriocins by lactic acid bacteria; an example of biological warfare and communication. *Antonie Van Leeuwenhoek.* 2002;81: 639–654. doi:10.1023/A:1020582211262
 92. Risøen PA, Brurberg MB, Eijsink VGH, Nes IF. Functional analysis of promoters involved in quorum sensing-based regulation of bacteriocin production in *Lactobacillus*. *Mol Microbiol.* 2000;37: 619–628. doi:10.1046/j.1365-2958.2000.02029.x
 93. Blanchard AE, Liao C, Lu T. An Ecological Understanding of Quorum Sensing-Controlled Bacteriocin Synthesis. *Cell Mol Bioeng.* 2016;9: 443–454. doi:10.1007/s12195-016-0447-6
 94. Cornforth DM, Foster KR. Competition sensing: the social side of bacterial stress responses. *Nat Rev Microbiol.* 2013;11: 285–293. doi:10.1038/nrmicro2977
 95. Cha C, Gao P, Chen Y-C, Shaw PD, Farrand SK. Production of Acyl-Homoserine Lactone Quorum-Sensing Signals by Gram-Negative Plant-Associated Bacteria. *Mol Plant Microbe Interact.* 1998;11: 1119–1129. doi:10.1094/MPMI.1998.11.11.1119
 96. Arwiyanto T, GOTO M, TAKIKAWA Y. Characterization of Bacteriocins Produced by *Pseudomonas solanacearum*. *Jpn J Phytopathol.* 1993;59: 114–122. doi:10.3186/jjphytopath.59.114
 97. Huerta AI, Milling A, Allen C. Tropical strains of *Ralstonia solanacearum* outcompete race 3 biovar 2 strains at lowland tropical temperatures. Drake HL, editor. *Appl Environ Microbiol.* 2015;81: 3542–3551. doi:10.1128/AEM.04123-14
 98. Yang R, Ray B. Factors influencing production of bacteriocins by lactic acid bacteria. *Food Microbiol.* 1994;11: 281–291.
 99. Riley MA, Wertz JE. Bacteriocins: Evolution, Ecology, and Application. *Annu Rev Microbiol.* 2002;56: 117–137. doi:10.1146/annurev.micro.56.012302.161024
 100. James R, Lazdunski C, Pattus F. Bacteriocins, Microcins and Lantibiotics. Springer-Verlag. 1991.
 101. Torreblanca M, Meseguer I, Ventosa A. Production of halocin is a practically universal feature of archaeal halophilic rods. *Lett Appl Microbiol.* 1994;19: 201–205. doi:10.1111/j.1472-765X.1994.tb00943.x
 102. Riley MA, Gordon DM. The ecological role of bacteriocins in bacterial competition. *Trends Microbiol.* 1999;7: 129–133. doi:10.1016/S0966-842X(99)01459-6
 103. Klaenhammer TR. Bacteriocins of lactic acid bacteria. *Biochimie.* 1988;70: 337–

349.

104. Klaenhammer TR. Genetics of bacteriocins produced by lactic acid bacteria. *FEMS Microbiol Rev.* 1993;12: 39–86. doi:10.1111/j.1574-6976.1993.tb00012.x
105. Kumariya R, Garsa AK, Rajput YS, Sood SK, Akhtar N, Patel S. Bacteriocins: Classification, synthesis, mechanism of action and resistance development in food spoilage causing bacteria. *Microb Pathog.* 2019;128: 171–177. doi:10.1016/j.micpath.2019.01.002
106. Szabo EA, Cahill ME. Nisin and ALTA 2341 inhibit the growth of *Listeria monocytogenes* on smoked salmon packaged under vacuum or 100% CO₂. *Lett Appl Microbiol.* 1999;28: 373–377. doi:10.1046/j.1365-2672.1999.00547.x
107. Mazzotta AS, Crandall AD, Montville TJ. Nisin resistance in *Clostridium botulinum* spores and vegetative cells. *Appl Environ Microbiol.* 1997;63: 6.
108. Le Lay C, Dridi L, Bergeron MG, Ouellette M, Fliss I. Nisin is an effective inhibitor of *Clostridium difficile* vegetative cells and spore germination. *J Med Microbiol.* 2016;65: 169–175. doi:10.1099/jmm.0.000202
109. Harris LJ, Fleming HP, Klaenhammer TR. Sensitivity and Resistance of *Listeria monocytogenes* ATCC 19115, Scott A, and UAL500 to Nisin. *J Food Prot.* 1991;54: 836–840. doi:10.4315/0362-028X-54.11.836
110. Férir G, Petrova MI, Andrei G, Huskens D, Hoorelbeke B, Snoeck R, et al. The lantibiotic peptide labyrinthopeptin A1 demonstrates broad anti-HIV and anti-HSV activity with potential for microbicidal applications. *PLOS ONE.* 2013;8: e64010. doi:10.1371/journal.pone.0064010
111. Mathur H, Fallico V, O'Connor PM, Rea MC, Cotter PD, Hill C, et al. Insights into the Mode of Action of the Sactibiotic Thuricin CD. *Front Microbiol.* 2017;8: 696. doi:10.3389/fmicb.2017.00696
112. Fimland G, Jack R, Jung G, Nes IF, Nissen-Meyer J. The bactericidal activity of pediocin PA-1 is specifically inhibited by a 15-mer fragment that spans the bacteriocin from the center toward the C terminus. *Appl Environ Microbiol.* 1998;64: 5057.
113. Kumariya R, Sood SK, Rajput YS, Garsa AK. Gradual pediocin PA-1 resistance in *Enterococcus faecalis* confers cross-protection to diverse pore-forming cationic antimicrobial peptides displaying changes in cell wall and mannose PTS expression. *Ann Microbiol.* 2015; 721–732.
114. Trinetta V, Morleo A, Sessa F, Iametti S, Bonomi F, Ferranti P. Purified sakacin A shows a dual mechanism of action against *Listeria* spp: proton motive force dissipation and cell wall breakdown. *FEMS Microbiol Lett.* 2012;334: 143–149. doi:10.1111/j.1574-6968.2012.02630.x
115. Chen H, Tian F, Li S, Xie Y, Zhang H, Chen W. Cloning and heterologous expression of a bacteriocin sakacin P from *Lactobacillus sakei* in *Escherichia coli*. *Appl Microbiol Biotechnol.* 2012;94: 1061–1068. doi:10.1007/s00253-012-3872-z
116. Balay DR, Dangeti RV, Kaur K, McMullen LM. Purification of leucocin A for use on wieners to inhibit *Listeria monocytogenes* in the presence of spoilage organisms. *Int J Food Microbiol.* 2017;255: 25–31. doi:10.1016/j.ijfoodmicro.2017.05.016
117. Oppegard C, Rogne P, Kristiansen PE, Nissen-Meyer J. Structure analysis of the two-peptide bacteriocin lactococcin G by introducing D-amino acid residues. *Microbiology.* 2010;156: 1883–1889. doi:10.1099/mic.0.038430-0
118. Ekblad B, Kyriakou PK, Oppegård C, Nissen-Meyer J, Kaznessis YN, Kristiansen PE. Structure–function analysis of the Two-Peptide bacteriocin plantaricin EF. *Biochemistry.* 2016;55: 5106–5116. doi:10.1021/acs.biochem.6b00588

119. Ekblad B, Nissen-Meyer J, Kristensen T. Whole-genome sequencing of mutants with increased resistance against the two-peptide bacteriocin plantaricin JK reveals a putative receptor and potential docking site. *PLOS ONE*. 2017;12: e0185279. doi:10.1371/journal.pone.0185279
120. Pandey N, Malik RK, Kaushik JK, Singroha G. Gassericin A: a circular bacteriocin produced by Lactic acid bacteria *Lactobacillus gasserii*. *World J Microbiol Biotechnol*. 2013;29: 1977–1987. doi:10.1007/s11274-013-1368-3
121. Grande Burgos M, Pulido R, del Carmen López Aguayo M, Gálvez A, Lucas R. The Cyclic Antibacterial Peptide Enterocin AS-48: Isolation, Mode of Action, and Possible Food Applications. *Int J Mol Sci*. 2014;15: 22706–22727. doi:10.3390/ijms151222706
122. Borrero J, Brede DA, Skaugen M, Diep DB, Herranz C, Nes IF, et al. Characterization of Garvicin ML, a novel circular bacteriocin produced by *Lactococcus garvieae* DCC43, isolated from Mallard Ducks (*Anas platyrhynchos*). *Appl Environ Microbiol*. 2011;77: 369. doi:10.1128/AEM.01173-10
123. O’Shea EF, O’Connor PM, O’Sullivan O, Cotter PD, Ross RP, Hill C. Bactofencin A, a new type of cationic bacteriocin with unusual immunity. Cossart PF, editor. *mBio*. 2013;4: e00498-13. doi:10.1128/mBio.00498-13
124. O’Connor PM, O’Shea EF, Cotter PD, Hill C, Ross RP. The potency of the broad spectrum bacteriocin, bactofencin A, against staphylococci is highly dependent on primary structure, N-terminal charge and disulphide formation. *Sci Rep*. 2018;8: 11833. doi:10.1038/s41598-018-30271-6
125. Uzelac G, Miljkovic M, Lozo J, Radulovic Z, Tosic N, Kojic M. Expression of bacteriocin LsbB is dependent on a transcription terminator. *Microbiol Res*. 2015;179: 45–53. doi:10.1016/j.micres.2015.06.011
126. Sun Z, Wang X, Zhang X, Wu H, Zou Y, Li P, et al. Class III bacteriocin Helveticin-M causes sublethal damage on target cells through impairment of cell wall and membrane. *J Ind Microbiol Biotechnol*. 2018;45: 213–227. doi:10.1007/s10295-018-2008-6
127. Thompson JK, Collins MA, Mercer WD. Characterization of a proteinaceous antimicrobial produced by *Lactobacillus helveticus* CNRZ450. *J Appl Bacteriol*. 1996;80: 338–348. doi:10.1111/j.1365-2672.1996.tb03229.x
128. Nilsen T, Nes IF, Holo H. Enterolysin A, a cell wall-degrading bacteriocin from *enterococcus faecalis* LMG 2333. *Appl Environ Microbiol*. 2003;69: 2975. doi:10.1128/AEM.69.5.2975-2984.2003
129. Chavan MA, Riley MA. Molecular Evolution of Bacteriocins in Gram-Negative Bacteria. In: Riley MA, Chavan MA, editors. *Bacteriocins*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. pp. 19–43. doi:10.1007/978-3-540-36604-1_3
130. Duquesne S, Destoumieux-Garzón D, Peduzzi J, Rebuffat S. Microcins, gene-encoded antibacterial peptides from enterobacteria. *Nat Prod Rep*. 2007;24: 708. doi:10.1039/b516237h
131. Baquero F, Moreno F. The microcins. *FEMS Microbiol Lett*. 1984;23: 117–124. doi:10.1111/j.1574-6968.1984.tb01046.x
132. Little JW, Mount DW. The SOS regulatory system of *Escherichia coli*. *Cell*. 1982;29: 11–22. doi:10.1016/0092-8674(82)90085-X
133. Nedialkova LP, Sidstedt M, Koeppl MB, Spriewald S, Ring D, Gerlach RG, et al. Temperate phages promote colicin-dependent fitness of *S almonella enterica* serovar Typhimurium: Temperate phages promote colicin release. *Environ Microbiol*. 2016;18: 1591–1603. doi:10.1111/1462-2920.13077

134. Campbell A. The future of bacteriophage biology. *Nat Rev Genet.* 2003;4: 471–477. doi:10.1038/nrg1089
135. Parret AHA, De Mot R. Bacteria killing their own kind: novel bacteriocins of *Pseudomonas* and other γ -proteobacteria. *Trends Microbiol.* 2002;10: 107–112. doi:10.1016/S0966-842X(02)02307-7
136. Scholl D. Phage Tail–Like Bacteriocins. *Annu Rev Virol.* 2017;4: 453–467. doi:10.1146/annurev-virology-101416-041632
137. Ghequire MGK, De Mot R. The Tailocin Tale: Peeling off Phage Tails. *Trends Microbiol.* 2015;23: 587–590. doi:10.1016/j.tim.2015.07.011
138. Bradley DE. Ultrastructure of Bacteriophages and Bacteriocins. *Bacteriol Rev.* 1967;31: 230–314.
139. Dorosky RJ, Pierson LS, Pierson EA. *Pseudomonas chlororaphis* Produces Multiple R-Tailocin Particles That Broaden the Killing Spectrum and Contribute to Persistence in Rhizosphere Communities. Vieille C, editor. *Appl Environ Microbiol.* 2018;84: e01230-18, /aem/84/18/e01230-18.atom. doi:10.1128/AEM.01230-18
140. Matsui H, Sano Y, Ishihara H, Shinomiya T. Regulation of pyocin genes in *Pseudomonas aeruginosa* by positive (prtN) and negative (prtR) regulatory genes. *J Bacteriol.* 1993;175: 1257–1263. doi:10.1128/JB.175.5.1257-1263.1993
141. Gebhart D, Williams SR, Bishop-Lilly KA, Govoni GR, Willner KM, Butani A, et al. Novel High-Molecular-Weight, R-Type Bacteriocins of *Clostridium difficile*. *J Bacteriol.* 2012;194: 6240–6247. doi:10.1128/JB.01272-12
142. Gebhart D, Lok S, Clare S, Tomas M, Stares M, Scholl D, et al. A Modified R-Type Bacteriocin Specifically Targeting *Clostridium difficile* Prevents Colonization of Mice without Affecting Gut Microbiota Diversity. Vidaver AK, editor. *mBio.* 2015;6: e02368-14. doi:10.1128/mBio.02368-14
143. Weiß AS, Götz A, Opitz M. Dynamics of ColicinE2 production and release determine the competitive success of a toxin-producing bacterial population. *Sci Rep.* 2020;10: 4052. doi:10.1038/s41598-020-61086-z
144. Smidt ML, Vidaver AK. Bacteriocin production by *Pseudomonas syringae* PsW-1 in plant tissue. *Can J Microbiol.* 1982;28: 600–604. doi:10.1139/m82-089
145. Kommineni S, Bretl DJ, Lam V, Chakraborty R, Hayward M, Simpson P, et al. Bacteriocin production augments niche competition by enterococci in the mammalian gastrointestinal tract. *Nature.* 2015;526: 719–722. doi:10.1038/nature15524
146. Neumann G, Schuster S. Continuous model for the rock–scissors–paper game between bacteriocin producing bacteria. *J Math Biol.* 2007;54: 815–846.
147. Gálvez A, Abriouel H, López RL, Omar NB. Bacteriocin-based strategies for food biopreservation. *Int J Food Microbiol.* 2007;120: 51–70. doi:10.1016/j.ijfoodmicro.2007.06.001
148. Ennahar S, Sashihara T, Sonomoto K, Ishizaki A. Class IIa bacteriocins: biosynthesis, structure and activity. *FEMS Microbiol Rev.* 2000;24: 85–106.
149. Paul Ross R, Morgan S, Hill C. Preservation and fermentation: past, present and future. *Int J Food Microbiol.* 2002;79: 3–16. doi:10.1016/S0168-1605(02)00174-5
150. Zhang J, Zhong J. The journey of nisin development in China, a natural-green food preservative. *Protein Cell.* 2015;6: 709–711. doi:10.1007/s13238-015-0214-9
151. Riley MA. MOLECULAR MECHANISMS OF BACTERIOCIN EVOLUTION. *Annu Rev Genet.* 1998;32: 255–278. doi:10.1146/annurev.genet.32.1.255
152. Oliveira DC, Tomasz A, de Lencastre H. Secrets of success of a human pathogen: molecular evolution of pandemic clones of meticillin-resistant *Staphylococcus*

- aureus. *Lancet Infect Dis.* 2002;2: 180–189. doi:10.1016/S1473-3099(02)00227-X
153. De Kwaadsteniet M, Doeschate KT, Dicks LMT. Nisin F in the treatment of respiratory tract infections caused by *Staphylococcus aureus*. *Lett Appl Microbiol.* 2009;48: 65–70. doi:10.1111/j.1472-765X.2008.02488.x
 154. Aslam S, Hamill RJ, Musher DM. Treatment of *Clostridium difficile*-associated disease: old therapies and new strategies. *Lancet Infect Dis.* 2005;5: 549–557. doi:10.1016/S1473-3099(05)70215-2
 155. Rea MC, Dobson A, O’Sullivan O, Crispie F, Fouhy F, Cotter PD, et al. Effect of broad- and narrow-spectrum antimicrobials on *Clostridium difficile* and microbial diversity in a model of the distal colon. *Proc Natl Acad Sci.* 2011;108: 4639–4644. doi:10.1073/pnas.1001224107
 156. Yamada T, Hiraoka Y, Ikehata M, Kimbara K, Avner BS, Das Gupta TK, et al. Apoptosis or growth arrest: Modulation of tumor suppressor p53’s specificity by bacterial redox protein azurin. *Proc Natl Acad Sci.* 2004;101: 4770–4775. doi:10.1073/pnas.0400899101
 157. Aranha C, Gupta S, Reddy KVR. Contraceptive efficacy of antimicrobial peptide Nisin: in vitro and in vivo studies. *Contraception.* 2004;69: 333–338. doi:10.1016/j.contraception.2003.11.002
 158. Sutyak KE, Anderson RA, Dover SE, Feathergill KA, Aroutcheva AA, Faro S, et al. Spermicidal Activity of the Safe Natural Antimicrobial Peptide Subtilosin. *Infect Dis Obstet Gynecol.* 2008;2008: 1–6. doi:10.1155/2008/540758
 159. Kaur B, Balgir PP, Mittu B, Kumar B, Garg N. Biomedical Applications of Fermenticin HV6b Isolated from *Lactobacillus fermentum* HV6b MTCC10770. *BioMed Res Int.* 2013;2013: 1–8. doi:10.1155/2013/168438
 160. Silkin L, Hamza S, Kaufman S, Cobb SL, Vederas JC. Spermicidal bacteriocins: Lacticin 3147 and subtilosin A. *Bioorg Med Chem Lett.* 2008;18: 3103–3106. doi:10.1016/j.bmcl.2007.11.024
 161. Soltani S, Hammami R, Cotter PD, Rebuffat S, Said LB, Gaudreau H, et al. Bacteriocins as a new generation of antimicrobials: toxicity aspects and regulations. *FEMS Microbiol Rev.* 2021;45: fuaa039. doi:10.1093/femsre/fuaa039
 162. Kamada N, Chen GY, Inohara N, Núñez G. Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol.* 2013;14: 685–690. doi:10.1038/ni.2608
 163. Stavric S, D’Aoust J-Y. Undefined and Defined Bacterial Preparations for the Competitive Exclusion of *Salmonella* in Poultry - A Review. *J Food Prot.* 1993;56: 173–180. doi:10.4315/0362-028X-56.2.173
 164. Cutler SA, Lonergan SM, Cornick N, Johnson AK, Stahl CH. Dietary Inclusion of Colicin E1 Is Effective in Preventing Postweaning Diarrhea Caused by F18-Positive *Escherichia coli* in Pigs. *Antimicrob Agents Chemother.* 2007;51: 3830–3835. doi:10.1128/AAC.00360-07
 165. EFSA Panel on Additives and Products or Substances used in Animal Feed (FEEDAP), Bampidis V, Azimonti G, Bastos M de L, Christensen H, Dusemund B, et al. Assessment of the application for renewal of authorisation of Bactocell® (*Pediococcus acidilactici* CNCM I-4622) as a feed additive for all fish and shrimps and its extension of use for all crustaceans. *EFSA J.* 2019;17: e05691. doi:10.2903/j.efsa.2019.5691
 166. Gauthier, Rouleau-Breton, Charette, Derome. Stimulated Growth and Innate Immunity in Brook Charr (*Salvelinus fontinalis*) Treated with a General Probiotic (Bactocell®) and Two Endogenous Probiotics That Inhibit *Aeromonas salmonicida* In Vitro. *Microorganisms.* 2019;7: 193. doi:10.3390/microorganisms7070193

167. Scientific Opinion on Bactocell PA 10 (*Pediococcus acidilactici*) as a feed additive for laying hens. EFSA J. [cited 6 Aug 2021]. doi:10.2903/j.efsa.2010.1865
168. Scientific Opinion on the safety and efficacy of Bactocell PA 10 (*Pediococcus acidilactici*) as a feed additive for piglets. EFSA J. [cited 6 Aug 2021]. doi:10.2903/j.efsa.2010.1660
169. Holtsmark I, Eijsink VGH, Brurberg MB. Bacteriocins from plant pathogenic bacteria. FEMS Microbiol Lett. 2008;280: 1–7. doi:10.1111/j.1574-6968.2007.01010.x
170. Grinter R, Milner J, Walker D. Bacteriocins active against plant pathogenic bacteria. Biochem Soc Trans. 2012;40: 1498–1502. doi:10.1042/BST20120206
171. Kerr A, Htay K. Biological control of crown gall through bacteriocin production. Physiol Plant Pathol. 1974;4: 37–44. doi:10.1016/0048-4059(74)90042-3
172. Jones DA, Ryder MH, Clare BG, Farrand SK, Kerr A. Construction of a Tra–deletion mutant of pAgK84 to safeguard the biological control of crown gall. Mol Gen Genet MGG. 1988;212: 207–214. doi:10.1007/BF00334686
173. Kerr A, Bullard G. Biocontrol of Crown Gall by *Rhizobium rhizogenes*: Challenges in Biopesticide Commercialisation. Agronomy. 2020;10. doi://doi.org/10.3390/agronomy10081126
174. Lavermicocca P, Lonigro SL, Valerio F, Evidente A, Visconti A. Reduction of Olive Knot Disease by a Bacteriocin from *Pseudomonas syringae* pv. *ciccaronei*. Appl Environ Microbiol. 2002;68: 1403–1407. doi:10.1128/AEM.68.3.1403-1407.2002
175. Hammami I, Triki MA, Rebai A. PURIFICATION AND CHARACTERIZATION OF THE NOVEL BACTERIOCIN BAC IH7 WITH ANTIFUNGAL AND ANTIBACTERIAL PROPERTIES. J Plant Pathol. 2011;93: 443–454.
176. Rooney WM, Grinter RW, Correia A, Parkhill J, Walker DC, Milner JJ. Engineering bacteriocin-mediated resistance against the plant pathogen *Pseudomonas syringae*. Plant Biotechnol J. 2020;18: 1296–1306. doi:10.1111/pbi.13294
177. USDA/APHIS. Agricultural Bioterrorism Protection Act of 2002; Possession, Use, and Transfer of Biological Agents and Toxins. 2005 p. 52.
178. Stackebrandt E, Murray RGE, Truper HG. Proteobacteria classis nov., a Name for the Phylogenetic Taxon That Includes the “Purple Bacteria and Their Relatives.” Int J Syst Bacteriol. 1988;38: 321–325. doi:10.1099/00207713-38-3-321
179. Denny TP, Hayward AC. *Ralstonia solanacearum*. American Phytopathological Society (APS Press). Laboratory guide for identification of plant pathogenic bacteria. American Phytopathological Society (APS Press). St. Paul: Schaad, N. W.; Jones, J. B.; Chun, W.; 2001. pp. 151–174.
180. Kelman A. The bacterial wilt caused by *Pseudomonas solanacearum*. Tech Bull N C Agric Exp Stn. 1953.
181. Zheng X, Zhu Y, Wang J, Wang Z, Liu B. Combined use of a microbial restoration substrate and avirulent *Ralstonia solanacearum* for the control of tomato bacterial wilt. 2019; 11.
182. Prasannakumar MK, Chandrashekara KN, Deepa M, Vani A. Davana: A New Host Plant for *Ralstonia solanacearum* from India. 2011; 9.
183. Salanoubat M, Genin S, Artiguenave F, Gouzy J, Mangenot S, Arlat M, et al. Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature. 2002;415: 497–502. doi:10.1038/415497a
184. Yabuuchi E, Kosako Y, Yano I, Hotta H, Nishiuchi Y. Transfer of Two *Burkholderia* and An *Alcaligenes* Species to *Ralstonia* Gen. Nov.: Proposal of *Ralstonia pickettii* (Ralston, Palleroni and Doudoroff 1973) Comb. Nov., *Ralstonia solanacearum* (Smith

- 1896) Comb. Nov. and *Ralstonia eutro*. Microbiol Immunol. 1995;39: 897–904. doi:10.1111/j.1348-0421.1995.tb03275.x
185. Yabuuchi E, Kosako Y, Oyaizu H, Yano I, Hotta H, Hashimoto Y, et al. Proposal of Burkholderia gen. nov. and transfer of seven species of the genus Pseudomonas homology group II to the new genus, with the type species Burkholderia cepacia (Palleroni and Holmes 1981) comb. nov. Microbiol Immunol. 1992;36: 1251–1275.
186. Safni I, Cleenwerck I, De Vos P, Fegan M, Sly L, Kappler U. Polyphasic taxonomic revision of the *Ralstonia solanacearum* species complex: proposal to emend the descriptions of *Ralstonia solanacearum* and *Ralstonia syzygii* and reclassify current *R. syzygii* strains as *Ralstonia syzygii subsp. syzygii subsp. nov.*, *R. solanacearum* phylotype IV strains as *Ralstonia syzygii subsp. indonesiensis subsp. nov.*, banana blood disease bacterium strains as *Ralstonia syzygii subsp. celebensis subsp. nov.* and *R. solanacearum* phylotype I and III strains as *Ralstonia pseudosolanacearum sp. nov.* Int J Syst Evol Microbiol. 2014;64: 3087–3103. doi:10.1099/ij.s.0.066712-0
187. Prior P, Ailloud F, Dalsing BL, Remenant B, Sanchez B, Allen C. Genomic and proteomic evidence supporting the division of the plant pathogen *Ralstonia solanacearum* into three species. BMC Genomics. 2016;17: 90. doi:10.1186/s12864-016-2413-z
188. Fegan M, Prior P. How complex is the “*Ralstonia solanacearum* Species Complex.” APS Press. Bacterial wilt disease and the *Ralstonia solanacearum* species complex. APS Press. Minnesota: Allen, C., Prior, P., Hayward, A.C.; 2005. pp. 449–461.
189. Digat B, Caffier D. “Pourriture brune” de la pomme de terre, flétrissement bactérien sur la tomate...Alerte face à une redoutable maladie des solanacées. Phytoma- Déf Végétaux. 1996;482: 33–37.
190. Vasse J, Danoun S, Trigalet A. Microscopic studies of root infection in resistant tomato cultivar Hawaii 7996. APS Press. Bacterial wilt disease and the *Ralstonia solanacearum* species complex. APS Press. St. Paul: Allen, C., Prior, P., Hayward, A.C.; 2005. pp. 285–291.
191. Lowe-Power TM, Khokhani D, Allen C. How *Ralstonia solanacearum* Exploits and Thrives in the Flowing Plant Xylem Environment. Trends Microbiol. 2018;26: 929–942. doi:10.1016/j.tim.2018.06.002
192. Smith EF. The brown rot of solanaceae. W.B. Saunders Company. Bacterial diseases of plants. W.B. Saunders Company. Londre: Smith; 1920. pp. 177–291.
193. Genin S. Molecular traits controlling host range and adaptation to plants in *Ralstonia solanacearum*: Research review. New Phytol. 2010;187: 920–928. doi:10.1111/j.1469-8137.2010.03397.x
194. AVRDC. Bacterial Wilt management in Tomato. Training course guide. : 18.
195. Saddler GS. Management of bacterial wilt disease. MN: APS press. Bacterial Wilt Disease and the *Ralstonia solanacearum* Species Complex. MN: APS press. Saint Paul: Allen C., Prior P., Hayward A.C.; 2005. pp. 121–132.
196. Lebeau A. Resistance de la tomate, l’aubergine et le piment à *Ralstonia solanacearum*: interactions entre les géniteurs de résistance et la diversité bactérienne, caractérisation et cartographie des facteurs génétiques impliqués chez l’aubergine. 2010; 226.
197. Aliye N, Fininsa C, Hiskias Y. Evaluation of rhizosphere bacterial antagonists for their potential to bioprotect potato (*Solanum tuberosum*) against bacterial wilt (*Ralstonia solanacearum*). Biol Control. 2008;47: 282–288. doi:10.1016/j.biocontrol.2008.09.003
198. Anuratha CS, Gnanamanickam SS. Biological control of bacterial wilt caused by

- Pseudomonas solanacearum* in India with antagonistic bacteria. *Plant Soil*. 1990;124: 109–116.
199. Liu Y, Shi J, Feng Y, Yang X, Li X, Shen Q. Tobacco bacterial wilt can be biologically controlled by the application of antagonistic strains in combination with organic fertilizer. *Biol Fertil Soils*. 2013;49: 447–464.
 200. Ramesh R, Joshi AA, Ghanekar MP. *Pseudomonas*: major antagonistic endophytic bacteria to suppress bacterial wilt pathogen, *Ralstonia solanacearum* in the eggplant (*Solanum melongena* L.). *World J Microbiol Biotechnol*. 2009;25: 47–55. doi:10.1007/s11274-008-9859-3
 201. Álvarez B, Biosca EG. Bacteriophage-Based Bacterial Wilt Biocontrol for an Environmentally Sustainable Agriculture. *Front Plant Sci*. 2017;8: 1218. doi:10.3389/fpls.2017.01218
 202. Tjou-Tam-Sin NNA, van de Bilt JIJ, Westenberg M, Bergsma-Vlami M, Korpershoek HJ, Vermunt AMW, et al. First Report of Bacterial Wilt Caused by *Ralstonia solanacearum* in Ornamental Rosa sp. *Plant Dis*. 2017;101: 378–378. doi:10.1094/PDIS-02-16-0250-PDN
 203. Yahiaoui N, Chéron J-J, Ravelomanantsoa S, Hamza AA, Petrouse B, Jeetah R, et al. Genetic diversity of the *Ralstonia solanacearum* species complex in the South-West Indian Ocean islands. *Front Plant Sci*. 2017;8: 2139. doi:10.3389/fpls.2017.02139
 204. Okabe N. Studies on *Pseudomonas solanacearum*. V. Antagonism among the strains of *P. solanacearum*. *Rep Fac Of Agric Shizuoka Univ*. 1954;4: 37–40.
 205. Cuppels DA, Hanson RS, Kelman A. Isolation and characterization of a bacteriocin produced by *Pseudomonas solanacearum*. *J Gen Microbiol*. 1978;109: 295–303. doi:10.1099/00221287-109-2-295
 206. He LY, Fan YL, Lie DX, Fang F. Intraspecific antagonism and bacteriocin production of *Pseudomonas solanacearum*. *Fifth Int Congr Plant Pathog*. 1988; 20–27.
 207. Frey P, Smith JJ, Albar L, Prior P, Saddler GS, Trigalet-Demery D, et al. Bacteriocin typing of *Burkholderia (Pseudomonas) solanacearum* race 1 of the French West Indies and correlation with genomic variation of the pathogen. *Appl Environ Microbiol*. 1996;62: 7.
 208. Fabre A. Activités et typage bactériocines du complexe d'espèces *Ralstonia solanacearum*, agent du flétrissement bactérien, dans le Sud-Ouest de l'Océan Indien. Université de La Réunion; 2016 p. 62.
 209. Gross DC, Vidaver AK. Bacteriocins. *Methods Phytobacteriology Akad Kiado Bp Hung*. 1990; 245–249.
 210. Frey P, Prior P, Marie C, Kotoujansky A, Trigalet-Demery D, TRIGALETI A. Hrp-mutants of *Pseudomonas solanacearum* as potential biocontrol agents of tomato bacterial wilt. *Appl Environ Microbiol*. 1994;60: 7.
 211. Wilson K. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol*. 2001;56: 2.4. 1-2.4. 5.
 212. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*. 2017;27: 722–736.
 213. Seppy M, Manni M, Zdobnov EM. BUSCO: Assessing Genome Assembly and Annotation Completeness. In: Kollmar M, editor. *Gene Prediction*. New York, NY: Springer New York; 2019. pp. 227–245. doi:10.1007/978-1-4939-9173-0_14
 214. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;29: 1072–1075.

215. Cabanettes F, Klopp C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ*. 2018;6: e4958.
216. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30: 2068–2069.
217. van Heel AJ, de Jong A, Song C, Viel JH, Kok J, Kuipers OP. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res*. 2018;46: W278–W281. doi:10.1093/nar/gky383
218. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020;37: 1530–1534.
219. Seemann T. Snippy: fast bacterial variant calling from NGS reads. 2015.
220. Huerta AI dissertant. Identification and characterization of interstrain chemical weapons in a major plant pathogen. [Madison, Wis.] : [University of Wisconsin--Madison], 2015.; 2015. Available: <https://search.library.wisc.edu/catalog/9911176396102121>
221. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31: 3691–3693.
222. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics*. 2015;16: 1–7.
223. Rasoamanana H, Ravelomanantsoa S, Yahiaoui N, Dianzinga N, Rébert E, Gauche M-M, et al. Contrasting genetic diversity and structure among Malagasy *Ralstonia pseudosolanacearum* phylotype I populations inferred from an optimized Multilocus Variable Number of Tandem Repeat Analysis scheme. Lee S-W, editor. *PLOS ONE*. 2020;15: e0242846. doi:10.1371/journal.pone.0242846
224. Li Z. A simple plant high-molecular-weight DNA extraction method suitable for single-molecule technologies. 2020; 6.
225. Mayjonade B, Gouzy J, Donnadiou C, Pouilly N, Marande W, Callot C, et al. Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. *BioTechniques*. 2016;61: 203–205. doi:10.2144/000114460
226. Tsai JW, Hsu S-T, Chen L-C. Bacteriocin-producing strains of *Pseudomonas solanacearum* and their effect on development of bacterial wilt of tomato. 1985.
227. Yoshida K, Burbano HA, Krause J, Thines M, Weigel D, Kamoun S. Mining Herbaria for Plant Pathogen Genomes: Back to the Future. Heitman J, editor. *PLoS Pathog*. 2014;10: e1004028. doi:10.1371/journal.ppat.1004028
228. FAO. International year of plant health. 2020 p. 20.
229. Sabat AJ, Budimir A, Nashev D, Sá-Leão R, van Dijl JM, Laurent F, et al. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Eurosurveillance*. 2013;18. doi:10.2807/ese.18.04.20380-en
230. Farber JM. An Introduction to the Hows and Whys of Molecular Typing†. *J Food Prot*. 1996;59: 1091–1101. doi:10.4315/0362-028X-59.10.1091
231. Wicker E, Lefeuvre P, de Cambiaire J-C, Lemaire C, Poussier S, Prior P. Contrasting recombination patterns and demographic histories of the plant pathogen *Ralstonia solanacearum* inferred from MLSA. *ISME J*. 2012;6: 961–974. doi:10.1038/ismej.2011.160
232. Francisco AP, Bugalho M, Ramirez M, Carriço JA. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics*. 2009;10: 1–15.
233. Pitel F, Riquet J. Les marqueurs anonymes et la détection de leur polymorphisme.

- INRA Prod Anim Numéro Hors Sér «Génétique Moléculaire Principes Appl Aux Popul Anim. 2000; 45–53.
234. Najimi B, El Jaafari S, Jlibène M, Jacquemin J-M. Applications des marqueurs moléculaires dans l'amélioration du blé tendre pour la résistance aux maladies et aux insectes. BASE. 2003.
 235. Fournet-Guérin C, Vacchiani-Marcuzzo C. Vers de nouvelles interactions entre les Suds? Rev Tiers Monde. 2013; 141–160.
 236. Mathieu M, Zaïdani M. Typologie des échanges commerciaux entre Mayotte et sa région entre 1994 et 2001. Inst D'Emission Dép D'outre-Mer. 2002;5: 5.
 237. Rabemanambola M, Rakotoarisoa J, Rieutort L. Entre ville et campagne: les adaptations du maraîchage paysan sur les Hautes Terres centrales malgaches. Cah D'Outre-Mer. 2009; 27.
 238. Yahiaoui N. Epidémiologie moléculaire du complexe d'espèces *Ralstonia solanacearum*, agent du flétrissement bactérien, dans les îles du Sud-Ouest de l'Océan Indien. Université de La Réunion; 2018.
 239. Saenger W. Chapter 714 - Proteinase K. In: Rawlings ND, Salvesen G, editors. Handbook of Proteolytic Enzymes (Third Edition). Academic Press; 2013. pp. 3240–3242. doi:10.1016/B978-0-12-382219-2.00714-6
 240. Farrell RE. Chapter 2 - Creating a Ribonuclease-Free Environment. In: Farrell RE, editor. RNA Methodologies (Fifth Edition). Academic Press; 2017. pp. 55–74. doi:10.1016/B978-0-12-804678-4.00002-6
 241. Berić T, Stanković S, Draganić V, Kojić M, Lozo J, Fira D. Novel antilisterial bacteriocin licheniocin 50.2 from *Bacillus licheniformis* VPS 50.2 isolated from soil sample. J Appl Microbiol. 2014;116: 502–510.
 242. Zhao R, Duan G, Yang T, Niu S, Wang Y. Purification, characterization and antibacterial mechanism of bacteriocin from *Lactobacillus acidophilus* XH1. Trop J Pharm Res. 2015;14: 989–995.
 243. Chai R, Rooney WM, Milner JJ, Walker D. Challenges of using protein antibiotics for pathogen control. Pest Manag Sci. 2021;77: 3836–3840. doi:10.1002/ps.6312
 244. Schulz S, Stephan A, Hahn S, Bortesi L, Jarczowski F, Bettmann U, et al. Broad and efficient control of major foodborne pathogenic strains of *Escherichia coli* by mixtures of plant-produced colicins. Proc Natl Acad Sci. 2015;112: E5454–E5460.
 245. Chuang D-Y, Kyeremeh AG, Gunji Y, Takahara Y, Ehara Y, Kikumoto T. Identification and cloning of an *Erwinia carotovora* subsp. *carotovora* bacteriocin regulator gene by insertional mutagenesis. J Bacteriol. 1999;181: 1953–1957.
 246. Príncipe A, Fernandez M, Torasso M, Godino A, Fischer S. Effectiveness of tailocins produced by *Pseudomonas fluorescens* SF4c in controlling the bacterial-spot disease in tomatoes caused by *Xanthomonas vesicatoria*. Microbiol Res. 2018;212–213: 94–102. doi:10.1016/j.micres.2018.05.010
 247. Curpanen E. Isolement de bactériocines produites par des lignées du complexe d'espèces *Ralstonia solanacearum*, agent responsable du flétrissement bactérien dans le Sud-Ouest de l'Océan Indien. Université de La Réunion: Université de La Réunion; 2021 p. 64.

ANNEXES

Annexes chapitre 1

S1 Table. Characteristics of the 148 Malagasy isolates used in the study. The isolates used for the comparative analysis of MLST-7 and RS1-MLVA14 are in bold.

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN0304	2006	I-18	<i>Solanum lycopersicum</i>	1273.3	Analamanga	16	Ambatomirahavavy	1	MCC1	MT001
RUN0305	2006	I-18	<i>Solanum lycopersicum</i>	1273.3	Analamanga	16	Ambatomirahavavy	1	MCC1	MT002
RUN0306	2006	I-18	<i>Capsicum annuum</i>	74.6	Atsinanana	3	Fanandrana	1	MCC1	NA
RUN0307	2006	I-18	<i>Capsicum annuum</i>	74.6	Atsinanana	3	Fanandrana	1	MCC1	NA
RUN0308	2006	I-18	<i>Capsicum annuum</i>	74.6	Atsinanana	3	Fanandrana	1	MCC1	MT003
RUN0309	2006	I-18	<i>Capsicum annuum</i>	74.6	Atsinanana	3	Fanandrana	1	MCC1	MT005
RUN0310	2006	I-18	<i>Capsicum annuum</i>	74.6	Atsinanana	3	Fanandrana	1	MCC1	MT003
RUN0311	2006	I-18	<i>Capsicum annuum</i>	74.6	Atsinanana	3	Fanandrana	1	MCC1	MT003
RUN0312	2006	I-18	<i>Solanum melongena</i>	11.9	Atsinanana	2	Melville	1	MCC1	MT004
RUN0313	2006	I-18	<i>Solanum melongena</i>	11.9	Atsinanana	2	Melville	1	MCC1	MT006
RUN0314	2006	I-18	<i>Solanum melongena</i>	11.9	Atsinanana	2	Melville	1	MCC1	MT007
RUN0315	2006	I-18	<i>Solanum melongena</i>	11.9	Atsinanana	2	Melville	1	MCC1	MT008
RUN0316	2006	I-18	<i>Solanum melongena</i>	11.9	Atsinanana	2	Melville	1	MCC1	MT008
RUN0317	2006	I-18	<i>Solanum scabrum</i>	13.5	Atsinanana	4	Mangarano	1	MCC1	MT009
RUN0318	2006	I-18	<i>Solanum melongena</i>	13.5	Atsinanana	4	Mangarano	1	MCC1	MT009
RUN0319	2006	I-46	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	2	MCC3	MT010
RUN0320	2006	I-46	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	2	MCC3	MT010
RUN0321	2006	I-46	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	2	MCC3	MT011
RUN0322	2006	I-18	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	1	MCC1	MT012
RUN0323	2006	I-46	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	2	MCC3	MT011
RUN0324	2006	I-46	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	2	Singleton	MT013

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN0325	2006	I-18	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	1	MCC1	MT012
RUN0326	2006	I-46	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	2	MCC3	MT014
RUN0327	2006	I-46	<i>Capsicum annuum</i>	8	Atsinanana	1	Vohidrotra	2	MCC3	MT010
RUN0328	2006	I-18	<i>Capsicum annuum</i>	74.6	Atsinanana	3	Fanandrana	1	MCC1	MT003
RUN0329	2006	I-18	<i>Solanum scabrum</i>	13.5	Atsinanana	4	Mangarano	1	MCC1	MT009
RUN0330	2006	I-18	<i>Solanum melongena</i>	13.5	Atsinanana	4	Mangarano	1	MCC1	MT009
RUN3075	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3076	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3077	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3078	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3080	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT007
RUN3082	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC2	MT015
RUN3083	2013	I-na	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3084	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT007
RUN3085	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3086	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3087	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT007
RUN3089	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN3108	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3109	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT016
RUN3110	2013	I-na	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3111	2013	I-na	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3112	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT007
RUN3113	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT007
RUN3114	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3116	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3117	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3118	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3119	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT017
RUN3120	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT017
RUN3122	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004
RUN3124	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC2	MT015
RUN3125	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT017

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN3126	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT016
RUN3128	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Analamanga	14	Fiadanankely	1	MCC2	MT018
RUN3129	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Analamanga	14	Fiadanankely	1	MCC2	MT018
RUN3130	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Analamanga	14	Fiadanankely	1	MCC2	MT019
RUN3131	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Analamanga	14	Fiadanankely	1	MCC1	MT020
RUN3134	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Analamanga	14	Fiadanankely	1	MCC2	MT015
RUN3136	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Analamanga	14	Fiadanankely	1	MCC1	MT020
RUN3137	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Analamanga	14	Fiadanankely	1	MCC1	MT020
RUN3138	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Analamanga	14	Fiadanankely	1	MCC2	MT015
RUN3142	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT020
RUN3148	2013	I-33	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	2	MCC4	MT021
RUN3149	2013	I-33	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	2	MCC4	MT021
RUN3151	2013	I-33	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	2	MCC4	MT021
RUN3159	2013	I-18	<i>Solanum lycopersicum</i>	1267.3	Analamanga	15	Mandrosoa	1	MCC1	MT022
RUN3162	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Analamanga	15	Mandrosoa	1	MCC1	MT022
RUN3163	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Analamanga	15	Mandrosoa	1	MCC2	MT015
RUN3164	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Analamanga	15	Mandrosoa	1	MCC1	MT023
RUN3165	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Analamanga	15	Mandrosoa	1	MCC1	MT023

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN3168	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Analamanga	15	Mandrosoa	1	MCC2	MT018
RUN3170	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Analamanga	15	Mandrosoa	1	MCC1	MT023
RUN3171	2013	I-18	<i>Solanum nigrum</i>	1267.3	Analamanga	15	Mandrosoa	1	MCC1	MT023
RUN3180	2013	I-18	<i>Solanum tuberosum</i>	1476.9	Itasy	11	Ambohitaza Imerinatsimo	1	MCC1	MT023
RUN3208	2013	I-18	<i>Solanum tuberosum</i>	1484	Itasy	10	Ambohitaza Imerinatsimo	1	MCC1	MT024
RUN3210	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT004
RUN3211	2013	I-33	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	2	MCC4	MT025
RUN3212	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT016
RUN3213	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT006
RUN3214	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT004
RUN3215	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT004
RUN3216	2013	I-33	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	2	MCC4	MT025
RUN3217	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT004
RUN3218	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT004
RUN3219	2013	I-na	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT006
RUN3220	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT006

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN3221	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT006
RUN3222	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT006
RUN3223	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	Singleton	MT026
RUN3224	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT006
RUN3226	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT004
RUN3228	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Itasy	5	Analavory	1	MCC1	MT006
RUN3243	2013	I-18	<i>Capsicum annuum</i>	1190.3	Itasy	6	Analavory	1	MCC1	MT020
RUN3245	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3246	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3247	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3256	2013	I-18	<i>Solanum tuberosum</i>	1461.5	Itasy	9	Tamoronala	1	MCC2	MT015
RUN3257	2013	I-18	<i>Solanum tuberosum</i>	1461.5	Itasy	9	Tamoronala	1	MCC2	MT015
RUN3264	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3265	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3266	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT027
RUN3267	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN3268	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3269	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3270	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3271	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3272	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3273	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3274	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3275	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3276	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3277	2013	I-na	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3278	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3279	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3280	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3281	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3282	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN3283	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3284	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3285	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC2	MT015
RUN3286	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	1	MCC1	MT028
RUN3287	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	1	MCC1	MT007
RUN3288	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	1	MCC1	MT023
RUN3289	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	1	MCC1	MT024
RUN3290	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	1	MCC1	MT028
RUN3291	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	1	MCC1	MT023
RUN3292	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Analamanga	13	Bemasoandro	1	MCC1	MT028
RUN3457	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT029
RUN3458	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3459	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT030
RUN3460	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3461	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3462	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3463	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3464	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT031

RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Region	Plot number	Locality	Cluster	Malagasy Clonal Complex	Haplotype
RUN3465	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3466	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT027
RUN3467	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3468	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3469	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3470	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3471	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3472	2013	I-18	<i>Solanum aethiopicum</i>	1220	Itasy	8	Moratsiazo	1	MCC1	MT007
RUN3480	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3484	2013	I-18	<i>Solanum aethiopicum</i>	1224	Itasy	7	Ambanitavy	1	MCC1	MT007
RUN3507	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Itasy	12	Ampangabe	1	MCC1	MT004

S2 Table. Characteristics of the 145 worldwide isolates used in the study. The isolates used for the comparative analysis of MLST-7 and RS1-MLVA14 are in bold.

RUN.ID	World	Country	Year	Phylotype - Sequevar	Host	Haplotype
RUN0033	Asia	Japan	1965	I-16	<i>Nicotinia tabacum</i>	MT032
RUN0038	West Indies	Guadeloupe	1983	I-18	<i>Solanum melongena</i>	MT033
RUN0044	Australia	Australia	1966	I-16	<i>Zingiber officinale</i>	MT032
RUN0054	America (South)	French Guiana	1978	I-18	<i>Solanum lycopersicum</i>	MT034
RUN0078	Asia	Malaysia	na	I-48	<i>Solanum lycopersicum</i>	MT036
RUN0094	Asia	Indonesia	na	I-18	<i>Solanum lycopersicum</i>	MT034
RUN0095	Asia	Indonesia	na	I-15	<i>Solanum lycopersicum</i>	MT037
RUN0104	Africa	South Africa	1986	I-31	<i>Solanum lycopersicum</i>	MT035
RUN0114	America (South)	Brazil	1999	I-18	<i>Solanum lycopersicum</i>	MT038
RUN0119	Africa	Cameroon	2005	I-13	<i>Solanum lycopersicum</i>	MT039
RUN0157	Asia	Taiwan	1988	I-15	<i>Solanum lycopersicum</i>	MT040
RUN0194	Africa	Cameroon	2005	I-13	<i>Amaranthus</i>	MT041
RUN0205	Africa	Cameroon	2005	I-44	<i>Solanum lycopersicum</i>	MT042
RUN0215	Africa	Cameroon	2005	I-13	<i>Solanum scabrum</i>	MT044
RUN0257	Asia	Taiwan	1996	I-34	<i>Solanum lycopersicum</i>	MT045
RUN0300	America (South)	Brazil	2003	I-17	<i>Begonia semperflorens</i>	MT046
RUN0353	West Indies	Martinique	2002	I-18	<i>Capsicum annuum</i>	MT047
RUN0471	Indian Ocean	Reunion Island	1993	I-31	<i>Pelargonium asperum</i>	MT035
RUN0486	America (South)	Peru	1966	I-18	<i>Solanum lycopersicum</i>	MT048
RUN0487	Indian Ocean	Seychelles	na	I-31	<i>Solanum melongena</i>	MT035
RUN0504	Indian Ocean	Seychelles	na	I-31	<i>Capsicum annuum</i>	MT035
RUN0528	West Indies	Trinidad	2003	I-14	<i>Solanum lycopersicum</i>	MT049
RUN0577	Asia	Malaysia	1998	I-18	<i>Arachis hypogaea</i>	MT050
RUN0578	America (Central)	Costa Rica	1972	I-14	<i>Capsicum annuum</i>	MT051
RUN0608	Indian Ocean	Reunion Island	na	I-31	<i>Solanum tuberosum</i>	MT052
RUN0616	Asia	China	2004	I-15	<i>Solanum lycopersicum</i>	MT053

RUN.ID	World	Country	Year	Phylotype - Sequevar	Host	Haplotype
RUN0617	Asia	China	2004	I-34	<i>Solanum melongena</i>	MT054
RUN0771	West Indies	Martinique	2002	I-18	<i>Solanum lycopersicum</i>	MT055
RUN0796	West Indies	Martinique	2002	I-18	<i>Petroselinum crispum</i>	MT056
RUN0821	West Indies	Martinique	2003	I-18	<i>Cucurbita pepo</i>	MT057
RUN1068	West Indies	Martinique	2003	I-18	<i>Physalis angulata</i>	MT058
RUN1074	West Indies	Guadeloupe	2006	I-18	<i>Musa acuminata x balbisiana AAA</i>	MT059
RUN1110	America (Central)	Honduras	1997	I-14	<i>Solanum lycopersicum</i>	MT060
RUN1161	West Indies	Martinique	2003	I-18	<i>Anthurium andreanum</i>	MT061
RUN1401	West Indies	Guadeloupe	2007	I-18	<i>Capsicum annuum</i>	MT062
RUN1404	West Indies	Guadeloupe	2007	I-18	<i>Solanum lycopersicum</i>	MT063
RUN1431	Indian Ocean	Mayotte	2006	I-31	<i>Solanum lycopersicum</i>	MT064
RUN1432	Indian Ocean	Mayotte	2006	I-31	<i>Solanum lycopersicum</i>	MT035
RUN1439	Indian Ocean	Mayotte	2006	I-31	<i>Solanum lycopersicum</i>	MT035
RUN1495	West Indies	Trinidad	2008	I-14	<i>Solanum lycopersicum</i>	MT065
RUN1523	America (South)	Guyana	2008	I-18	<i>Solanum lycopersicum</i>	MT066
RUN1528	America (Central)	Guatemala	2004	I-14	<i>Solanum lycopersicum</i>	MT067
RUN1771	Africa	Ivory Coast	2010	I-18	<i>Solanum lycopersicum</i>	MT068
RUN1798	Africa	Ivory Coast	2010	I-13	<i>Solanum melongena</i>	MT069
RUN1854	Africa	Ivory Coast	2011	I-14	<i>Solanum lycopersicum</i>	MT070
RUN1869	Africa	Ivory Coast	2011	I-44	<i>Solanum lycopersicum</i>	MT071
RUN1874	Africa	Ivory Coast	2011	I-46	<i>Solanum lycopersicum</i>	MT072
RUN1897	Africa	Ivory Coast	2011	I-46	<i>Solanum lycopersicum</i>	MT073
RUN1906	Africa	Ivory Coast	2011	I-46	<i>Solanum lycopersicum</i>	MT074
RUN1930	Indian Ocean	Mayotte	2011	I-31	<i>Solanum lycopersicum</i>	MT043
RUN1948	America (South)	French Guiana	2011	I-17	<i>Capsicum annuum</i>	MT075
RUN1952	America (South)	French Guiana	2011	I-17	<i>Capsicum annuum</i>	MT076
RUN1958	America (South)	French Guiana	2011	I-18	<i>Capsicum annuum</i>	MT077
RUN1972	America (South)	French Guiana	2011	I-14	<i>Capsicum annuum</i>	MT078

RUN.ID	World	Country	Year	Phylotype - Sequevar	Host	Haplotype
RUN1975	America (South)	French Guiana	2011	I-18	<i>Solanum aethiopicum</i>	MT079
RUN2083	Indian Ocean	Mayotte	2012	I-18	<i>Capsicum frutescens</i>	MT080
RUN2127	Indian Ocean	Mayotte	2012	I-46	<i>Capsicum annuum</i>	MT081
RUN2144	Indian Ocean	Mayotte	2012	I-15	<i>Solanum melongena</i>	MT082
RUN2177	America (South)	French Guiana	2012	I-47	<i>Solanum lycopersicum</i>	MT083
RUN3489	Indian Ocean	Seychelles	2005	I-31	<i>Solanum melongena</i>	MT084
RUN3490	Indian Ocean	Seychelles	1998	I-31	<i>Solanum lycopersicum</i>	MT035
RUN3491	Indian Ocean	Comoros	2005	I-31	<i>Solanum lycopersicum</i>	MT043
RUN3492	Indian Ocean	Comoros	2005	I-31	<i>Phaseolus vulgaris</i>	MT043
RUN3493	Indian Ocean	Comoros	2005	I-31	<i>Solanum tuberosum</i>	MT043
RUN3494	Indian Ocean	Comoros	2005	I-31	<i>Solanum tuberosum</i>	MT043
RUN3495	Indian Ocean	Comoros	2005	I-31	<i>Solanum lycopersicum</i>	MT043
RUN3501	Indian Ocean	Mauritius	1957	I-33	<i>Vicia faba</i>	MT025
RUN3505	Indian Ocean	Mauritius	1957	I-33	<i>Brassica oleracea</i>	MT085
RUN3510	Indian Ocean	Mauritius	1957	I-31	<i>Casuarina equisetifolia</i>	MT035
RUN3556	Asia	Philippines	2010	I-17	<i>Capsicum annuum</i>	MT086
RUN3558	Asia	Philippines	2008	I-17	<i>Cucurbita maxima</i>	MT087
RUN3559	Asia	Philippines	2003	I-15	<i>Solanum lycopersicum</i>	MT037
RUN3560	Asia	Thailand	2009	I-31	<i>Momordica charantia</i>	MT088
RUN3563	Asia	Thailand	2007	I-31	<i>Solanum melongena</i>	MT088
RUN3564	Asia	Thailand	2011	I-31	<i>Tagetes sp.</i>	MT088
RUN3567	Asia	Thailand	2010	I-47	<i>Capsicum annuum</i>	MT089
RUN3573	Indian Ocean	Reunion Island	2014	I-33	<i>Solanum lycopersicum</i>	MT085
RUN3633	Africa	Bénin	2014	I-31	<i>Solanum macrocarpon</i>	MT052
RUN3663	America (Central)	El Salvador	2014	I-14	<i>Solanum lycopersicum</i>	MT090
RUN3711	Indian Ocean	Reunion Island	2014	I-31	<i>Pelargonium graveolens</i>	MT043
RUN3743	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum melongena</i>	MT035
RUN3756	Indian Ocean	Reunion Island	2014	I-31	<i>Solanum melongena</i>	MT035

RUN.ID	World	Country	Year	Phylotype - Sequevar	Host	Haplotype
RUN3770	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum melongena</i>	MT035
RUN3815	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum melongena</i>	MT035
RUN3841	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	MT035
RUN3856	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	MT043
RUN3869	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	MT035
RUN3884	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	MT043
RUN3893	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	MT035
RUN3928	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	MT035
RUN4076	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum nigrum</i>	MT035
RUN4079	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	MT035
RUN4082	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	MT043
RUN4105	Indian Ocean	Reunion Island	2015	I-31	<i>Capsicum annuum</i>	MT035
RUN4112	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum melongena</i>	MT035
RUN4196	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	MT035
RUN4210	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	MT035
RUN4257	Indian Ocean	Reunion Island	2015	I-31	<i>Phaseolus vulgaris</i>	MT035
RUN4263	Indian Ocean	Reunion Island	2015	I-31	<i>Amaranthus viridis</i>	MT035
RUN4292	Indian Ocean	Seychelles	2015	I-31	<i>Solanum melongena</i>	MT035
RUN4303	Indian Ocean	Seychelles	2015	I-31	<i>Capsicum annuum</i>	MT035
RUN4308	Indian Ocean	Seychelles	2015	I-31	<i>Solanum melongena</i>	MT035
RUN4311	Indian Ocean	Seychelles	2015	I-31	<i>Solanum lycopersicum</i>	MT035
RUN4315	Indian Ocean	Seychelles	2015	I-31	<i>Solanum melongena</i>	MT035
RUN4325	Indian Ocean	Seychelles	2015	I-31	<i>Amaranthus viridis</i>	MT035
RUN4326	Indian Ocean	Seychelles	2015	I-31	<i>Solanum lycopersicum</i>	MT035
RUN4337	Indian Ocean	Seychelles	2015	I-31	<i>Solanum melongena</i>	MT035
RUN4376	Indian Ocean	Mauritius	2015	I-31	<i>Solanum melongena</i>	MT035
RUN4392	Indian Ocean	Mauritius	2015	I-15	<i>Solanum lycopersicum</i>	MT091
RUN4393	Indian Ocean	Mauritius	2015	I-15	<i>Capsicum annuum</i>	MT091

RUN.ID	World	Country	Year	Phylotype - Sequevar	Host	Haplotype
RUN4400	Indian Ocean	Mauritius	2015	I-33	<i>Solanum lycopersicum</i>	MT025
RUN4406	Indian Ocean	Mauritius	2015	I-33	<i>Amaranthus</i>	MT025
RUN4407	Indian Ocean	Mauritius	2015	I-15	<i>Solanum lycopersicum</i>	MT092
RUN4438	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum muricatum</i>	MT035
RUN4450	Indian Ocean	Reunion Island	2015	I-31	<i>Capsicum annum</i>	MT035
RUN4584	Indian Ocean	Mauritius	2015	I-18	<i>Solanum tuberosum</i>	MT093
RUN4603	Indian Ocean	Mauritius	2015	I-33	<i>Solanum tuberosum</i>	MT025
RUN4645	Indian Ocean	Mauritius	2015	I-18	<i>Solanum lycopersicum</i>	MT093
RUN4664	Indian Ocean	Mauritius	2015	I-15	<i>Solanum tuberosum</i>	MT091
RUN4680	Indian Ocean	Mauritius	2015	I-33	<i>Solanum tuberosum</i>	MT094
RUN5007	Indian Ocean	Mauritius	2015	I-15	<i>Solanum tuberosum</i>	MT095
RUN5135	Indian Ocean	Mauritius	2015	I-18	<i>Cyphostemma mappia</i>	MT096
RUN5137	Indian Ocean	Mauritius	2015	I-15	<i>Capsicum frutescens</i>	MT097
RUN5253	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	MT035
RUN5277	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum tuberosum</i>	MT035
RUN5385	Indian Ocean	Mauritius	2016	I-33	<i>Anthurium</i>	MT025
RUN5445	Indian Ocean	Rodrigues	2016	I-33	<i>Capsicum frutescens</i>	MT098
RUN5453	Indian Ocean	Mauritius	2016	I-31	<i>Solanum lycopersicum</i>	MT043
RUN5454	Indian Ocean	Mauritius	2016	I-33	<i>Anthurium</i>	MT025
RUN5479	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	MT043
RUN5480	Indian Ocean	Comoros	2016	I-31	<i>Solanum tuberosum</i>	MT043
RUN5481	Indian Ocean	Comoros	2016	I-31	<i>Solanum melongena</i>	MT035
RUN5482	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	MT043
RUN5485	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	MT099
RUN5489	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	MT100
RUN5495	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	MT043
RUN5508	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	MT043
RUN5516	Indian Ocean	Comoros	2016	I-31	<i>Solanum melongena</i>	MT035

RUN.ID	World	Country	Year	Phylotype - Sequevar	Host	Haplotype
RUN5521	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	MT035
RUN5525	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	MT043
RUN5564	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum lycopersicum</i>	MT101
RUN5565	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum tuberosum</i>	MT102
RUN5566	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum lycopersicum</i>	MT103
RUN5567	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum lycopersicum</i>	MT104
RUN5570	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum tuberosum</i>	MT105

S3 Table. Screening of the 14 selected VNTR loci in the 114 publicly available RSSC genomes.

GenBank assembly accession	Phylotype- Sequevar	VNTR													
		GMIch_0581	RS3L19	GMIch_3461	GMIch_1457	GMImp_0487	RS1L14	GMImp_1343	GMImp_0449	RS3L33	GMImp_0740	GMImp_0618	GMImp_0266	RS1L10	GMImp_1618
GCA_003576625.1	I-13	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001876975.1	I-13	Yes	Yes	Yes*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001484095.1	I-13	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001920895.1	I-14	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes*	Yes*	No	Yes*	Yes	Yes
GCA_000348545.1	I-14	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001644865.1	I-14	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001645725.1	I-14	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001663415.1	I-14	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_008271775.1	I-15	Yes	Yes	Yes*	Yes	Yes	Yes	Yes*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_008271845.1	I-15	Yes	Yes	Yes*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_008271855.1	I-15	Yes	Yes	Yes*	Yes	Yes	Yes	Yes*	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001887535.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001920905.1	I-15	Yes	Yes	Yes	Yes*	Yes	Yes	Yes*	Yes	Yes*	Yes*	No	Yes	Yes	Yes
GCA_001870825.1	I-15	Yes	Yes	Yes	Yes*	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
GCA_001876985.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515405.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515205.1	I-15	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes

GenBank assembly accession	Phylotype- Sequevar	VNTR													
		GMIch_0581	RS3L19	GMIch_3461	GMIch_1457	GMImp_0487	RS1L14	GMImp_1343	GMImp_0449	RS3L33	GMImp_0740	GMImp_0618	GMImp_0266	RS1L10	GMImp_1618
GCA_003515625.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515365.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515605.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515585.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515245.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515565.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes
GCA_003595325.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515545.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515285.1	I-15	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001708525.1	I-15	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
GCA_002155245.1	I-16	Yes	Yes	Yes	Yes	Yes	Yes*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_000430925.2	I-16	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001855495.1	I-17	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_002220465.1	I-17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_002162015.1	I-17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_000009125.1	I-18	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_002930085.2	I-30	Yes	Yes	No	Yes	Yes	Yes	Yes*	Yes	Yes	Yes*	Yes*	Yes	Yes	Yes
GCA_001267515.1	I-30	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_000671335.1	I-33	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

GenBank assembly accession	Phylotype- Sequevar	VNTR													
		GMIch_0581	RS3L19	GMIch_3461	GMIch_1457	GMImp_0487	RS1L14	GMImp_1343	GMImp_0449	RS3L33	GMImp_0740	GMImp_0618	GMImp_0266	RS1L10	GMImp_1618
GCA_001891105.1	I-34	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001879565.1	I-34	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001870805.1	I-34	Yes	Yes	Yes*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_001854265.1	I-44	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515345.1	I-44	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515305.1	I-44	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_002894285.1	I-45	Yes	Yes	Yes	Yes	Yes	Yes*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_000671315.1	I-48	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_000223115.2	I-54	Yes	Yes	Yes	Yes*	Yes	Yes	Yes*	Yes	Yes*	Yes	Yes	Yes	Yes	Yes
GCA_003999715.1	I-x	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515225.1	I-x	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_003515465.1	I-x	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GCA_000825785.2	IIA-24	No	No	No	No	No	No	No	No	Yes	No	No	Yes	Yes	Yes
GCA_001373275.1	IIA-24	No	No	No	No	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes
GCA_001644805.1	IIA-38	No	No	No	No	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes
GCA_003337165.1	IIA-38	No	No	No	No	Yes	No	No	Yes	Yes	No	No	Yes	No	Yes
GCA_001644795.1	IIA-38	No	No	No	No	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes
GCA_003590585.1	IIA-53	No	Yes	No	No	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes
GCA_003860665.1	IIA-6	No	No	No	No	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes

GenBank assembly accession	Phylotype- Sequevar	VNTR													
		GMIch_0581	RS3L19	GMIch_3461	GMIch_1457	GMImp_0487	RS1L14	GMImp_1343	GMImp_0449	RS3L33	GMImp_0740	GMImp_0618	GMImp_0266	RS1L10	GMImp_1618
		GCA_003860685.1	IIA-6	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	No
GCA_000825845.2	IIA-6	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001901665.1	IIA-7	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002251605.2	IIA-7	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002251695.1	IIA-7	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003576005.1	IIA-7	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002549815.1	IIA-x	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_009832805.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_000750585.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	No	Yes
GCA_001299555.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002029865.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002029885.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002029895.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002251655.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002501565.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002894765.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002894785.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002894795.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002894845.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes

GenBank assembly accession	Phylotype- Sequevar	VNTR													
		GMIch_0581	RS3L19	GMIch_3461	GMIch_1457	GMImp_0487	RS1L14	GMImp_1343	GMImp_0449	RS3L33	GMImp_0740	GMImp_0618	GMImp_0266	RS1L10	GMImp_1618
		GCA_000710135.3	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No
GCA_000710695.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001050995.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_002894775.1	IIB-1	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_000750575.1	IIB-2	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001373315.1	IIB-25	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	No	No	Yes
GCA_009832785.1	IIB-25	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001373335.1	IIB-25	No	Yes	No	No	Yes	No	No	Yes	No	Yes	No	No	Yes	Yes
GCA_000749995.1	IIB-27	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_008271875.1	IIB-3	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_000212635.2	IIB-3	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_000825925.2	IIB-3	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_000825825.2	IIB-3	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_000215325.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_000525615.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_000825805.2	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	No	No	Yes	Yes
GCA_001587135.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001587155.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001644815.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes

GenBank assembly accession	Phylotype- Sequevar	VNTR													
		GMIch_0581	RS3L19	GMIch_3461	GMIch_1457	GMImp_0487	RS1L14	GMImp_1343	GMImp_0449	RS3L33	GMImp_0740	GMImp_0618	GMImp_0266	RS1L10	GMImp_1618
		GCA_003860705.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No
GCA_003860725.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003860745.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003860765.1	IIB-4	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001373255.1	IIB-51	No	Yes	No	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001644855.1	III-23	No	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
GCA_000427195.1	III-29	No	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
GCA_006088755.1	III-x	No	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
GCA_000283475.1	IV-10	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515185.1	IV-11	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515265.1	IV-11	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_001586135.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515145.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515165.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515325.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515385.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515425.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515445.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515485.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

GenBank assembly accession	Phylotype- Sequevar	VNTR													
		GMIch_0581	RS3L19	GMIch_3461	GMIch_1457	GMImp_0487	RS1L14	GMImp_1343	GMImp_0449	RS3L33	GMImp_0740	GMImp_0618	GMImp_0266	RS1L10	GMImp_1618
GCA_003515505.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GCA_003515525.1	IV-8	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

S4 Table. Correspondence between sequevar, ST and haplotype for the comparative analysis of MLST-7 and RS1-MLVA14.

Sequevar	ST	Haplotype
I-13	STI-24	MT044
I-14	STI-2	MT067
I-15	STI-11	MT082, MT091, MT092, MT095, MT097
I-16	STI-14	MT032
I-18	STI-6	MT080
	STI-7	MT034
	STI-12	MT093, MT096
I-33	STI-9	MT021, MT025, MT085, MT094, MT098
I-34	STI-25	MT045
I-46	STI-27	MT081
I-31	STI-5	MT064
	STI-13	MT035, MT043, MT052, MT084, MT099, MT100
	ST-4	MT052
	ST-10	MT088

S5 Table. Frequency of the Malagasy MLVA type.

Haplotype	Number of strains	Frequency of haplotype per strains (%)
MT001	1	0.685
MT002	1	0.685
MT003	4	2.740
MT004	24	16.438
MT005	1	0.685
MT006	8	5.479
MT007	44	30.137
MT008	2	1.370
MT009	4	2.740
MT010	3	2.055
MT011	2	1.370
MT012	2	1.370
MT013	1	0.685
MT014	1	0.685
MT015	8	5.479
MT016	3	2.055
MT017	3	2.055
MT018	3	2.055
MT019	1	0.685
MT020	5	3.425
MT021	3	2.055
MT022	2	1.370
MT023	7	4.795
MT024	2	1.370
MT025	2	1.370
MT026	1	0.685
MT027	2	1.370
MT028	3	2.055
MT029	1	0.685
MT030	1	0.685
MT031	1	0.685
TOTAL	146	

S6 Table. Diversity intra-population indexes.

Allelic richness Madagascar	Average over loci	GMICH_0581	RS3L19	GMICH3461	GMICH_1457	GMIMP_0487	RS1L14	GMIMP_1343	GMIMP_0449	RS3L33	GMIMP_0740	GMIMP_0618	GMIMP_0266	RS1L10	GMIMP_1618
Analamanga	2.32	2.1717	1.9999	2.3889	1.8889	1	5.57	2.3889	1.8889	2.8889	2.886	1.8889	1.8889	2.6481	1
Itasy	1.54	1.5648	1.8327	1.2744	1.2744	1	3.6158	1.2744	1.2744	1.4217	1.9795	1.2744	1.2744	1.2744	1
Atsinanana	2.98	3.2179	2.816	1.56	1.56	6.6874	4.0932	2.5574	1.9993	4.2091	2.816	1.9993	2.9739	2.816	1
Private allelic richness Madagascar	Average over loci	GMICH_0581	RS3L19	GMICH3461	GMICH_1457	GMIMP_0487	RS1L14	GMIMP_1343	GMIMP_0449	RS3L33	GMIMP_0740	GMIMP_0618	GMIMP_0266	RS1L10	GMIMP_1618
Analamanga	0.49	0.6376	0.3147	0.4264	0.645	0	1.7375	1.145	0.0004	0.7853	1.0859	0.0004	0	0.1396	0
Itasy	0.05	0.2121	0.0002	0.0758	0.0305	0	0.312	0.0305	0	0.0305	0.0591	0	0	0	0
Atsinanana	1.28	1.4235	1.6687	0.9211	0.56	5.9727	0.2737	1.5574	0.0806	2.3736	1.6956	0.0806	1.0545	0.2772	0
Allelic richness Madagascar and worldwide	Average over loci	GMICH_0581	RS3L19	GMICH3461	GMICH_1457	GMIMP_0487	RS1L14	GMIMP_1343	GMIMP_0449	RS3L33	GMIMP_0740	GMIMP_0618	GMIMP_0266	RS1L10	GMIMP_1618
Madagascar and worldwide	3.24	3.7991	4.0389	2.8959	3.4052	3.4105	4.8046	3.403	1.9998	5.0135	4.2273	1.9998	2.9057	2.3097	1.1776

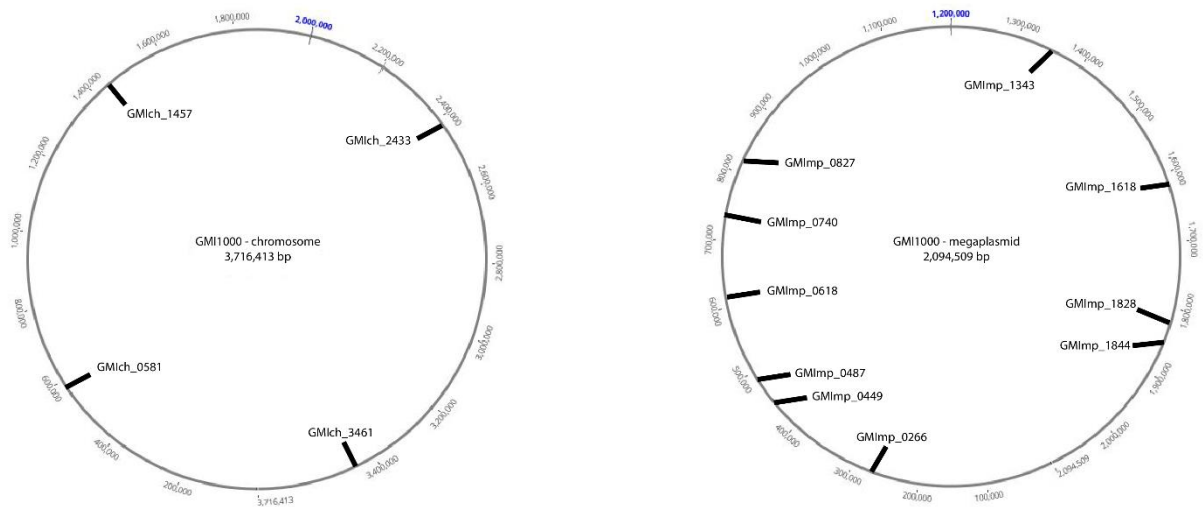
	Region of collect	Number of population	Hnb	Number of Haplotypes	The most prevalent Haplotype	Average allelic Richness	Average private Allelic richness	Mean number Of alleles	Mean number of Private alleles
Pop 1	Analamanga	28	0.318	12	MT023 (21.43%)	2.32	0.49	2.571	0.214
Pop 2	Itasy	95	0.109	15	MT007 (44.21%)	1.54	0.05	2.714	0.071
Pop 3	Atsinanana	25	0.43	12	MT003 (16%) - MT009 (16%)	2.98	1.28	2.857	0.786

S7 Table. FST and RST between populations from the three regions where data was collected and between populations from lowlands, mid-elevation and highlands.

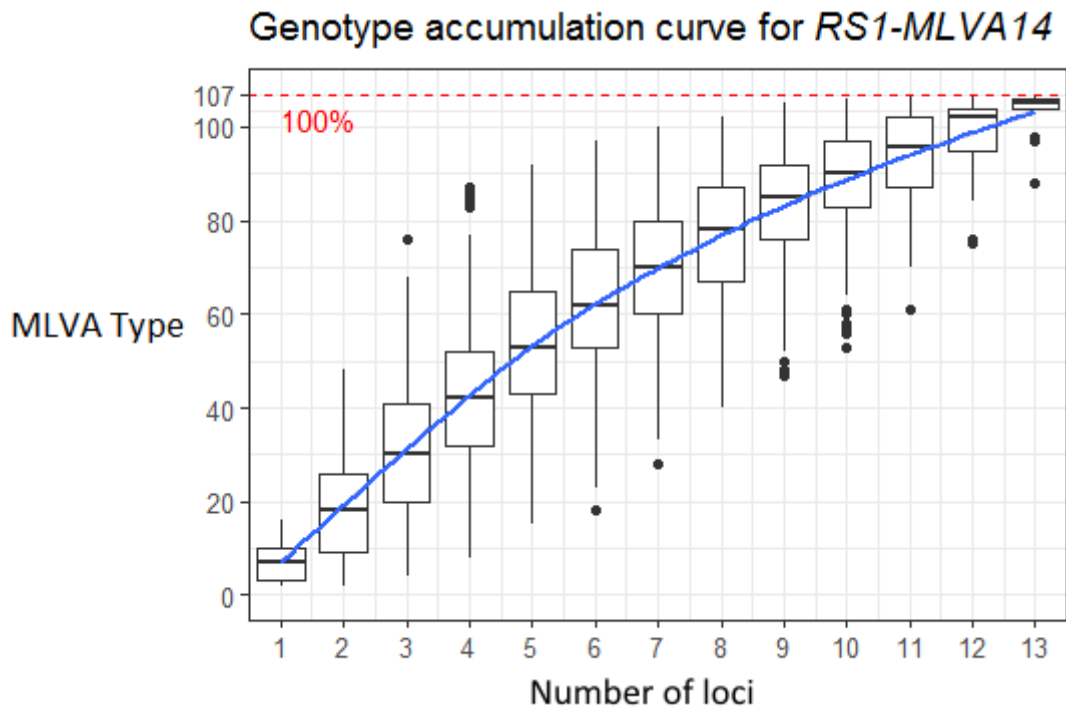
	<i>Analamanga</i>		<i>Atsinanana</i>		<i>Itasy</i>	
	FST	RST	FST	RST	FST	RST
<i>Analamanga</i>	0	0				
<i>Atsinanana</i>	0.109***	0.0897*	0	0		
<i>Itasy</i>	0.166***	0.252***	0.288***	0.381***	0	0

	<i>Lowland</i>		<i>Mid-elevation</i>		<i>Highland</i>	
	FST	RST	FST	RST	FST	RST
<i>Lowland (0 - 100m)</i>	0	0				
<i>Mid-elevation (900 - 1200m)</i>	0.095**	0.106	0	0		
<i>Highland (1200 - 1500m)</i>	0.242***	0.336***	0.079***	0.00085	0	0

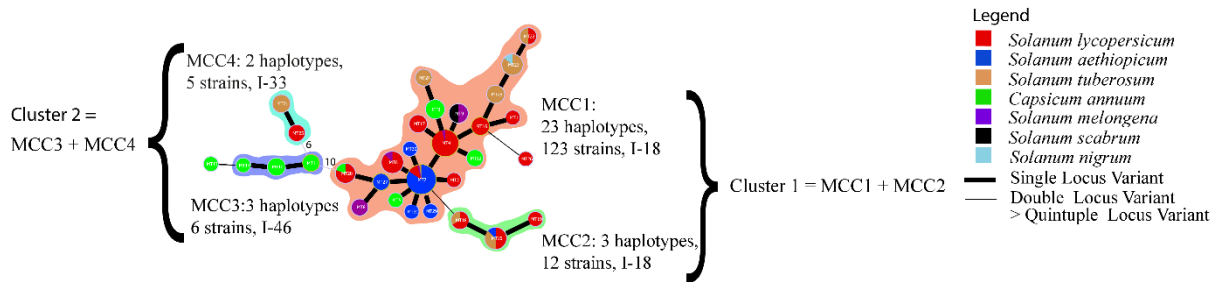
S1 Fig. Localization of the 14 VNTRs on the chromosome and megaplasmid of the reference phylotype Istrain GMI1000.



S2 Fig. Genotype accumulation curve for 291 strains of RSSC phylotype Istrains genotyped over 14 loci. The horizontal axis represents the number of loci randomly sampled without replacement up to $n-1$ loci, the vertical axis shows the number of unique MLVA types observed ($n=107$) in the data set. The level of 100% of unique MLVA types detected is indicated with adotted red line.



S3 Fig. Minimum spanning tree (MST) of the Malagasy phylotype according to the host of collect. The haplotypes were identified by using goeBURST full MST in PHYLOVIZ. Each MLVA type (MT) is displayed as a circle, the size of which is proportional to the number of isolates represented. The different colours indicate the sample host. Branch thickness indicates the number of locus differences between the neighbouring haplotypes. MCC1, MCC2, MCC3, MCC4 represent the Malagasy Clonal Complexes 1, 2, 3 and 4. A clonal complex is composed of haplotypes that differ only by one VNTR locus. I-18, I-33 and I-46 represent the phylotype - sequence of the isolates.



Annexes chapitre 2

S1 Table. Characteristics of the C1 and C2 Malagasy isolates used in the study.

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C1	304	2006	I-18	<i>Solanum lycopersicum</i>	1273.3	Antananarivo	16	1	MCC1	MT001
C1	305	2006	I-18	<i>Solanum lycopersicum</i>	1273.3	Antananarivo	16	1	MCC1	MT002
C1	306	2006	I-18	<i>Capsicum annuum</i>	74.6	Toamasina	3		MCC1	
C1	307	2006	I-18	<i>Capsicum annuum</i>	74.6	Toamasina	3		MCC1	
C1	308	2006	I-18	<i>Capsicum annuum</i>	74.6	Toamasina	3	1	MCC1	MT003
C1	309	2006	I-18	<i>Capsicum annuum</i>	74.6	Toamasina	3	1	MCC1	MT005
C1	310	2006	I-18	<i>Capsicum annuum</i>	74.6	Toamasina	3	1	MCC1	MT003
C1	311	2006	I-18	<i>Capsicum annuum</i>	74.6	Toamasina	3	1	MCC1	MT003
C1	312	2006	I-18	<i>Solanum melongena</i>	11.9	Toamasina	2	1	MCC1	MT004
C1	313	2006	I-18	<i>Solanum melongena</i>	11.9	Toamasina	2	1	MCC1	MT006
C1	314	2006	I-18	<i>Solanum melongena</i>	11.9	Toamasina	2	1	MCC1	MT007
C1	315	2006	I-18	<i>Solanum melongena</i>	11.9	Toamasina	2	1	MCC1	MT008
C1	316	2006	I-18	<i>Solanum melongena</i>	11.9	Toamasina	2	1	MCC1	MT008
C1	317	2006	I-18	<i>Solanum scabrum</i>	13.5	Toamasina	4	1	MCC1	MT009
C1	318	2006	I-18	<i>Solanum melongena</i>	13.5	Toamasina	4	1	MCC1	MT009
C1	319	2006	I-46	<i>Capsicum annuum</i>	8	Toamasina	1	2	MCC3	MT010
C1	320	2006	I-46	<i>Capsicum annuum</i>	8	Toamasina	1	2	MCC3	MT010
C1	321	2006	I-46	<i>Capsicum annuum</i>	8	Toamasina	1	2	MCC3	MT011
C1	322	2006	I-18	<i>Capsicum annuum</i>	8	Toamasina	1	1	MCC1	MT012
C1	323	2006	I-46	<i>Capsicum annuum</i>	8	Toamasina	1	2	MCC3	MT011
C1	324	2006	I-46	<i>Capsicum annuum</i>	8	Toamasina	1	2	Singleton	MT013
C1	325	2006	I-18	<i>Capsicum annuum</i>	8	Toamasina	1	1	MCC1	MT012
C1	326	2006	I-46	<i>Capsicum annuum</i>	8	Toamasina	1	2	MCC3	MT014
C1	327	2006	I-46	<i>Capsicum annuum</i>	8	Toamasina	1	2	MCC3	MT010
C1	328	2006	I-18	<i>Capsicum annuum</i>	74.6	Toamasina	3	1	MCC1	MT003
C1	329	2006	I-18	<i>Solanum scabrum</i>	13.5	Toamasina	4	1	MCC1	MT009

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C1	330	2006	I-18	<i>Solanum melongena</i>	13.5	Toamasina	4	1	MCC1	MT009
C1	3075	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3076	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3077	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3078	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3080	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT007
C1	3082	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC2	MT015
C1	3083	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3084	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT007
C1	3085	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3086	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3087	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT007
C1	3089	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3108	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3109	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT016
C1	3110	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3111	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3112	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT007
C1	3113	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT007
C1	3114	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3116	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3117	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3118	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3119	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT017
C1	3120	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT017
C1	3122	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C1	3124	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC2	MT015
C1	3125	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT017

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C1	3126	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT016
C1	3128	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Antananarivo	14	1	MCC2	MT018
C1	3129	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Antananarivo	14	1	MCC2	MT018
C1	3130	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Antananarivo	14	1	MCC2	MT015
C1	3131	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Antananarivo	14	1	MCC1	MT020
C1	3134	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Antananarivo	14	1	MCC2	MT015
C1	3136	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Antananarivo	14	1	MCC1	MT020
C1	3137	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Antananarivo	14	1	MCC1	MT020
C1	3138	2013	I-18	<i>Solanum lycopersicum</i>	1278.3	Antananarivo	14	1	MCC2	MT015
C1	3142	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT020
C1	3148	2013	I-33	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	2	MCC4	MT021
C1	3149	2013	I-33	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	2	MCC4	MT021
C1	3151	2013	I-33	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	2	MCC4	MT021
C1	3159	2013	I-18	<i>Solanum lycopersicum</i>	1267.3	Antananarivo	15	1	MCC1	MT022
C1	3162	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Antananarivo	15	1	MCC1	MT022
C1	3163	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Antananarivo	15	1	MCC2	MT015
C1	3164	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Antananarivo	15	1	MCC1	MT023
C1	3165	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Antananarivo	15	1	MCC1	MT023
C1	3168	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Antananarivo	15	1	MCC2	MT018
C1	3170	2013	I-18	<i>Solanum tuberosum</i>	1267.3	Antananarivo	15	1	MCC1	MT023
C1	3171	2013	I-18	<i>Solanum nigrum</i>	1267.3	Antananarivo	15	1	MCC1	MT023
C1	3180	2013	I-18	<i>Solanum tuberosum</i>	1476.9	Antananarivo	11	1	MCC1	MT023
C1	3208	2013	I-18	<i>Solanum tuberosum</i>	1484	Antananarivo	10	1	MCC1	MT024
C1	3210	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT004
C1	3211	2013	I-33	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	2	MCC4	MT025
C1	3212	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT016
C1	3213	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT006
C1	3214	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT004

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C1	3215	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT004
C1	3216	2013	I-33	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	2	MCC4	MT025
C1	3217	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT004
C1	3218	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT004
C1	3219	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT006
C1	3220	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT006
C1	3221	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT006
C1	3222	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT006
C1	3223	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	Singleton	MT026
C1	3224	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT006
C1	3226	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT004
C1	3228	2013	I-18	<i>Solanum lycopersicum</i>	1176.6	Antananarivo	5	1	MCC1	MT006
C1	3243	2013	I-18	<i>Capsicum annuum</i>	1190.3	Antananarivo	6	1	MCC1	MT020
C1	3245	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3246	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3247	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3256	2013	I-18	<i>Solanum tuberosum</i>	1461.5	Antananarivo	9	1	MCC2	MT015
C1	3257	2013	I-18	<i>Solanum tuberosum</i>	1461.5	Antananarivo	9	1	MCC2	MT015
C1	3264	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3265	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3266	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT027
C1	3267	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3268	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3269	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3270	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3271	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3272	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3273	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C1	3274	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3275	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3276	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3277	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3278	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3279	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3280	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3281	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3282	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3283	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3284	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3285	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC2	MT015
C1	3286	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	1	MCC1	MT028
C1	3287	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	1	MCC1	MT007
C1	3288	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	1	MCC1	MT023
C1	3289	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	1	MCC1	MT024
C1	3290	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	1	MCC1	MT028
C1	3291	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	1	MCC1	MT023
C1	3292	2013	I-18	<i>Solanum tuberosum</i>	1275.7	Antananarivo	13	1	MCC1	MT028
C1	3457	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT029
C1	3458	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3459	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT030
C1	3460	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3461	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3462	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3463	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3464	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT031
C1	3465	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C1	3466	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT027
C1	3467	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3468	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3469	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3470	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3471	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3472	2013	I-18	<i>Solanum aethiopicum</i>	1220	Antananarivo	8	1	MCC1	MT007
C1	3480	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3484	2013	I-18	<i>Solanum aethiopicum</i>	1224	Antananarivo	7	1	MCC1	MT007
C1	3507	2013	I-18	<i>Solanum lycopersicum</i>	1348.7	Antananarivo	12	1	MCC1	MT004
C2	5917	2018	I-46	-	5.2	Toamasina	27	2	MCC3	MT014
C2	5918	2018	I-46	<i>Capsicum annuum</i>	10.7	Toamasina	28	2	Singleton	MT327
C2	5919	2018	I-18	<i>Solanaceae</i>	5.2	Toamasina	27	1	Singleton	MT326
C2	5920	2018	I-18	<i>Solanaceae</i>	5.2	Toamasina	27	1	Singleton	MT346
C2	5921	2018	I-18	-	?					
C2	5922	2018	I-18	<i>Solanaceae</i>	5.2	Toamasina				
C2	5923	2018	I-18	<i>Solanum melongena</i>	5.2	Toamasina	27	1	Singleton	MT343
C2	5924	2018	I-46	<i>Helianthus</i>	5.2	Toamasina	27	2	Singleton	MT353
C2	5925	2018	I-18	<i>Cucurbita pepo</i>	18.5	Toamasina	18			
C2	5926	2018	I-46	<i>Capsicum annuum</i>	5.2	Toamasina	27	2	Singleton	MT354
C2	5927	2018	I-18	<i>Solanaceae</i>	5.2	Toamasina	27			
C2	5928	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18			
C2	5929	2018	I-18	<i>Colocasia esculenta</i>	6.7	Toamasina	21			
C2	5930	2018	I-46	-	14.3	Toamasina	20			
C2	5931	2018	I-46	<i>Capsicum annuum</i>	10.7	Toamasina	28	2	Singleton	MT345
C2	5932	2018	I-18	<i>Colocasia esculenta</i>	14.2	Toamasina	19	1	Singleton	MT325
C2	5933	2018	I-18	<i>Citrullus lanatus</i>	18.5	Toamasina				
C2	5934	2018	I-46	<i>Fabaceae</i>	5.2	Toamasina	27	2	MCC3	MT014

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	5935	2018	I-46	<i>Capsicum annuum</i>	5.2	Toamasina	27	2	MCC3	MT014
C2	5936	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18			
C2	5937	2018	I-18	<i>Solanum melongena</i>	5.2	Toamasina	27	1	Singleton	MT254
C2	5939	2018	I-18	<i>Solanaceae</i>	14.2	Toamasina	19			
C2	5940	2018	I-18	<i>Solanum melongena</i>	5.2	Toamasina	27	1	Singleton	MT254
C2	5941	2018	I-46	<i>Solanum melongena</i>	5.2	Toamasina	27	2	MCC3	MT014
C2	5942	2018	I-46	<i>Musa</i>	5.2	Toamasina	27			
C2	5943	2018	I-18	<i>Solanaceae</i>	5.2	Toamasina	27	1	Singleton	MT326
C2	5944	2018	I-46	-	-			2	Singleton	MT327
C2	5945	2018	I-46	<i>Helianthus</i>	5.2	Toamasina	27	2	MCC3	MT014
C2	5946	2018	I-46	-	14.2	Toamasina	19	2	MCC3	MT328
C2	5947	2018	I-46	<i>Capsicum annuum</i>	10.7	Toamasina	28	2	MCC3	MT014
C2	5948	2018	I-46	-	5.2	Toamasina	27			
C2	5949	2018	I-46	<i>Solanaceae</i>	5.2	Toamasina	27	2	MCC3	MT014
C2	5950	2018	I-46	<i>Capsicum annuum</i>	10.7	Toamasina				
C2	5951	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18	1	Singleton	MT341
C2	5952	2018	I-46	<i>Solanum melongena</i>	10.7	Toamasina	28	2	Singleton	MT355
C2	5953	2018	I-46	<i>Solanum melongena</i>	18.5	Toamasina	18	2	MCC9	MT356
C2	5954	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina				
C2	5955	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18			
C2	5956	2018	I-46	<i>Solanaceae</i>	14.2	Toamasina	19	2	MCC3	MT328
C2	5957	2018	I-46	<i>Solanum americanum</i>	18.5	Toamasina	18	2	MCC3	MT329
C2	5958	2018	I-46	-	14.3	Toamasina	20	2	MCC10	MT349
C2	5959	2018	I-18	<i>Phaseolus vulgaris</i>	1405.9	Antananarivo	33			
C2	5960	2018	I-18	<i>Phaseolus vulgaris</i>	1405.9	Antananarivo	33	1	MCC1	MT334
C2	5961	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC1	MT334
C2	5962	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC1	MT334
C2	5963	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23			

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	5964	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina				
C2	5965	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC1	MT334
C2	5966	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina				
C2	5967	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina				
C2	5968	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC1	MT028
C2	5969	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC1	MT334
C2	5970	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC1	MT334
C2	5971	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23			
C2	5972	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23			
C2	5973	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC1	MT334
C2	5974	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23			
C2	5975	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina				
C2	5976	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23			
C2	5977	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC1	MT028
C2	5978	2018	I-18	<i>Capsicum annuum</i>	14	Toamasina	23	1	MCC8	MT347
C2	5979	2018	I-18	<i>Mimosoideae</i>	14	Toamasina	23	1	MCC8	MT348
C2	5980	2018	I-18	<i>Mimosoideae</i>	14	Toamasina	23	1	MCC8	MT347
C2	5981	2018	I-46	<i>Capsicum annuum</i>	79.9	Toamasina	14	2	MCC4	MT358
C2	5982	2018	I-46	<i>Solanum macrocarpon</i>	79.9	Toamasina	14	2	Singleton	MT362
C2	5983	2018	I-46	<i>Solanum macrocarpon</i>	79.9	Toamasina	14	2	MCC3	MT330
C2	5984	2018	I-46	<i>Solanum mauritianum</i>	177.8	Toamasina	15	2	MCC3	MT330
C2	5985	2018	I-46	-	202.8	Toamasina	17	2	MCC4	MT359
C2	5986	2018	I-18	<i>Solanum melongena</i>	202.8	Toamasina	17	1	MCC1	MT027
C2	5987	2018	I-46	<i>Solanum melongena</i>	202.8	Toamasina	17	1	MCC6	MT351
C2	5988	2018	I-46	<i>Solanum melongena</i>	202.8	Toamasina	17	2	MCC4	MT360
C2	5989	2018	I-18	<i>Asteraceae</i>	192	Toamasina	16	1	MCC6	MT352
C2	5990	2018	I-46	-	14.3	Toamasina	20	2	MCC10	MT350
C2	5991	2018	I-46	-	14.3	Toamasina	20			

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	5992	2018	I-46	<i>Solanaceae</i>	14.2	Toamasina	19			
C2	5993	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina				
C2	5994	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18			
C2	5995	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina				
C2	5996	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18			
C2	5997	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18			
C2	5998	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18			
C2	5999	2018	I-18	<i>Solanum americanum</i>	18.5	Toamasina	18			
C2	6000	2018	I-46	<i>Solanum americanum</i>	18.5	Toamasina	18	2	MCC9	MT357
C2	6001	2018	I-46	-	18.5	Toamasina	18	2	Singleton	MT361
C2	6249	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT239
C2	6250	2019	I-18	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT235
C2	6251	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6252	2019	I-18	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6253	2019	I-18	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT237
C2	6254	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6255	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6256	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6257	2019	I-18	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6258	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6259	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6260	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6261	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6262	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6263	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6264	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6265	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT236
C2	6266	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT238

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6267	2019	I-18	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT238
C2	6268	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10			
C2	6270	2019	I	<i>Solanum melongena</i>	12	Mahajanga	10	1	MCC2	MT238
C2	6271	2019	I	<i>Solanum melongena</i>	15	Mahajanga	9	1	MCC2	MT239
C2	6272	2019	I	<i>Solanum nigrum</i>	15	Mahajanga	9	1	MCC2	MT249
C2	6273	2019	I-18	<i>Solanum nigrum</i>	15	Mahajanga	9	1	MCC2	MT239
C2	6274	2019	I-18	<i>Solanum nigrum</i>	12	Mahajanga	8	1	Singleton	MT240
C2	6275	2019	I-18	<i>Solanum nigrum</i>	12	Mahajanga	8	1	MCC7	MT241
C2	6276	2019	I	<i>Solanum nigrum</i>	12	Mahajanga	8	1	MCC7	MT241
C2	6277	2019	I	<i>Solanum nigrum</i>	12	Mahajanga	8	1	Singleton	MT340
C2	6278	2019	I	<i>Solanum nigrum</i>	12	Mahajanga	8	1	Singleton	MT240
C2	6279	2019	I-18	<i>Capsicum annuum</i>	12	Mahajanga	8	1	MCC7	MT242
C2	6280	2019	I-18	<i>Solanum aethiopicum</i>	44	Mahajanga	11	1	MCC1	MT243
C2	6281	2019	I	<i>Solanum aethiopicum</i>	44	Mahajanga	11	1	MCC1	MT004
C2	6282	2019	I-18	<i>Solanum aethiopicum</i>	44	Mahajanga	11	1	MCC1	MT243
C2	6283	2019	I-18	<i>Solanum aethiopicum</i>	44	Mahajanga	11	1	MCC1	MT243
C2	6284	2019	I	<i>Solanum aethiopicum</i>	44	Mahajanga	11	1	MCC1	MT243
C2	6285	2019	I-31	<i>Solanum aethiopicum</i>	44	Mahajanga	11	2	MCC11	MT244
C2	6286	2019	I-18	<i>Solanum aethiopicum</i>	44	Mahajanga	11	1	MCC1	MT245
C2	6287	2019	I	<i>Solanum aethiopicum</i>	44	Mahajanga	11	1	MCC1	MT245
C2	6288	2019	I	<i>Solanum aethiopicum</i>	44	Mahajanga	11	1	MCC1	MT245
C2	6289	2019	I	<i>Nicotiana tabacum</i>	44	Mahajanga	11	1	MCC1	MT004
C2	6290	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT004
C2	6291	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT246
C2	6292	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT004
C2	6293	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT004
C2	6294	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6295	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT245

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6296	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT007
C2	6297	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6298	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6299	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT245
C2	6300	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6301	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6302	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6303	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC2	MT247
C2	6304	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC2	MT247
C2	6305	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT004
C2	6306	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6307	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6308	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT004
C2	6309	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC2	MT251
C2	6310	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6311	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC2	MT248
C2	6312	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC2	MT251
C2	6313	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6314	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC2	MT248
C2	6315	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6316	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6317	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6318	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6319	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6320	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6321	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT003
C2	6322	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT007
C2	6323	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6			

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6324	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT004
C2	6325	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT004
C2	6326	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC2	MT249
C2	6327	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT004
C2	6328	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	6	1	MCC1	MT007
C2	6329	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT249
C2	6330	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC1	MT004
C2	6331	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	Singleton	MT254
C2	6332	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC1	MT004
C2	6333	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC1	MT003
C2	6334	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC1	MT004
C2	6335	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT249
C2	6336	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT249
C2	6337	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC1	MT252
C2	6338	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	Singleton	MT254
C2	6339	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT250
C2	6340	2019	I-31	<i>Solanum lycopersicum</i>	48	Mahajanga	5	2	MCC11	MT363
C2	6341	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT251
C2	6342	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT249
C2	6343	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT247
C2	6344	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT247
C2	6345	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC1	MT003
C2	6346	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT249
C2	6347	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC2	MT249
C2	6348	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC1	MT252
C2	6349	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	Singleton	MT254
C2	6350	2019	I	<i>Solanum lycopersicum</i>	48	Mahajanga				
C2	6351	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	MCC1	MT253

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6352	2019	I-18	<i>Solanum lycopersicum</i>	48	Mahajanga	5	1	Singleton	MT254
C2	6353	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	Singleton	MT331
C2	6354	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT332
C2	6355	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT249
C2	6356	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT249
C2	6357	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT249
C2	6358	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT249
C2	6359	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT251
C2	6360	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT251
C2	6361	2019	I-18	<i>Physalis</i>	44	Mahajanga	4	1	MCC2	MT255
C2	6362	2019	I	<i>Physalis</i>	44	Mahajanga	4	1	MCC2	MT247
C2	6363	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT003
C2	6364	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6365	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6366	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT249
C2	6367	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT029
C2	6368	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	Singleton	MT342
C2	6369	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4			
C2	6370	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga				
C2	6371	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT251
C2	6372	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6373	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6374	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6375	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6376	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6377	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4			
C2	6378	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6379	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT027

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6380	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT257
C2	6381	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT249
C2	6382	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT249
C2	6383	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6384	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT004
C2	6385	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6386	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6387	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6388	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6389	2019	I-18	<i>Physalis</i>	44	Mahajanga	4	1	MCC2	MT249
C2	6390	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT256
C2	6391	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC2	MT257
C2	6392	2019	I	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6393	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT007
C2	6394	2019	I-18	<i>Solanum lycopersicum</i>	44	Mahajanga	4	1	MCC1	MT004
C2	6395	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC2	MT267
C2	6396	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6397	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC2	MT267
C2	6398	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	Singleton	MT344
C2	6399	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC2	MT258
C2	6400	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC2	MT251
C2	6401	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC2	MT267
C2	6402	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	Singleton	MT333
C2	6403	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC2	MT267
C2	6404	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6405	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6406	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC2	MT259
C2	6407	2019	I-18	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC2	MT259

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6408	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC2	MT259
C2	6409	2019	I-18	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC2	MT260
C2	6410	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT020
C2	6411	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT020
C2	6412	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT020
C2	6413	2019	I	<i>Capsicum annuum</i>	49	Mahajanga	7	1	MCC1	MT020
C2	6414	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6415	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT007
C2	6416	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6417	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6418	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6419	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6420	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC2	MT259
C2	6421	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6422	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6423	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6424	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT007
C2	6425	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	Singleton	MT261
C2	6426	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6427	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6428	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT007
C2	6429	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	Singleton	MT261
C2	6430	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	Singleton	MT261
C2	6431	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT007
C2	6432	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6433	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6434	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT007
C2	6435	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC2	MT259

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6436	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC2	MT259
C2	6437	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6438	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC2	MT259
C2	6439	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6440	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	Singleton	MT261
C2	6441	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT004
C2	6442	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT007
C2	6443	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3	1	MCC1	MT007
C2	6444	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	3			
C2	6445	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6446	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6447	2019	I-18	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	Singleton	MT262
C2	6448	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6449	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6450	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC1	MT004
C2	6451	2019	I	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC1	MT004
C2	6452	2019	I-18	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6453	2019	I-18	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6454	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6455	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC2	MT249
C2	6456	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT017
C2	6457	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6458	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6459	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6460	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6461	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6462	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6464	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6465	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6466	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	Singleton	MT263
C2	6467	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6468	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6469	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6470	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6471	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6472	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6473	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6474	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT264
C2	6475	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6476	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2			
C2	6477	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6478	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6479	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6480	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT245
C2	6481	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6482	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6483	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	Singleton	MT265
C2	6484	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6485	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6486	2019	I-18	<i>Solanum lycopersicum</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6487	2019	I	<i>Solanum lycopersicum</i>	21	Mahajanga	2			
C2	6488	2019	I	<i>Physalis</i>	21	Mahajanga	2	1	MCC1	MT004
C2	6489	2019	I-18	<i>Solanum melongena</i>	8	Mahajanga	1	1	MCC2	MT266
C2	6490	2019	I	<i>Solanum melongena</i>	8	Mahajanga	1	1	MCC2	MT249
C2	6491	2019	I	<i>Solanum melongena</i>	8	Mahajanga	1	1	MCC2	MT249
C2	6492	2019	I	<i>Solanum melongena</i>	8	Mahajanga	1	1	MCC2	MT249

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6493	2019	I	<i>Solanum melongena</i>	8	Mahajanga	1	1	MCC2	MT266
C2	6494	2019	I	<i>Solanum melongena</i>	8	Mahajanga	1	1	MCC2	MT266
C2	6495	2019	I	<i>Solanum melongena</i>	8	Mahajanga	1	1	MCC2	MT266
C2	6496	2019	I	<i>Solanum melongena</i>	8	Mahajanga	1	1	MCC2	MT249
C2	6497	2019	I	<i>Brassica</i>	8	Mahajanga	1			
C2	6498	2019	I-18	<i>Solanum aethiopicum</i>	8	Mahajanga	1	1	MCC2	MT249
C2	6499	2019	I	<i>Solanum aethiopicum</i>	8	Mahajanga	1	1	MCC2	MT249
C2	6500	2019	I-18	<i>Physalis</i>	44	Mahajanga	4	1	MCC1	MT252
C2	6508	2019	I-18	<i>Solanum aethiopicum</i>	49	Mahajanga	7	1	MCC2	MT267
C2	6509	2019	I	<i>Solanum lycopersicum</i>	49	Mahajanga	7	1	MCC1	MT007
C2	6531	2019	I	<i>Solanum lycopersicum</i>	765	Toamasina	12	1	MCC1	MT007
C2	6532	2019	I	<i>Solanum lycopersicum</i>	765	Toamasina	12			
C2	6534	2019	I	<i>Capsicum annuum</i>	787	Toamasina	13	1	MCC1	MT004
C2	6537	2019	I	<i>Capsicum annuum</i>	787	Toamasina	13	1	MCC1	MT004
C2	6540	2019	I	<i>Capsicum annuum</i>	787	Toamasina				
C2	6541	2019	I	<i>Capsicum annuum</i>	787	Toamasina				
C2	6543	2019	I-18	<i>Capsicum annuum</i>	787	Toamasina				
C2	6002	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6003	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6004	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6005	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6006	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6007	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6008	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6009	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6010	2018	IIB-1	<i>Solanum tuberosum</i>	1276	Antananarivo				
C2	6503	2019	II	<i>Solanum tuberosum</i>	1503	Antananarivo				
C2	6504	2019	II	<i>Solanum tuberosum</i>	1503	Antananarivo				

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6505	2019	II	<i>Solanum tuberosum</i>	1503	Antananarivo				
C2	6506	2019	II	<i>Solanum tuberosum</i>	1503	Antananarivo				
C2	6507	2019	IIB-1	<i>Solanum tuberosum</i>	1503	Antananarivo				
C2	6510	2019	IIB-1	<i>Solanum tuberosum</i>	913	Antananarivo				
C2	6511	2019	II	<i>Solanum tuberosum</i>	913	Antananarivo				
C2	6512	2019	IIB-1	<i>Solanum tuberosum</i>	792	Toamasina				
C2	6513	2019	IIB-1	<i>Solanum tuberosum</i>	902	Antananarivo				
C2	6514	2019	IIB-1	<i>Solanum tuberosum</i>	902	Antananarivo				
C2	6515	2019	IIB-1	<i>Solanum tuberosum</i>	902	Antananarivo				
C2	6516	2019	II	<i>Solanum tuberosum</i>	902	Antananarivo				
C2	6517	2019	IIB-1	<i>Solanum tuberosum</i>	902	Antananarivo				
C2	6518	2019	II	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6519	2019	II	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6520	2019	II	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6521	2019	II	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6522	2019	IIB-1	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6523	2019	IIB-1	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6524	2019	II	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6525	2019	IIB-1	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6526	2019	IIB-1	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6527	2019	II	<i>Solanum tuberosum</i>	764	Toamasina				
C2	6528	2019	IIB-1	<i>Capsicum annuum</i>	766	Toamasina				
C2	6529	2019	II	<i>Solanum lycopersicum</i>	765	Toamasina				
C2	6530	2019	II	<i>Solanum lycopersicum</i>	765	Toamasina				
C2	6533	2019	III-60	<i>Capsicum annuum</i>	787	Toamasina				
C2	6535	2019	II	<i>Capsicum annuum</i>	787	Toamasina				
C2	6536	2019	IIB-1	<i>Capsicum annuum</i>	787	Toamasina				
C2	6538	2019	III-60	<i>Capsicum annuum</i>	787	Toamasina				

Collection name	RUN.ID	Year	Phylotype - Sequevar	Host	Altitude	Province	Plot number	Cluster	Malagasy Clonal Complex	Haplotype
C2	6539	2019	III-60	<i>Capsicum annuum</i>	787	Toamasina				
C2	6542	2019	III-60	<i>Capsicum annuum</i>	787	Toamasina				
C2	6545	2019	II	<i>Solanum tuberosum</i>	783	Toamasina				
C2	6546	2019	II	<i>Solanum tuberosum</i>	783	Toamasina				
C2	6547	2019	II	<i>Solanum tuberosum</i>	783	Toamasina				
C2	6548	2019	IIB-1	<i>Solanum tuberosum</i>	783	Toamasina				
C2	6549	2019	IIB-1	<i>Solanum tuberosum</i>	783	Toamasina				
C2	6550	2019	IIB-1	<i>Solanum tuberosum</i>	783	Toamasina				
C2	6551	2019	II	<i>Solanum tuberosum</i>	786	Toamasina				
C2	6552	2019	IIB-1	<i>Solanum tuberosum</i>	786	Toamasina				
C2	6553	2019	II	<i>Solanum tuberosum</i>	786	Toamasina				
C2	6554	2019	II	<i>Solanum tuberosum</i>	792	Toamasina				
C2	6556	2019	IIB-1	<i>Solanum tuberosum</i>	792	Toamasina				
C2	6557	2019	IIB-1	<i>Solanum tuberosum</i>	792	Toamasina				
C2	6558	2019	II	<i>Solanum tuberosum</i>	792	Toamasina				
C2	6559	2019	IIB-1	<i>Solanum tuberosum</i>	792	Toamasina				
C2	6560	2019	IIB-1	<i>Solanum tuberosum</i>	792	Toamasina				
C2	6562	2019	IIB-1	<i>Solanum tuberosum</i>	792	Toamasina				
C2	6565	2019	III-19	<i>Phaseolus vulgaris</i>	1240	Antananarivo				

S2 Table. Characteristics of the C3 isolates collection used in the study.

RUN.ID	World	Country	Year	Phylotype - sequevar	Host	Plant families	Haplotype
RUN33	Asia	Japan	1965	I-16	<i>Nicotinia tabacum</i>	Solanaceae	MT032
RUN38	West Indies	Guadeloupe	1983	I-18	<i>Solanum melongena</i>	Solanaceae	MT033
RUN44	Australia	Australia	1966	I-16	<i>Zingiber officinale</i>	Zingiberaceae	MT032
RUN54	America (South)	French Guiana	1978	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT034
RUN78	Asia	Malaysia	na	I-48	<i>Solanum lycopersicum</i>	Solanaceae	MT036
RUN87	Asia	Taiwan, Province Of China	1997	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT213
RUN94	Asia	Indonesia	na	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT034
RUN95	Asia	Indonesia	na	I-15	<i>Solanum lycopersicum</i>	Solanaceae	MT037
RUN100	Africa	Kenya	1974	I-31	<i>Capsicum annuum</i>	Solanaceae	MT168
RUN102	Africa	South Africa	1986	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT004
RUN104	Africa	South Africa	1986	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN105	Africa	South Africa	1986	I-31	<i>Capsicum annuum</i>	Solanaceae	MT035
RUN106	Africa	South Africa	1986	I-31	<i>Nicotiana tabacum</i>	Solanaceae	MT168
RUN107	Africa	South Africa	1986	I-31	<i>Nicotinia tabacum</i>	Solanaceae	MT035
RUN108	Africa	South Africa	1986	I-31	<i>Nicotiana tabacum</i>	Solanaceae	MT035
RUN114	America (South)	Brazil	1999	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT038
RUN116	America	Brazil	1998	I-18	<i>Capsicum annuum</i>	Solanaceae	MT191
RUN119	Africa	Cameroon	2005	I-13	<i>Solanum lycopersicum</i>	Solanaceae	MT039
RUN157	Asia	Taiwan	1988	I-15	<i>Solanum lycopersicum</i>	Solanaceae	MT040
RUN194	Africa	Cameroon	2005	I-13	<i>Amaranthus</i>	Amaranthaceae	MT041
RUN205	Africa	Cameroon	2005	I-44	<i>Solanum lycopersicum</i>	Solanaceae	MT042
RUN215	Africa	Cameroon	2005	I-13	<i>Solanum scabrum</i>	Solanaceae	MT044
RUN257	Asia	Taiwan	1996	I-34	<i>Solanum lycopersicum</i>	Solanaceae	MT045

RUN.ID	World	Country	Year	Phylotype - sequevar	Host	Plant families	Haplotype
RUN300	America (South)	Brazil	2003	I-17	<i>Begonia semperflorens</i>	Begoniaceae	MT046
RUN353	West Indies	Martinique	2002	I-18	<i>Capsicum annuum</i>	Solanaceae	MT047
RUN471	Indian Ocean	Reunion Island	1993	I-31	<i>Pelargonium asperum</i>	Geraniaceae	MT035
RUN486	America (South)	Peru	1966	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT048
RUN487	Indian Ocean	Seychelles	na	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN504	Indian Ocean	Seychelles	na	I-31	<i>Capsicum annuum</i>	Solanaceae	MT035
RUN528	West Indies	Trinidad	2003	I-14	<i>Solanum lycopersicum</i>	Solanaceae	MT049
RUN577	Asia	Malaysia	1998	I-18	<i>Arachis hypogaea</i>	Fabaceae	MT050
RUN578	America (Central)	Costa Rica	1972	I-14	<i>Capsicum annuum</i>	Solanaceae	MT051
RUN608	Indian Ocean	Reunion Island	na	I-31	<i>Solanum tuberosum</i>	Solanaceae	MT052
RUN616	Asia	China	2004	I-15	<i>Solanum lycopersicum</i>	Solanaceae	MT053
RUN617	Asia	China	2004	I-34	<i>Solanum melongena</i>	Solanaceae	MT054
RUN771	West Indies	Martinique	2002	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT055
RUN796	West Indies	Martinique	2002	I-18	<i>Petroselinum crispum</i>	Apiaceae	MT056
RUN821	West Indies	Martinique	2003	I-18	<i>Cucurbita pepo</i>	Cucurbitaceae	MT057
RUN1068	West Indies	Martinique	2003	I-18	<i>Physalis angulata</i>	Solanaceae	MT058
RUN1074	West Indies	Guadeloupe	2006	I-18	<i>Musa acuminata x balbisiana</i> AAA	Musaceae	MT059
RUN1110	America (Central)	Honduras	1997	I-14	<i>Solanum lycopersicum</i>	Solanaceae	MT060
RUN1161	West Indies	Martinique	2003	I-18	<i>Anthurium andreanum</i>	Araceae	MT061
RUN1401	West Indies	Guadeloupe	2007	I-18	<i>Capsicum annuum</i>	Solanaceae	MT062
RUN1404	West Indies	Guadeloupe	2007	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT063
RUN1431	Indian Ocean	Mayotte	2006	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT064
RUN1432	Indian Ocean	Mayotte	2006	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN1439	Indian Ocean	Mayotte	2006	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035

RUN.ID	World	Country	Year	Phylotype - sequevar	Host	Plant families	Haplotype
RUN1495	West Indies	Trinidad	2008	I-14	<i>Solanum lycopersicum</i>	Solanaceae	MT065
RUN1523	America (South)	Guyana	2008	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT066
RUN1528	America (Central)	Guatemala	2004	I-14	<i>Solanum lycopersicum</i>	Solanaceae	MT067
RUN1537	Africa	Ivory Coast	2010	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN1743	Africa	Ivory Coast	2010	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN1746	Africa	Ivory Coast	2010	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN1771	Africa	Ivory Coast	2010	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT068
RUN1798	Africa	Ivory Coast	2010	I-13	<i>Solanum melongena</i>	Solanaceae	MT069
RUN1801	Africa	Ivory Coast	2010	I-31	<i>Capsicum annuum</i>	Solanaceae	MT035
RUN1818	Africa	Ivory Coast	2010	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN1829	Africa	Ivory Coast	2010	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN1831	Africa	Ivory Coast	2010	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN1854	Africa	Ivory Coast	2011	I-14	<i>Solanum lycopersicum</i>	Solanaceae	MT070
RUN1869	Africa	Ivory Coast	2011	I-44	<i>Solanum lycopersicum</i>	Solanaceae	MT071
RUN1874	Africa	Ivory Coast	2011	I-46	<i>Solanum lycopersicum</i>	Solanaceae	MT072
RUN1890	Africa	Ivory Coast	2011	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN1897	Africa	Ivory Coast	2011	I-46	<i>Solanum lycopersicum</i>	Solanaceae	MT073
RUN1906	Africa	Ivory Coast	2011	I-46	<i>Solanum lycopersicum</i>	Solanaceae	MT074
RUN1912	Africa	Ivory Coast	2011	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN1914	Africa	Ivory Coast	2011	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN1930	Indian Ocean	Mayotte	2011	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043
RUN1948	America (South)	French Guiana	2011	I-17	<i>Capsicum annuum</i>	Solanaceae	MT075
RUN1952	America (South)	French Guiana	2011	I-17	<i>Capsicum annuum</i>	Solanaceae	MT076
RUN1958	America (South)	French Guiana	2011	I-18	<i>Capsicum annuum</i>	Solanaceae	MT077

RUN.ID	World	Country	Year	Phylotype - sequevar	Host	Plant families	Haplotype
RUN1972	America (South)	French Guiana	2011	I-14	<i>Capsicum annuum</i>	Solanaceae	MT078
RUN1975	America (South)	French Guiana	2011	I-18	<i>Solanum aethiopicum</i>	Solanaceae	MT079
RUN2083	Indian Ocean	Mayotte	2012	I-18	<i>Capsicum frutescens</i>	Solanaceae	MT080
RUN2127	Indian Ocean	Mayotte	2012	I-46	<i>Capsicum annuum</i>	Solanaceae	MT081
RUN2144	Indian Ocean	Mayotte	2012	I-15	<i>Solanum melongena</i>	Solanaceae	MT082
RUN2177	America (South)	French Guiana	2012	I-47	<i>Solanum lycopersicum</i>	Solanaceae	MT083
RUN3489	Indian Ocean	Seychelles	2005	I-31	<i>Solanum melongena</i>	Solanaceae	MT084
RUN3490	Indian Ocean	Seychelles	1998	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN3491	Indian Ocean	Comoros	2005	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043
RUN3492	Indian Ocean	Comoros	2005	I-31	<i>Phaseolus vulgaris</i>	Fabaceae	MT043
RUN3493	Indian Ocean	Comoros	2005	I-31	<i>Solanum tuberosum</i>	Solanaceae	MT043
RUN3494	Indian Ocean	Comoros	2005	I-31	<i>Solanum tuberosum</i>	Solanaceae	MT043
RUN3495	Indian Ocean	Comoros	2005	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043
RUN3501	Indian Ocean	Mauritius	1957	I-33	<i>Vicia faba</i>	Fabaceae	MT025
RUN3505	Indian Ocean	Mauritius	1957	I-33	<i>Brassica oleracea</i>	Brassicaceae	MT085
RUN3510	Indian Ocean	Mauritius	1957	I-31	<i>Casuarina equisetifolia</i>	Casuarinaceae	MT035
RUN3556	Asia	Philippines	2010	I-17	<i>Capsicum annuum</i>	Solanaceae	MT086
RUN3558	Asia	Philippines	2008	I-17	<i>Cucurbita maxima</i>	Cucurbitaceae	MT087
RUN3559	Asia	Philippines	2003	I-15	<i>Solanum lycopersicum</i>	Solanaceae	MT037
RUN3560	Asia	Thailand	2009	I-31	<i>Momordica charantia</i>	Cucurbitaceae	MT088
RUN3563	Asia	Thailand	2007	I-31	<i>Solanum melongena</i>	Solanaceae	MT088
RUN3564	Asia	Thailand	2011	I-31	<i>Tagetes sp.</i>	Asteraceae	MT088
RUN3567	Asia	Thailand	2010	I-47	<i>Capsicum annuum</i>	Solanaceae	MT089
RUN3573	Indian Ocean	Reunion Island	2014	I-33	<i>Solanum lycopersicum</i>	Solanaceae	MT085
RUN3633	Africa	Bénin	2014	I-31	<i>Solanum macrocarpon</i>	Solanaceae	MT052
RUN3657	Africa	Bénin	2014	I-31	<i>Solanum macrocarpon</i>	Solanaceae	MT035

RUN.ID	World	Country	Year	Phylotype - sequevar	Host	Plant families	Haplotype
RUN3663	America (Central)	El Salvador	2014	I-14	<i>Solanum lycopersicum</i>	Solanaceae	MT090
RUN3711	Indian Ocean	Reunion Island	2014	I-31	<i>Pelargonium graveolens</i>	Geraniaceae	MT043
RUN3743	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN3756	Indian Ocean	Reunion Island	2014	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN3770	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN3815	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN3841	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	Geraniaceae	MT035
RUN3856	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	Geraniaceae	MT043
RUN3869	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	Geraniaceae	MT035
RUN3884	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	Geraniaceae	MT043
RUN3893	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN3928	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	Geraniaceae	MT035
RUN4076	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum nigrum</i>	Solanaceae	MT035
RUN4079	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN4082	Indian Ocean	Reunion Island	2015	I-31	<i>Pelargonium graveolens</i>	Geraniaceae	MT043
RUN4105	Indian Ocean	Reunion Island	2015	I-31	<i>Capsicum annuum</i>	Solanaceae	MT035
RUN4112	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN4196	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN4210	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN4257	Indian Ocean	Reunion Island	2015	I-31	<i>Phaseolus vulgaris</i>	Fabaceae	MT035
RUN4263	Indian Ocean	Reunion Island	2015	I-31	<i>Amaranthus viridis</i>	Amaranthaceae	MT035
RUN4292	Indian Ocean	Seychelles	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN4303	Indian Ocean	Seychelles	2015	I-31	<i>Capsicum annuum</i>	Solanaceae	MT035
RUN4308	Indian Ocean	Seychelles	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN4311	Indian Ocean	Seychelles	2015	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN4315	Indian Ocean	Seychelles	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN4325	Indian Ocean	Seychelles	2015	I-31	<i>Amaranthus viridis</i>	Amaranthaceae	MT035

RUN.ID	World	Country	Year	Phylotype - sequevar	Host	Plant families	Haplotype
RUN4326	Indian Ocean	Seychelles	2015	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN4337	Indian Ocean	Seychelles	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN4376	Indian Ocean	Mauritius	2015	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN4392	Indian Ocean	Mauritius	2015	I-15	<i>Solanum lycopersicum</i>	Solanaceae	MT091
RUN4393	Indian Ocean	Mauritius	2015	I-15	<i>Capsicum annuum</i>	Solanaceae	MT091
RUN4400	Indian Ocean	Mauritius	2015	I-33	<i>Solanum lycopersicum</i>	Solanaceae	MT025
RUN4406	Indian Ocean	Mauritius	2015	I-33	<i>Amaranthus</i>	Amaranthaceae	MT025
RUN4407	Indian Ocean	Mauritius	2015	I-15	<i>Solanum lycopersicum</i>	Solanaceae	MT092
RUN4438	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum muricatum</i>	Solanaceae	MT035
RUN4450	Indian Ocean	Reunion Island	2015	I-31	<i>Capsicum annuum</i>	Solanaceae	MT035
RUN4584	Indian Ocean	Mauritius	2015	I-18	<i>Solanum tuberosum</i>	Solanaceae	MT093
RUN4603	Indian Ocean	Mauritius	2015	I-33	<i>Solanum tuberosum</i>	Solanaceae	MT025
RUN4645	Indian Ocean	Mauritius	2015	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT093
RUN4664	Indian Ocean	Mauritius	2015	I-15	<i>Solanum tuberosum</i>	Solanaceae	MT091
RUN4680	Indian Ocean	Mauritius	2015	I-33	<i>Solanum tuberosum</i>	Solanaceae	MT094
RUN5007	Indian Ocean	Mauritius	2015	I-15	<i>Solanum tuberosum</i>	Solanaceae	MT095
RUN5135	Indian Ocean	Mauritius	2015	I-18	<i>Cyphostemma mappia</i>	Vitaceae	MT096
RUN5137	Indian Ocean	Mauritius	2015	I-15	<i>Capsicum frutescens</i>	Solanaceae	MT097
RUN5253	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN5277	Indian Ocean	Reunion Island	2015	I-31	<i>Solanum tuberosum</i>	Solanaceae	MT035
RUN5385	Indian Ocean	Mauritius	2016	I-33	<i>Anthurium</i>	Araceae	MT025
RUN5445	Indian Ocean	Rodrigues	2016	I-33	<i>Capsicum frutescens</i>	Solanaceae	MT098
RUN5453	Indian Ocean	Mauritius	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043
RUN5454	Indian Ocean	Mauritius	2016	I-33	<i>Anthurium</i>	Araceae	MT025
RUN5479	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043
RUN5480	Indian Ocean	Comoros	2016	I-31	<i>Solanum tuberosum</i>	Solanaceae	MT043
RUN5481	Indian Ocean	Comoros	2016	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN5482	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043

RUN.ID	World	Country	Year	Phylotype - sequevar	Host	Plant families	Haplotype
RUN5485	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT099
RUN5489	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT100
RUN5495	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043
RUN5508	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043
RUN5516	Indian Ocean	Comoros	2016	I-31	<i>Solanum melongena</i>	Solanaceae	MT035
RUN5521	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT035
RUN5525	Indian Ocean	Comoros	2016	I-31	<i>Solanum lycopersicum</i>	Solanaceae	MT043
RUN5536	Pacific islands	New-Caledonia	2016	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT160
RUN5564	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT101
RUN5565	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum tuberosum</i>	Solanaceae	MT102
RUN5566	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT103
RUN5567	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum lycopersicum</i>	Solanaceae	MT104
RUN5570	Pacific islands	New-Caledonia	2017	I-18	<i>Solanum tuberosum</i>	Solanaceae	MT105

S3 Table. F_{ST} and R_{ST} between populations from the three regions.

	Antananarivo		Mahajanga		Toamasina	
	Fst	Rst	Fst	Rst	Fst	Rst
Antananarivo	0					
Mahajanga	0.19***	0.24***	0			
Toamasina	0.28***	0.19***	0.19***	0.29***	0	0

Gege

S4 Table. F_{ST} and R_{ST} between populations from the three regions.

	<i>Solanum lycopersicum</i>		<i>Capsicum annuum</i>		<i>Solanum melongena</i>		<i>Solanum tuberosum</i>		<i>Solanum aethiopicum</i>	
	Fst	Rst	Fst	Rst	Fst	Rst	Fst	Rst	Fst	Rst
<i>Solanum lycopersicum</i>	0	0								
<i>Capsicum annuum</i>	0.15966***	0.25717***	0	0						
<i>Solanum melongena</i>	0.32285***	0.25503***	0.2589***	0.1214***	0	0				
<i>Solanum tuberosum</i>	0.12639***	0.29444***	0.0834***	0.00687*	0.33802***	0.27941***	0	0		
<i>Solanum aethiopicum</i>	0.0235***	0.00015*	0.1939***	0.16387***	0.40045***	0.27056***	0.1608***	0.32763***	0	0

S5 Table. Frequency of the Malagasy MLVA type.

Haplotype	Number of strains	Frequency of haplotype per strains %
MT001	1	0.228310502
MT002	1	0.228310502
MT003	24	5.479452055
MT004	83	18.94977169
MT005	1	0.228310502
MT006	8	1.826484018
MT007	81	18.49315068
MT008	2	0.456621005
MT009	4	0.913242009
MT010	3	0.684931507
MT011	2	0.456621005
MT012	2	0.456621005
MT013	1	0.228310502
MT014	8	1.826484018
MT015	9	2.054794521
MT016	3	0.684931507
MT017	4	0.913242009
MT018	3	0.684931507
MT020	9	2.054794521
MT021	3	0.684931507
MT022	2	0.456621005
MT023	7	1.598173516
MT024	2	0.456621005
MT025	2	0.456621005
MT026	1	0.228310502
MT027	4	0.913242009
MT028	5	1.141552511
MT029	2	0.456621005
MT030	1	0.228310502
MT031	1	0.228310502
MT235	1	0.228310502
MT236	14	3.196347032
MT237	1	0.228310502
MT238	3	0.684931507
MT239	3	0.684931507
MT240	2	0.456621005
MT241	2	0.456621005
MT242	1	0.228310502
MT243	4	0.913242009
MT244	1	0.228310502
MT245	6	1.369863014
MT246	1	0.228310502
MT247	5	1.141552511
MT248	2	0.456621005

Haplotype	Number of strains	Frequency of haplotype per strains %
MT249	21	4.794520548
MT250	1	0.228310502
MT251	7	1.598173516
MT252	3	0.684931507
MT253	1	0.228310502
MT254	5	1.141552511
MT255	1	0.228310502
MT256	1	0.228310502
MT257	2	0.456621005
MT258	1	0.228310502
MT259	7	1.598173516
MT260	1	0.228310502
MT261	4	0.913242009
MT262	1	0.228310502
MT263	1	0.228310502
MT264	1	0.228310502
MT265	1	0.228310502
MT266	4	0.913242009
MT267	5	1.141552511
MT325	1	0.228310502
MT326	2	0.456621005
MT327	2	0.456621005
MT328	2	0.456621005
MT329	1	0.228310502
MT330	2	0.456621005
MT331	1	0.228310502
MT332	1	0.228310502
MT333	1	0.228310502
MT334	7	1.598173516
MT340	1	0.228310502
MT341	1	0.228310502
MT342	1	0.228310502
MT343	1	0.228310502
MT344	1	0.228310502
MT345	1	0.228310502
MT346	1	0.228310502
MT347	2	0.456621005
MT348	1	0.228310502
MT349	1	0.228310502
MT350	1	0.228310502
MT351	1	0.228310502
MT352	1	0.228310502
MT353	1	0.228310502
MT354	1	0.228310502
MT355	1	0.228310502
MT356	1	0.228310502

Haplotype	Number of strains	Frequency of haplotype per strains %
MT357	1	0.228310502
MT358	1	0.228310502
MT359	1	0.228310502
MT360	1	0.228310502
MT361	1	0.228310502
MT362	1	0.228310502
MT363	1	0.228310502
TOTAL	438	