



Génomique comparative des bactéries *Dickeya solani* et *Pectobacterium wasabiae*, pathogènes émergents chez *Solanum tuberosum*

Slimane Khayi

► **To cite this version:**

Slimane Khayi. Génomique comparative des bactéries *Dickeya solani* et *Pectobacterium wasabiae*, pathogènes émergents chez *Solanum tuberosum*. Sciences agricoles. Université Paris-Saclay, 2015. Français. <NNT : 2015SACLS048>. <tel-01295341>

HAL Id: tel-01295341

<https://tel.archives-ouvertes.fr/tel-01295341>

Submitted on 30 Mar 2016

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

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

THÈSE DE DOCTORAT
de
l'UNIVERSITÉ MOULAY ISMAÏL
et de
l'UNIVERSITÉ PARIS-SACLAY
préparée à l'Université Paris Sud

ÉCOLE DOCTORALE N° 567
Sciences du végétal : du gène à l'écosystème

CEDoc : SCIENCES FONDAMENTALES ET APPLIQUÉES
Sciences Biologiques et leurs Applications

Spécialité de doctorat : Biologie

Par

M. Slimane KHAYI

Génomique comparative des bactéries *Dickeya solani* et *Pectobacterium wasabiae*,
pathogènes émergents chez *Solanum tuberosum*

Thèse présentée et soutenue à Gif-sur-Yvette, le Mercredi 09 Décembre 2015 :

Composition du Jury :

Marie-Agnès JACQUES,
William NASSER,
Jean-Luc PERNODET,
Yves LE HINGRAT,
Mohammed El HASSOUNI,
Mohieddine MOUMNI,
Denis FAURE,

Directrice de Recherche INRA, Angers,
Directeur de recherche CNRS, Lyon,
Directeur de recherche CNRS, Gif sur Yvette,
Responsable R&D FN3PT/RD3PT, Paris,
Professeur USMBA, Fès,
Professeur UMI, Meknès,
Directeur de Recherche CNRS, Gif sur Yvette,

Rapporteur
Rapporteur
Président du jury
Examineur
Examineur
Co-Directeur de thèse
Directeur de thèse



"In the name of Allah, the Merciful, the Compassionate"

Dédicace
À mes très chers parents,
Ma mère Sfia Ellouli
Mon père Moha Khayi
À ma sœur adorée, mon frère
et à toute ma famille à qui j'exprime tout mon amour

Remerciements

Tout d'abord je tiens à remercier le président de l'Université Paris sud, Pr. Jacques Bittoun et le président de l'Université Moulay Ismaïl, Pr. Hassane SAHBI pour avoir autorisé la cotutelle de ma thèse au sein de leurs universités. Je tiens à remercier le Pr. Jacqui Shykoff, directrice de l'école doctorale Sciences du végétal (ED 567) pour le suivi de ma formation doctorale.

Je tiens à remercier infiniment mes directeurs de thèse Denis Faure et Mohieddine Mounni pour m'avoir accordé leur confiance et m'avoir encadré durant cette thèse en cotutelle Franco-Marocaine. Je remercie Denis Faure pour m'avoir accueilli au sein de son équipe durant ces 3 années de thèse et pour m'avoir offert l'opportunité de vivre ces moments de sciences et de convivialité, merci pour son encadrement, ses conseils et orientations qui m'ont permis de mener à terme et en bonnes conditions mes travaux de recherches. Je remercie Mohieddine Mounni pour son soutien et ses conseils. Merci à tous deux pour vos encouragements précieux tout au long de ma thèse.

Je remercie Marie-agnès Jacques (INRA, Angers) et William Nasser (CNRS, Lyon) d'avoir accepté d'être les rapporteurs de ce travail de thèse, ainsi que Jean-Luc Pernodet (I2BC, Gif sur Yvette), Yves Le Hingrat (FN3PT/FD3PT, Paris) et Mohammed El Hassouni (USMBA, Fès).

Je tiens à remercier toute l'équipe Interactions Plantes-Bactéries pour leur soutien durant ma thèse, je remercie Yves Dessaux pour ses conseils et ses orientations précieux et pour avoir pris le temps de relire plusieurs de mes manuscrits. Merci Yves infiniment pour ta gentillesse et ton sourire.

Je remercie Samuel Mondy et Anthony Kwasiborski pour m'avoir accompagné et encadré durant mes analyses bioinformatiques, merci d'avoir pris le temps pour résoudre les problèmes de scripts qui ne tournaient jamais la première fois !

Un grand merci à Catherine Grandclement pour ses conseils et encouragements durant ces 3 années de thèse et particulièrement lors de la rédaction de mon manuscrit.

Je tiens également à remercier mes très chère(s) ami(e)s : Yannick Raoul Des Essarts, Jérémy Cigna, Pauline Blin, Almudena Gonzalez-Mula, Fabienne Pierre, Claudie Barnier, Elisa Violeta Bertini, Julien Lang qui ont su instaurer une ambiance amicale au cours de ces années.

Une pensée pour celles et ceux qui sont partis vers de nouvelles aventures et qui ont contribué directement ou indirectement au bon déroulement de ma thèse : Mélanie Tannières, Sara Planamente, Noura Yahyaoui, Teik-Min Chong, Kar Wai Hong, Nadia Mhedbi-Hajri, Samy Carbonnel et Nicolas Mothe.

Un grand merci à mes collègues au sein du laboratoire GCTMI, Said Oulghazi, Asmae Benhida et Nada Lafkih ainsi qu'à mes amis Abdelmounaim Mouhajir, Adnane El Yaacoubi, Omar Mohaddach et Imane Boularab pour leur soutien et leurs encouragements.

Finalement, au terme de ce travail je tiens à remercier toutes les personnes que j'ai rencontrées lors de mes études doctorales et qui ont contribué de près ou de loin au bon déroulement de cette thèse.

Résumé

Des bactéries pectinolytiques appartenant aux genres *Pectobacterium* et *Dickeya* sont des agents pathogènes chez *Solanum tuberosum* et sur d'autres cultures ornementales et horticoles. Ces bactéries sont responsables de la maladie de la jambe noire et de la pourriture molle lors de la culture et du stockage des tubercules. Ce travail de thèse est divisé en deux axes : 1) Etude de la diversité d'une population du pathogène *D. solani* par approche de génomique comparée afin de mieux comprendre la structure génomique de cette espèce émergente, 2) L'assemblage du génome et la caractérisation génomique des facteurs de virulence chez *Pectobacterium wasabiae* RNS 08421A.

L'analyse des génomes de 20 isolats de *D. solani* issus d'environnements différents, par une approche de génomique comparative associée à des analyses fonctionnelles, a révélé une forte homogénéité génétique au sein de la majorité des souches (16/20). De plus, cette analyse a permis de caractériser un nouveau sous-groupe au sein de l'espèce *D. solani*, représenté par la souche 0512 (1/20). En revanche, d'autres isolats (3/20) montrent des variations de quelques centaines à quelques milliers de SNPs/InDels qui sont regroupés dans des îlots génomiques. Leur analyse phylogénétique révèle qu'ils proviennent d'autres pathogènes par transferts horizontaux. Par ailleurs, l'analyse des fonctions affectées par les SNPs/InDels a permis de prédire, puis de vérifier sur pomme de terre, qu'un des isolats était faiblement virulent.

La deuxième partie de mon travail porte sur l'assemblage, la caractérisation et l'analyse du génome de la souche RNS 08.42.1A de *P. wasabiae*, qui a été isolée en France. La génomique comparative avec 3 autres souches de *P. wasabiae* d'origines géographiques différentes, a révélé à la fois une forte similitude au niveau de la séquence génomique (ANI > 99%) et une synténie conservée des gènes de virulence. En outre, notre analyse a mis en évidence une nette distinction entre ces quatre souches de *P. wasabiae* (isolées de *S. tuberosum*) et la souche type japonaise *P. wasabiae* CFBP 3304^T (isolée du raifort). Dans *P. wasabiae* RNS 08.42.1A, les gènes de synthèse et de perception du quorum sensing, *expl/expR*, présentent une plus forte homologie avec leurs orthologues chez *P. atrosepticum* et *P. carotovum* (90%) qu'avec leurs homologues chez *P. wasabiae* (70%). Ceci suggère une acquisition de ces gènes par transfert horizontal au sein d'une population de pathogènes infectant la même plante hôte.

Mots clés : Jambe noire, *Pectobacterium*, *Dickeya*, Génomique, Transfert horizontal, Phylogénie

Abstract

Some pectolytic bacteria *Pectobacterium* and *Dickeya* species cause important diseases on *Solanum tuberosum* and other arable and horticultural crops. These bacteria are responsible for blackleg in the field and tuber soft rots in storage and in transit as well as in the field worldwide. The main objectives of this thesis are: 1) To study the diversity of a *D. solani* population using comparative genomics approaches with the aim of understanding the genomic structure and evolution of this emerging species, 2) Characterization and genomic analysis of virulence factors in *Pectobacterium wasabiae* RNS 08421A.

Using comparative genomics approaches combined with functional assays, the analysis of the genomes of 20 isolates of *D. solani* from different environments, revealed a strong genetic homogeneity within the majority of the strains (16/20). Moreover, this analysis allowed to characterize a new sub-group within *D. solani* species, represented by the strain 0512 (1/20). In contrast, some strains (3/20) showed variations from hundreds to a few thousand of SNPs/Indels which are grouped in what we called "Genomic islands". Phylogenetic analysis of these regions showed that they were acquired from other pathogens by HGT. Furthermore, the analysis of the functions affected by SNPs/Indels allowed predicting and then checking on potatoes, that one of these strain was less virulent.

The second part of my work involves assembly, characterization and analysis of the genome of the strain *P. wasabiae* RNS 08.42.1A, which was isolated in France. Comparative genomics with three other *P. wasabiae* strains from different geographical origins, revealed a strong similarity in the genome sequence (ANI > 99%) and conserved synteny of virulence genes. In addition, our analysis showed a clear distinction between the strains of *P. wasabiae* isolated from *S. tuberosum* and the type strain *P. wasabiae* CFBP 3304^T (isolated from horseradish). In *P. wasabiae* RNS 08.42.1A, the complex system for synthesis and perception of quorum sensing signal *expI/expR*, exhibit higher homology with their orthologs in *P. atrosepticum* and *P. carotovorum* (90%) than their homologs in *P. wasabiae* (70%). This suggests acquisition of these genes by HGT within a population of pathogens infecting the same host plant.

Keywords : Soft rot, *Dickeya*, *Pectobacterium*, Genomics, Horizontal transfer, Phylogeny

Table des matières

Remerciements.....	4
Résumé	6
Abstract.....	7
Table des matières	8
Liste des figures.....	10
Liste des tableaux	11
Liste des abréviations.....	12
Avant-propos.....	13
Introduction générale.....	14
Objectifs de la thèse.....	17
Chapitre I. Généralités sur les <i>Pectobacterium</i> et <i>Dickeya</i> spp.	18
I.1. Les phytopathogènes pectinolytiques <i>Pectobacterium</i> et <i>Dickeya</i> spp.....	19
I.1.1. Les maladies de la jambe noire et de la pourriture molle causées par <i>Pectobacterium</i> et <i>Dickeya</i> spp.....	20
I.1.2. Taxinomie et phylogénie des <i>Pectobacterium</i> et <i>Dickeya</i> spp.	21
I.2. Epidémiologie et conditions environnementales.....	24
I.2.1. Distribution des <i>Pectobacterium</i> et <i>Dickeya</i> spp. dans l'environnement.....	24
I.2.2. Implication des facteurs environnementaux dans la maladie	25
I.3. La lutte contre les bactéries pectinolytiques <i>Dickeya</i> et <i>Pectobacterium</i> spp.	26
I.3.1. Lutte chimique.....	26
I.3.2. Lutte génétique.....	27
I.3.3. Lutte biologique	29
I.4. Outils de la génomique appliqué aux bactéries pectinolytiques.....	31
I.4.1. Technologies de séquençage.....	31
I.4.2. Assemblage des génomes	37
I.5. Apport de la génomique comparative à l'étude des <i>Pectobacterium</i> et <i>Dickeya</i> spp.	40
I.5.1. Génomique et taxinomie.....	40
I.5.2. Contribution de la génomique à l'identification et à la détection des <i>Pectobacterium</i> et <i>Dickeya</i> spp.....	42
I.5.3. Génomique et évolution des <i>Pectobacterium</i> et <i>Dickeya</i> spp.	45
I.5.4. Génomique et virulence des <i>Pectobacterium</i> et <i>Dickeya</i> spp.	46
Chapitre II. Genome sequence of the emerging plant pathogen <i>Dickeya solani</i> strain RNS	
08.23.3.1A.....	49
II.1. Abstract	51
II.2. Introduction.....	51
II.3. Illumina sequencing and assembly.....	51
II.4. PacBio sequencing and assembly	52
II.4.1. Single Molecule Real Time (SMRT) Sequencing	52
II.4.2. De novo Assembly using Hierarchical Genome Assembly Process (Hgap)	53
II.5. Complete genome sequence	53
II.6. Nucleotide sequence accession numbers	54
II.7. Acknowledgments.....	55
II.8. References	55
Chapitre III. Population genomics reveals additive and replacing horizontal gene transfers in the emerging pathogen <i>Dickeya solani</i>	56
III.1. Abstract.....	59
III.2. Background.....	59
III.3. Results.....	61
III.3.1. Complete genome of the <i>D. solani</i> 3337	61

III.3.2. Positioning the sequenced <i>D. solani</i> strains within the <i>Dickeya</i> genus.....	62
III.3.3. Overview of the SNP and InDel variations in <i>D. solani</i> genomes	63
III.3.4. Heterogeneous distribution of the SNP and InDel variations in <i>D. solani</i> genes.....	63
III.3.5. The mosaic genome of <i>D. solani</i> 0512 might define a novel <i>D. solani</i> sub-group	66
III.3.6. Infra-species replacing HGT in Dsl 07-7	69
III.3.7. Inter-species replacing HGT in <i>D. solani</i> strains 9134 and 9019.....	70
III.3.8. Plasmid acquisition in <i>D. solani</i> strain 9019 from <i>Burkholderia</i>	72
III.3.9. In <i>D. solani</i> 3296, variations in flagellar genes correlated motility and virulence decrease	73
III.4. Discussion	74
III.5. Conclusions	76
III.6. Methods.....	76
III.6.1. Bacterial strains and growth conditions.....	76
III.6.2. DNA extraction and sequencing	77
III.6.3. Assembly, variants calling and genome sequence analysis.....	77
III.6.4. Nucleotide sequence accession number	78
III.6.5. Aggressiveness and motility assays	78
III.7. Supplementary data	79
III.8. Acknowledgements.....	88
III.9. References	88
Chapitre IV. Genomic overview of the phytopathogen <i>Pectobacterium wasabiae</i> strain RNS 08.42.1A suggests horizontal acquisition of quorum-sensing genes	92
IV.1. Abstract	95
IV.2. Introduction.....	95
IV.3. Materials and methods	97
IV.3.1. Isolation of the bacterial strain and culture conditions	97
IV.3.2. DNA extraction.....	97
IV.3.3. Molecular characterization and genome sequencing.....	98
IV.3.4. ORFs annotation, phylogenetic tree and genomic comparison.....	99
IV.3.5. Nucleotide sequence accession number	100
IV.4. Results and discussion	100
IV.4.1. Isolation and characterization of <i>P. wasabiae</i> RNS 08.42.1A.....	100
IV.4.2. <i>Pectobacterium wasabiae</i> RNS 08.42.1A genome sequence.....	101
IV.4.3. Positioning <i>P. wasabiae</i> RNS 08.42.1A within <i>Pectobacterium</i> and <i>Dickeya</i>	102
IV.4.4. Synteny relatedness among <i>P. wasabiae</i> strains	103
IV.4.5. PCWDEs and virulence determinants	104
IV.4.6. <i>P. wasabiae</i> RNS 08.42.1A displays an original <i>expl-expR1</i> system.....	107
IV.4.7. Secretions systems	110
IV.4.8. Toxin antitoxin HigA/HigB complex shared only by RNS 08.42.1A	110
IV.5. Conclusions	112
IV.6. Supplementary data	113
IV.7. Acknowledgements	117
IV.8. References	117
Chapitre V. Discussion générale et perspectives	122
V.1. Discussion générale	123
V.2. Perspectives.....	129
Références bibliographiques	132
Annexes	146

Liste des figures

Figure I.1 Structure de la paroi végétale.	19
Figure I.2 Les symptômes de la maladie de la jambe noire et de la pourriture molle sur <i>Solanum tuberosum</i>	21
Figure I.3 ddNTP et dNTP. L'absence du groupement hydroxyle sur le ddNTP bloque la progression de la synthèse du brin d'ADN.	32
Figure I.4 Présentation de la technologie 454.	35
Figure I.5 Résumé des étapes d'assemblage d'un génome.	40
Figure I.6 Evolution du nombre de génomes de <i>Pectobacterium</i> et <i>Dickeya</i> spp. déposés sur NCBI entre 2003 et 2015 (http://www.ncbi.nlm.nih.gov/genome).	42
Figure I.7 Schéma simplifié des différents déterminants de virulence connus chez <i>Pectobacterium atrosepticum</i> . D'après Toth <i>et al.</i> (2006)	48
Figure II.1 Map of <i>Dickeya solani</i> 3337 genome.	54
Figure III.1 MLSA and ANIs of <i>D. solani</i> strains.	62
Figure III.2 Number of genes affected by variations (SNPs and Indels).	65
Figure III.3 Mapping of the clustered and scattered SNP/InDel variations using Dsl 3337 as a reference genome	66
Figure III.4 Mapping and phylogeny of the Dsl 0512 variant genes.	68
Figure III.5 Replacing HGT region 14 (RGT ₁₄ 07-7) in <i>D. solani</i> 07-7.	69
Figure III.6 Replacing HGT region 4 (RGT ₄ 9134) in <i>D. solani</i> 9134.	71
Figure III.7 Replacing HGT region 7 (RGT ₇ 9019) in <i>D. solani</i> 9019.	72
Figure III.8 Motility and aggressiveness assays performed on potato tubers.	73
Figure III.S1 Synteny between the strain <i>D. solani</i> 3337 and the draft genome <i>D. solani</i> 0512.	79
Figure III.S2 Protein-based phylogenetic trees revealing Dsl 0512 as a member of a distinct sub-cluster within the <i>D. solani</i> species.	80
Figure III.S3 Protein-based phylogenetic trees of different RGTs in Dsl 07-7.	81
Figure III.S4 Protein-based phylogenetic trees of different RGTs in Dsl 9134.	82
Figure III.S5 Protein-based phylogenetic trees of different RGTs in Dsl 9019.	83
Figure III.S6 Local alignment of FliC and FliN proteins.	84
Figure IV.1 Blackleg symptoms in greenhouse plant assay.	101
Figure IV.2 MLSA-based (<i>fusA</i> , <i>gyrB</i> , <i>recA</i> , <i>dnaX</i>) relation tree and ANI values using <i>P. wasabiae</i> RNS 08.42.1A as a reference.	103
Figure IV.3 Synteny between <i>P. wasabiae</i> strains RNS 08.42.1A, SCC3193 and WPP163.	104
Figure IV.4 The QS regulatory system ExpI-ExpR1 of <i>P. wasabiae</i> RNS 08.23.1A.	108

Liste des tableaux

Tableau I.1 Récapitulatif de la nomenclature des espèces de <i>Pectobacterium</i> et <i>Dickeya</i>	23
Tableau I.2 Amorces utilisées pour la détection et l'identification des différentes espèces <i>Pectobacterium</i> et <i>Dickeya</i> spp. par PCR, qPCR et PCR multiplexe..	44
Table III.S1 <i>Dickeya solani</i> strains used in this study	85
Table III.S2 Sequencing data and mappings on the Dsl 3337 genome	86
Table III.S3 Variants distribution on the strains vs. Dsl 3337	86
Table III.S4 Non-synonymous variants.....	87
Table III.S5 Other genomes used in this study.	88
Table IV.1 List of the primers used in this study	98
Table IV.2 Some characteristics of the <i>P. wasabiae</i> strains and their genome sequences	102
Table IV.3 Plant cell wall degrading enzymes in <i>P. wasabiae</i> RNS 08.42.1A.....	106
Table IV.4 Virulence determinants in <i>P. wasabiae</i> RNS 08.42.1A.....	107
Table IV.5 Virulence regulators in <i>P. wasabiae</i> RNS 08.42.1A.....	109
Table IV.S1 Functional comparisons of specific features between RNS 08.42.1A and all other sequenced <i>P. wasabiae</i> strains	113

Liste des abréviations

AAI Average amino acids identity

ADN (DNA) Acide désoxyribonucléique
(Deoxyribonucleic acid)

ANI Average nucleotide identity

ARN (RNA) Acide ribonucléique
(Ribonucleic acid)

ATP Adenosine triphosphate

BLAST Basic local alignment search tool

CDS Coding DNA sequence

CFBP Collection française de bactéries
phytopathogènes

CNRS Centre national de la recherche
scientifique

CVP Cristal violet Pectate

DAPG 2,4-diacetylphloroglucinol

ddNTP Didésoxyribonucéotides
triphosphate

dNTP Désoxyribonucéotides triphosphate

DsI *Dickeya solani*

Dz *Dickeya zeae*

FAO Food and agriculture organization of the
united nations

FAOSTAT The Statistics division of FAO

Fig. Figure

FN3PT Fédération nationale des producteurs
de plants de pomme de terre

HGT Horizontal gene transfer

InDel Insertion/délétion

INRA Institut national de la recherche
agronomique

LPS Lipopolysaccharide

M Million

MLSA Multilocus sequence analysis

NAHL N-acyl homoserine lactone

NCBI National center for biotechnology
information

NGS Next generation sequencing

OGM Organisme modifié génétiquement

pb Paire de base

Pa *Pectobacterium atrosepticum*

Pcc *Pectobacterium carotovorum* subsp.
carotovorum

PCR Polymerase chain reaction

PCWDE Plant cell wall-degrading enzymes

Pwa *Pectobacterium wasabiae*

qPCR PCR en temps réel

QS Quorum sensing

RAST Rapid annotation using subsystem
technology

RGT Replacing HGT region

SNP Single nucleotide polymorphism

T Température

TY Tryptone yeast extract

UFC/CFU Unité formant colonie/Colony-
forming unit

UV Ultraviolet

Avant-propos

Ce travail de doctorat a été effectué dans le cadre d'une thèse en cotutelle entre l'Université Paris-sud et l'Université Moulay Ismaïl à Meknès. Les travaux de recherches ont été conduits au sein de l'équipe Interactions Plantes-Bactéries à l'Institut de Biologie Intégrative de la Cellule (I2BC) (anciennement Institut des Sciences du Végétal). Cette collaboration a été financée par plusieurs programmes de coopération et bourses:

- Projet Franco-Marocain " Volubilis no. MA/09/218, 2009-2012"
- Projet Franco-Marocain "PRAD 14-02, Campus France no. 30229 ZK, 2014-2016"
- Bourse d'excellence accordée par le ministère de l'enseignement supérieur marocain. (no. H011/007, 2013-2015)
- Bourse de soutien financier à la cotutelle accordée par l'Université Paris sud (2014-2015).

L'ensemble des organismes qui ont soutenus ce projet sont remerciés.

Introduction générale

De nos jours, la pomme de terre est cultivée dans de nombreux pays et représente la quatrième culture vivrière la plus importante à l'échelle mondiale après le maïs, le blé et le riz (Thomas et Sansonetti, 2009).

La culture de pomme de terre était pratiquée il y a déjà 8000 ans en Amérique du sud dans les montagnes des Andes. Vers les années 1570, la pomme de terre a été introduite en Europe par les espagnols, tout d'abord comme plante décorative, puis plus tard pour la consommation alimentaire animale puis humaine. La culture s'est alors rapidement répandue dans le reste de l'Europe et du monde (Hawkes et Francisco-Ortega, 1993)

La pomme de terre est actuellement produite dans 126 pays et les surfaces de culture sont en constante augmentation surtout dans les pays en développement (Leff *et al.*, 2004). Plus de 376M de tonnes sont produites dans le monde dont 80% en Europe et en Asie (Van Der Zaag et Horton, 1983) (FAOSTAT, 2013). En Afrique, cette culture occupe un rang moins important qu'en Europe. En 2013 environ 30M de tonnes y ont été produites ce qui représentait 7% de la production mondiale, mais plus de 50% de la récolte était réalisée dans seulement 7 pays, Algérie, Égypte, Malawi, Afrique du Sud, Rwanda, Kenya et Maroc (par ordre décroissant de production) (FAOSTAT, 2013).

La culture de la pomme de terre est sensible à différents pathogènes bactériens, viraux ou fongiques (Arora et Khurana, 2004). Les infections bactériennes dues aux bactéries des genres *Pectobacterium* et *Dickeya* causent des dégâts de plus en plus importants au champ et lors de la conservation des tubercules en provoquant la dégradation des tissus végétaux. Ces bactéries pectinolytiques causent aussi d'importants dégâts sur d'autres cultures légumières et ornementales (Hélias, 2008). Les bactéries nécrotrophes utilisent comme nutriments, pour leur croissance, les composés et les électrolytes cellulaires (Pérombelon, 2002). Ces bactéries sont classées parmi les dix phytopathogènes bactériennes les plus étudiés et les plus importants sur

le plan économique (Mansfield *et al.*, 2012). Actuellement, il n'existe aucun moyen de lutte efficace et aucune variété cultivée n'est totalement résistante à ces phytopathogènes (Charkowski, 2006 ; Latour *et al.*, 2008 ; Czajkowski *et al.*, 2011).

La caractérisation génétique et phénotypique ainsi que le processus de pathogénie des *Pectobacterium* et *Dickeya* spp. ont été largement étudiés ces dernières décennies (Perombelon et Kelman, 1980 ; Elphinstone, 1987 ; Pérombelon *et al.*, 1989 ; De Boer, 2004 ; Charkowski, 2006). Plus récemment, le développement des technologies de séquençage a donné un nouvel élan à la recherche sur les *Pectobacterium* et *Dickeya* spp. Les études menées sur les premiers génomes séquencés de *Pectobacterium* et *Dickeya* ont largement contribué à la compréhension de la structure génomique ainsi qu'à la caractérisation des déterminants de virulence de ces pathogènes (Bell *et al.*, 2004 ; Glasner *et al.*, 2011). Le développement de nouveaux outils de taxinomie a également permis d'identifier et de reclassifier plusieurs espèces et sous-espèces parmi ces pathogènes.

C'est ainsi qu'il y a quelques années, une nouvelle espèce a été identifiée et nommée *Dickeya solani* isolée de plusieurs plantes ornementales (hyacinth et iris) (Sławiak *et al.*, 2009) et sur pomme de terre. Celle-ci est responsable de dégâts importants et tend à devenir un pathogène redoutable causant les maladies de la jambe noire et de la pourriture molle de pomme de terre et aussi sur de nombreuses autres cultures en Europe (Toth *et al.*, 2011 ; van der Wolf *et al.*, 2014). Plusieurs études ont rapporté le caractère clonal de cette espèce et ont montré que les régulateurs et déterminants de virulence décrits dans *D. dadantii* sont également impliqués dans la virulence de *D. solani* (Czajkowski *et al.*, 2013a ; Garland *et al.*, 2013 ; Pédrón *et al.*, 2014 ; Potrykus *et al.*, 2014 ; Sławiak *et al.*, 2009). Jusqu'à présent, aucune étude n'a été faite pour comprendre la diversité et l'évolution génétique de cette espèce en prenant en compte l'ensemble du génome. Dans ce contexte, la première partie de mon travail de thèse s'intéresse à l'analyse génomique d'un ensemble de 20 souches de *D. solani* (dates d'isolement et lieux de prélèvement différents)

pour comprendre la structure génomique de cette espèce et l'évolution génétique qu'elle suit depuis son émergence. Dans une deuxième partie, nous nous sommes intéressés au séquençage et à l'assemblage du génome de la souche RNS 08.42.1A de *P. wasabiae*, isolée en France, puis à sa caractérisation par génomique comparative à l'aide des 4 autres souches de *P. wasabiae* disponibles dans les bases de données publiques (NCBI).

Le présent manuscrit de thèse est organisé en 5 chapitres. Le premier chapitre comprend une introduction bibliographique sur les pathogènes et les maladies de la pourriture molle et de la jambe noire ainsi qu'un aperçu sur l'apport de la génomique à l'étude de ces pathogènes. Le deuxième chapitre présente le séquençage et l'assemblage de la souche *Dickeya solani* 3337 qui représente le premier génome complet de l'espèce *D. solani*. Le troisième chapitre est dédié à l'analyse génomique de population de *D. solani* en mettant en évidence la diversification et l'évolution génétique de cette espèce. Le quatrième chapitre traite du séquençage et de l'assemblage ainsi que la caractérisation par génomique comparée de *P. wasabiae* RNS 08.42.1A. Le cinquième chapitre est consacré à la discussion générale et aux perspectives. Les chapitres 1 et 5 sont rédigés en français, les chapitres 2, 3 et 4 sont rédigés en anglais sous forme d'articles publiés dans des revues internationales à comité de lecture.

Objectifs de la thèse

Mon travail de thèse a comme objectif global de contribuer à la compréhension de la structure et de l'évolution génétique des pathogènes *Dickeya solani* et *Pectobacterium wasabiae* par des approches de génomique comparée et de génomique des populations.

La méthodologie de travail s'articule autour de quatre grands axes :

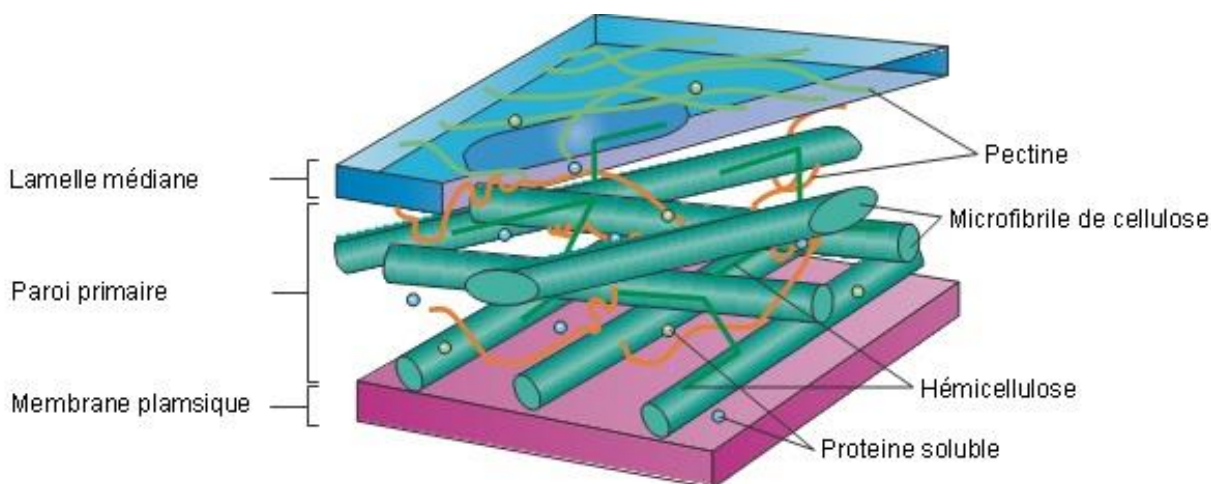
- Séquençage et assemblage du génome de référence *Dickeya solani* 3337.
- Analyse de la diversité et l'évolution génétique d'une population composée de 20 souches de *D. solani* par approche de génome *mapping* et recherche des SNPs/InDels.
- Séquençage et assemblage du génome de la souche de *Pectobacterium wasabiae* RNS 08.42.1A.
- Caractérisation de *P. wasabiae* RNS 08.42.1A par approche de génomique comparative par rapport à 4 souches de *P. wasabiae* disponibles sur NCBI.

Chapitre I. Généralités sur les *Pectobacterium* et *Dickeya* spp.

I.1. Les phytopathogènes pectinolytiques *Pectobacterium* et *Dickeya* spp.

Les bactéries pectinolytiques sont des micro-organismes capables de produire des enzymes PCWD (*plant cell wall-degrading enzymes*) impliquées notamment dans la dégradation de la pectine. Ce composant de la paroi des cellules végétales, est un polysaccharide caractérisé par un squelette d'acide α -D-galacturonique plus ou moins ramifié avec de faibles quantités de résidus de sucres, tels que α -L-rhamnose. La pectine joue plusieurs rôles dans la physiologie, le développement et la croissance des cellules végétales (Willats *et al.*, 2001) (Fig. I.1). Plusieurs bactéries des genres *Pectobacterium*, *Dickeya*, *Pseudomonas*, *Burkholderia* et *Xanthomonas* possèdent la faculté de produire des pectinases (Rombouts, 1972).

Figure I.1 Structure de la paroi végétale. La pectine est présente dans la lamelle moyenne de la paroi végétale où elle joue un rôle dans la structure et la physiologie de la cellule végétale. Modifiée d'après Sticklen (2008)



Les enzymes pectinolytiques sont bien identifiées et connues chez les bactéries des genres *Pectobacterium* et *Dickeya* (Hugouvieux-Cotte-Pattat *et al.*, 1996), où elles sont responsables de la dégradation de la paroi cellulaire de plusieurs plantes hôtes, en particulier la pomme de terre (*Solanum tuberosum*). Les bactéries des genres *Dickeya* et *Pectobacterium* sont particulièrement associées aux maladies de la jambe noire et de la pourriture molle sur les plantes hôtes alors que les espèces d'autres genres sont considérées comme des bactéries pectinolytiques secondaires qui interviennent après initiation de la maladie (Pérombelon et Lowe, 1975).

I.1.1. Les maladies de la jambe noire et de la pourriture molle causées par *Pectobacterium* et *Dickeya* spp.

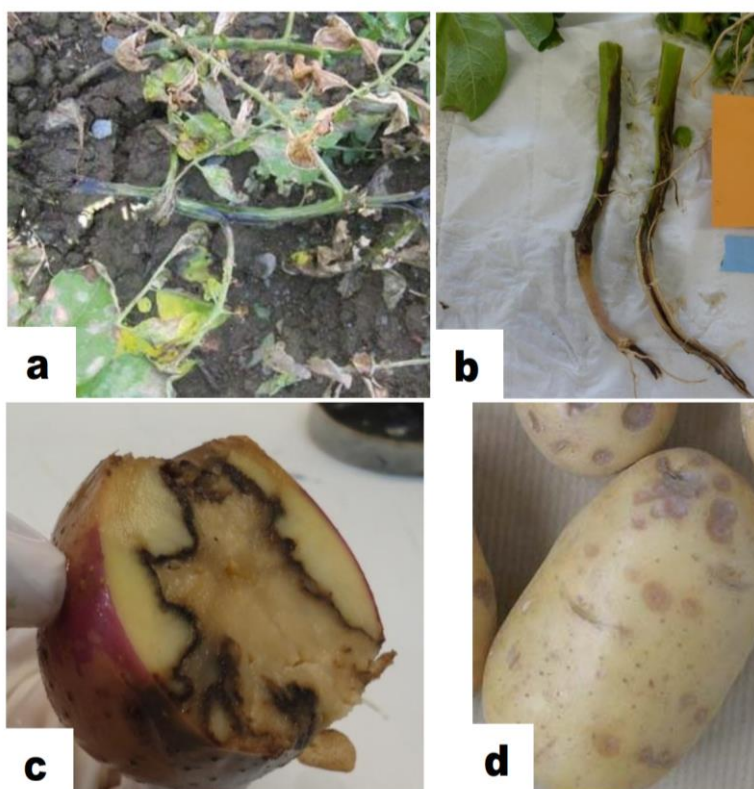
Les agents pathogènes peuvent demeurer latents dans le sol ou sur le tubercule mère et ils peuvent être propagés via différents vecteurs (eau d'irrigation ou de ruissellement, insectes..). Dès que les conditions pédoclimatiques deviennent favorables, le pathogène se multiplie et envahit les tissus de la plante (Pérombelon et Hyman, 1989). Les phytopathogènes *Dickeya* et *Pectobacterium* disposent d'un arsenal enzymatique important (pectinanes, cellulases, protéases et xylanases) capable de dégrader la paroi végétale et de causer les maladies de la pourriture molle et de la jambe noire (Sjöblom, 2009).

Chez ces phytopathogènes, comme chez de nombreuses bactéries, l'expression des facteurs de virulence, tels que les enzymes de macération et les effecteurs de la mort cellulaire, est régulé par *quorum sensing* (QS) et par des systèmes à deux composants (TCS : *Two-component system*) (Keller et Surette, 2006 ; Parkinson et Kofoed, 1992). Dans le cas des TCS, en général il s'agit d'un senseur du signal souvent membranaire associé à un régulateur cytoplasmique de réponse. Plus de 30 systèmes TCS ont été identifiés chez les *Pectobacterium* et *Dickeya* spp. tels que les systèmes ExpS/ExpA, PehR/PehS et PhoP/PhoQ (Hyytiäinen, 2005 ; Kettani Halabi, 2012). Le QS est un système qui couple l'expression de gènes cibles à la densité bactérienne via la synthèse et la perception des signaux moléculaires appelés autoinducteurs. Dans le cas de *Pectobacterium* et *Dickeya* spp., la molécule signal sécrétée fait partie de la classe des N-acyl homosérines lactones (NAHL) (Barnard et Salmond, 2007).

Suite à la dégradation des tissus de la plante, les symptômes peuvent s'exprimer dans plusieurs parties de la plante en végétation mais également lors de la conservation des tubercules après récolte. Le tubercule mère peut être porteur de l'agent pathogène qui demeure en dormance dans les lenticelles et les blessures des tubercules. À la faveur de conditions adéquates, le pathogène se multiplie et provoque la pourriture molle du tubercule mère avant ou durant la phase germinative provoquant des manques ou retards à la levée. Au stade plante, les premiers

symptômes apparaissent sous forme d'un feuillage vert pâle ou jaunâtre suivis d'un rabougrissement de la plante (Fig. 1.2a). Puis la maladie peut se manifester sous forme d'une nécrose humide sur les tiges qui prennent alors une couleur noire (Fig. 1.2b) d'où le nom de maladie de la jambe noire (Hélias, 2008 ; Pérombelon, 2002). Après la récolte, les tubercules en stockage peuvent développer des symptômes de pourriture molle (Fig. 1.2c) ou de pourriture lenticellaire dans le cas d'attaques localisées au niveau des lenticelles (Fig. 1.2d) (Bétencourt et Prunier, 1965 ; Hélias, 2008 ; Hélias *et al.*, 2000).

Figure 1.2 Les symptômes de la maladie de la jambe noire et de la pourriture molle sur *Solanum tuberosum* **a** et **b**: Symptômes de jambe noire sur tige et feuillage, d'après V. Hélias (a) et J. Cigna (b) **c**: Pourriture molle de tubercule, **d**: Pourriture lenticellaire, d'après V. Hélias (2008) (c et d)



1.1.2. Taxinomie et phylogénie des *Pectobacterium* et *Dickeya* spp.

Les bactéries pectinolytiques des genres *Dickeya* et *Pectobacterium* ont une forme de bâtonnet (0,5-1 μm de diamètre sur 1-3 μm de longueur), sont anaérobies facultatives et munies de flagelles péritriches (Charkowski, 2006). Elles sont anciennement classées dans le genre *Erwinia* proposé

par Erwin Frink Smith (Winslow *et al.*, 1920). Ce dernier contenait les bactéries phytopathogènes de la famille des *Enterobacteriaceae* y compris les *Erwinia* qui causent des nécroses sur les organes des plantes hôtes. Les *Erwinia* pectinolytiques comprenaient trois espèces (ou sous-espèces) principales: *Erwinia carotovora* sous-espèce *carotovora* ; *Erwinia carotovora* sous-espèce *atroseptica* et *Erwinia chrysanthemi* (Dye, 1969). La nomenclature des bactéries pectinolytiques de la pomme de terre a subi plusieurs changements ces dernières décennies. Ce fait est dû principalement au développement des outils de caractérisation moléculaire, l'hétérogénéité et au faible nombre des souches disponibles pour certaines espèces d'entérobactéries, ce qui rend difficile la détermination de la parenté de ces taxons (Charkowski, 2006).

Après la première proposition de changement de nomenclature faite par Waldee en 1942 (Waldee, 1942), les travaux de phylogénie, basés sur les données d'hybridation ADN-ADN, la sérologie, les caractéristiques phénotypiques et biochimiques, ainsi que l'analyse des séquences ADN (Hauben *et al.*, 1998 ; Kwon *et al.*, 1997) ont conduit au passage des *Erwinia* pectinolytiques dans le genre des *Pectobacterium*. Cependant, avec l'avancée des outils de taxinomie, une analyse multi-locus des séquences d'ADN associée à de précédentes données d'hybridation ADN-ADN (Brenner *et al.*, 1973) ont abouti à une nouvelle nomenclature des *Pectobacterium*. *Pectobacterium chrysanthemi* et *Pectobacterium carotovorum* ont été ainsi reclassifiés dans deux genres différents, *Dickeya* (pour *P. chrysanthemi*) et *Pectobacterium* (pour *P. carotovorum*) (Samson *et al.*, 2005). De plus, les travaux de Gardan *et al.* (2003) ont abouti à la transformation de certaines sous-espèces de *Pectobacterium carotovorum* en espèces à part entière. En 2004, une nouvelle sous-espèce de *Pectobacterium*, nommée *P. carotovorum* sous-espèce *brasiliensis*, a été identifiée comme agent pathogène associé à la pomme de terre au Brésil (Duarte *et al.*, 2004) puis en Afrique du sud (van der Merwe *et al.*, 2010) et en Corée du sud (Lee *et al.*, 2014).

In fine, les *Erwinia* pectinolytiques sont maintenant reclassifiés en deux genres, *Dickeya* et *Pectobacterium*, comprenant plusieurs espèces et sous-espèces : *Pectobacterium atrosepticum* (Pa) (anciennement *Erwinia carotovora* sous-espèce *atroseptica*), *Pectobacterium carotovorum* sous-espèce *carotovorum* (Pcc) (anciennement *Erwinia carotovora* sous-espèce *carotovora*) et *Dickeya chrysanthemi* (anciennement *Erwinia chrysanthemi*). Les espèces de *Dickeya* sont très diversifiées (Charkowski, 2006) et elles ont été divisées en plusieurs pathovars et biovars (Nassar *et al.*, 1996). Les principales espèces associées à la pourriture de la pomme de terre sont *D. dianthicola*, *D. dadantii*, *D. paradisiaca*, *D. chrysanthemi* et *D. zea* (Samson *et al.*, 2005). Récemment, une nouvelle espèce émergente, nommée *D. solani* a été identifiée dans plusieurs pays européens (van der Wolf *et al.*, 2014). Le tableau I.1 ci-dessous résume les changements de nomenclature des espèces *Dickeya* et *Pectobacterium* associées à la maladie sur la pomme de terre (*Solanum tuberosum*).

Tableau I.1 Récapitulatif de la nomenclature des espèces de *Pectobacterium* et *Dickeya*. Modifié d'après Charkowski (2006)

Ancienne nomenclature	Nomenclature proposée	Hôte	Référence
<i>Erwinia carotovora</i>			
subsp. <i>atroseptica</i>	<i>Pectobacterium atrosepticum</i>	Pomme de terre, tomate	Gardan <i>et al.</i> (2003)
subsp. <i>betavasculorum</i>	<i>Pectobacterium betavasculorum</i>	Betterave, tournesol, pomme de terre, artichaut	Gardan <i>et al.</i> (2003)
subsp. <i>brasiliensis</i>	<i>Pectobacterium carotovorum</i> subsp. <i>brasiliense</i>	Pomme de terre	Duarte <i>et al.</i> (2004), Nabhan <i>et al.</i> (2012)
subsp. <i>carotovora</i>	<i>Pectobacterium carotovorum</i>	Plusieurs hôtes	Hauben <i>et al.</i> (1998)
subsp. <i>odorifera</i>	<i>Pectobacterium carotovorum</i> subsp. <i>odoriferum</i>	Chicorée, poireau, céleri	Hauben <i>et al.</i> (1998)
subsp. <i>wasabiae</i>	<i>Pectobacterium wasabiae</i>	Raifort	Gardan <i>et al.</i> (2003)
<i>E. chrysanthemi</i>			
pv. <i>dieffenbachia</i> (2, 3)	<i>Pectobacterium chrysanthemum</i> <i>Dickeya Dieffenbachiae</i> , <i>Dickeya dadantii</i>	Plusieur hotes	Samson <i>et al.</i> (2005)
pv. <i>dianthicola</i> (1, 7, 9)	<i>Dickeya dianthicola</i>	Plusieurs hôtes	Samson <i>et al.</i> (2005)
pv. <i>paradisiaca</i> (3)	<i>Dickeya paradisiaca</i>	Plusieurs hôtes	Samson <i>et al.</i> (2005)
pv. <i>zea</i> (3, 8)	<i>Dickeya zea</i>	Plusieurs hôtes	Samson <i>et al.</i> (2005)
pv. <i>chrysanthemi</i> (5, 6)	<i>Dickeya chrysanthemi</i>	Plusieurs hôtes	Samson <i>et al.</i> (2005)
pv. <i>parthenii</i> (5, 6)	<i>Dickeya chrysanthemi</i> <i>Dickeya solani</i> sp. nov.	Plusieurs hôtes Pomme de terre	Samson <i>et al.</i> (2005) Van der Wolf <i>et al.</i> (2014)

I.2. Épidémiologie et conditions environnementales

I.2.1. Distribution des *Pectobacterium* et *Dickeya* spp. dans l'environnement

Les bactéries pathogènes *Pectobacterium* et *Dickeya* sont trouvées dans différents environnements. Elles ont été isolées à partir de nombreuses plantes hôtes, sols, insectes ainsi que dans les eaux de surface et souterraines (Charkowski, 2006). Cependant, selon les espèces, elles possèdent une faible capacité à survivre en dehors de l'association avec les tissus de la plante hôte (De Boer, 2004).

Le tubercule de pomme de terre constitue, avec les autres plantes hôtes et les apports extérieurs (via l'eau d'irrigation, les insectes, le machinisme...), une source primaire de dissémination du pathogène, il assure la transmission de l'infection d'une culture à l'autre et du tubercule mère aux tubercules fils. De plus, tout autre organe infecté de la plante constitue aussi une source non négligeable d'inoculum après décomposition à la surface du sol (Czajkowski *et al.*, 2010, 2013a). En se basant sur ces observations, les agriculteurs producteurs de plants de pomme de terre mettent en place un schéma très strict pour le contrôle et la certification des plants afin de limiter la propagation de la maladie. Pour la jambe noire et selon les normes françaises, les plants ne doivent pas présenter plus de 1% de jambe noire pour la certification lors des inspections au champ. Dans le cas contraire, le lot sera déclassé (<http://plantdepommedeterre.org>).

Plusieurs études ont rapporté que la présence des *Dickeya* et *Pectobacterium* dans le sol et la rhizosphère était liée à la disponibilité des nutriments provenant de la dégradation des débris végétaux et des exsudats racinaires des plantes hôtes. En effet, la viabilité de ces micro-organismes diminue fortement dès qu'ils se retrouvent dissociés des tissus de leurs plantes hôtes (Hélias, 2008). La contamination de la rhizosphère par ces micro-organismes et leur dissémination se fait en partie par la réintroduction de tubercules infectés, par les débris de plants infectés ainsi que par les eaux de pluies. La survie des *Pectobacterium* et *Dickeya* dans le sol est également fortement influencée par les facteurs d'humidité et de température (Hélias, 2008 ; Pérombelon, 1992).

La propagation des *Pectobacterium* et *Dickeya* peut aussi impliquer des insectes (Basset *et al.*, 2000) et les nématodes (Perombelon et Kelman, 1980 ; Czajkowski *et al.*, 2011). Récemment, une étude a mis en évidence la capacité des *Pectobacterium* et *Dickeya* spp. à adhérer et coloniser le système digestif de *Caenorhabditis elegans*, suggérant une association de mutualisme favorisant la dissémination de ces pathogènes (Nykyri *et al.*, 2014).

L'eau joue un rôle important dans la transmission de ces pathogènes, plusieurs auteurs ont rapporté l'isolement des *Dickeya* et *Pectobacterium* des eaux de rivières, lacs et eaux souterraines (Powelson, 1985 ; Harrison *et al.*, 1987 ; Pérombelon et Hyman, 1989). Des *Erwinia* ont été également isolés à partir d'aérosols (Cappaert et Powelson, 1987).

1.2.2. Implication des facteurs environnementaux dans la maladie

Plusieurs auteurs ont rapporté la difficulté de reproduire les symptômes de la maladie de la jambe noire sous serre ou dans un champ d'expérimentation par simple inoculation du pathogène. Ainsi, le développement des symptômes au champ est fonction de plusieurs facteurs tels que l'humidité, la texture du sol et la température ainsi que selon le matériel végétal et la conduite culturale. L'effet de la température sur le développement de la maladie a été largement étudié, les *P. atrosepticum* ne peuvent tolérer des températures supérieures à 33 °C et sont principalement identifiés comme agents causals de la maladie au printemps, alors que *P. carotovorum* et *Dickeya* spp. sont préférentiellement retrouvés en été ou dans les régions chaudes (Elphinstone, 1987 ; Harrison *et al.*, 1987 ; Hélias, 2008 ; Mendonca, 1979 ; Perombelon et Kelman, 1980 ; Powelson, 1985). De plus, plusieurs études ont montré l'importance des variations de la température pour l'expression des facteurs de virulence chez les agents pathogènes *Pectobacterium* et *Dickeya*, tels que les pectates lyase, pectine lyase ainsi que pour la formation de biofilms (Hugouvieux-Cotte-Pattat *et al.*, 1996 ; Smadja *et al.*, 2004). La texture du sol est étroitement liée à la capacité de rétention de l'eau, elle influence le taux d'humidité du sol. Lors du stockage des tubercules humides ou dans un lieu à taux élevé d'humidité, la pourriture molle peut se développer sur les tubercules. De plus,

le taux d'oxygène influence également fortement le développement de la maladie, il est à l'origine des conditions d'anaérobiose qui se forment à la surface du tubercule favorisant ainsi l'initiation du processus de macération par le pathogène (Pérombelon *et al.*, 1989 ; Gill *et al.*, 2014). En effet, un local de stockage peu aéré va favoriser le développement de la maladie (Lund et Kelman, 1977 ; Molina et Harrison, 1980 ; Webb et Wood, 1974).

En général, la quantité de l'inoculum et l'interaction entre l'humidité du sol et la température déterminent l'incidence de la maladie et le type des symptômes sur la plante hôte. Par temps humide à la plantation et température élevée du sol, l'incidence de la maladie se manifeste par un manque à la levée parfois associé à l'apparition de la jambe noire. Par contre, une forte humidité lors de la plantation, accompagnée de températures faibles à la levée, accélèrent le développement de la maladie de la jambe noire avec parfois un dessèchement des plants infectés si un temps plus sec survient après (Elphinstone, 1987 ; Moh, 2012).

I.3. La lutte contre les bactéries pectinolytiques *Dickeya* et *Pectobacterium* spp.

A ce jour, il n'existe aucun moyen de lutte curatif contre les bactéries responsables des maladies de la jambe noire et de la pourriture molle de la pomme de terre. En effet, la lutte actuelle se base essentiellement sur l'utilisation de mesures prophylactiques visant à diminuer la quantité et la dissémination du pathogène via les semences et les machines agricoles, afin de limiter les dégâts enregistrés (Priou et Jouan, 1996). Plusieurs stratégies de lutte sont toutefois disponibles ou en cours d'étude comme la lutte chimique, la lutte génétique et la lutte biologique (Latour *et al.*, 2008 ; Czajkowski *et al.*, 2011).

I.3.1. Lutte chimique

En plus des mesures prophylactiques, les agriculteurs, ont parfois recours à l'utilisation de produits chimiques peu efficaces et extrêmement dommageables pour l'environnement. Ces produits chimiques sont à base de cuivre (Rousselle *et al.*, 1996) et ils sont fortement déconseillés et interdits ou en cours d'interdiction dans plusieurs pays (Ordax *et al.*, 2006). Comme alternative,

Mills *et al.* (2006) ont montré, par des essais *in planta*, que plusieurs composés chimiques tels que l'acetate d'aluminium et le benzoate de sodium ont une efficacité significative contre le développement de la pourriture molle sur les tubercules de pomme de terre au moment du stockage. De plus, des études menées en laboratoire ont montré la capacité de nouveaux composés chimiques à altérer l'intégrité cellulaire bactérienne ou à inhiber la croissance de ces pathogènes tels que les solutions de chlorure d'aluminium et de métabisulfite de sodium (Yaganza *et al.*, 2004) ou des peptides de synthèses (Kamysz *et al.*, 2005). Néanmoins, ces produits chimiques, malgré leur efficacité sur les pathogènes pectinolytiques, demeurent une menace non négligeable pour la communauté microbienne bénéfique, et risquent d'altérer les équilibres écologiques dans la rhizosphère (Latour *et al.*, 2008).

1.3.2. Lutte génétique

Jusqu'à présent aucune variété de pomme de terre (*Solanum tuberosum*) n'est complètement résistante aux *Pectobacterium* et *Dickeya* spp., cependant certaines variétés moins sensibles à ces infections existent (Lyon, 1989). Une étude menée sur un panel de 16 variétés, a montré que la variété *Kerpondy* était peu sensible, tandis que la variété *Ackergesegen* était très sensible aux pourritures molles engendrées par *P. atrosepticum* (Pasco, 2005 ; Latour *et al.*, 2008). Plusieurs gènes sont impliqués dans les mécanismes de défense chez *S. tuberosum*, ces mécanismes de défense ciblent la croissance et l'invasion du pathogène dans les tissus du végétal en réprimant l'expression des déterminants de virulence impliqués dans le pouvoir pathogène (Lyon, 1989). Lors des processus de sélection variétale, plusieurs critères sont recherchés en plus de la résistance aux pathogènes macergènes : les caractères améliorant le rendement, la conservation, les critères organoleptiques ou la teneur en féculé ou sucres réducteurs. Ainsi, l'obtention de plantes résistantes par les méthodes classiques de sélection variétale reste difficile et demande beaucoup de temps (Latour *et al.*, 2008). De plus, jusqu'à présent, la résistance aux pathogènes

macergènes prend un rang inférieure dans les programmes de sélection variétale comparée à d'autres maladies ou caractères de qualité (Czajkowski *et al.*, 2011).

Les plantes transgéniques constituent une alternative dans la lutte génétique par création de variétés de plants de pomme de terre ayant intégré du matériel génétique étranger (un gène ou plusieurs) qui leur confère une résistance vis à vis des pathogènes. Un gène originaire d'un bactériophage, codant pour la production d'un lysozyme qui dégrade la membrane de certaines bactéries, a été utilisé pour produire des plantes transgéniques produisant ce lysozyme et le transportant dans les espaces intercellulaires. Ces propriétés confèrent à la plante une résistance significative aux attaques de *Pectobacterium atrosepticum* (During *et al.*, 1993 ; Latour *et al.*, 2008). La production des enzymes pectolytiques bactériennes est en partie induite par la présence de composés de dégradation des polymères de polygalacturonate, provenant de la dégradation de la pectine végétale. Ces composés constituent une source d'énergie pour les pathogènes. Dans ce contexte, l'utilisation de plantes transgéniques de pomme de terre capables de produire une enzyme dégradant ces dérivés pourrait réduire le développement de la maladie en limitant les ressources trophiques de la bactérie (Czajkowski *et al.*, 2011). Ainsi, Wegener (2001) a montré que, lors des essais effectués *in vitro* et *in planta* sur des plantes transgéniques capables de produire une pectate lyase, la sensibilité aux infections par *Pectobacterium* est significativement diminuée.

Enfin, la lactonase est une metalloenzyme produite par certaines espèces bactériennes qui cible et inactive les signaux *quorum sensing* de type N-acyl-homoserines lactones impliqués dans la régulation de divers processus notamment les gènes de virulences chez *Pectobacterium* spp. (Dong *et al.*, 2000). Le clonage et l'expression d'un gène de lactonase (AHL lactonase) de *Bacillus* sp. dans des plants pomme de terre transgéniques a montré un niveau très élevé de résistance contre *Pectobacterium carotovorum* (Dong *et al.*, 2001). L'utilisation des plantes transgéniques

est subordonnée à la législation sur les OGMs, actuellement aucune variété de pomme de terre transgénique n'est autorisée en culture en Europe.

I.3.3. Lutte biologique

La lutte biologique représente une stratégie alternative aux luttes chimique et génétique contre les *Pectobacterium* et *Dickeya*. Cette méthode comprend l'utilisation d'antagonistes ciblant le pathogène soit directement soit via différents mécanismes tels que l'antibiose, la compétition pour les nutriments ou la stimulation des défenses de la plante hôte (Howarth, 1991). Le développement des méthodes de lutte biologique contre *Pectobacterium* et *Dickeya* est limité pour l'instant à des essais *in vitro* ou sur tubercules ; très peu d'études ont été menées au champ pour prouver la capacité des antagonistes à réduire l'incidence de la maladie (Czajkowski *et al.*, 2011).

Dans le cas de l'antibiose, l'antagoniste produit des substances antibactériennes capables de limiter le développement du pathogène (Fravel, 1988). Les *Pseudomonas* spp. constituent un groupe taxinomique de grand intérêt dans la lutte biologique contre les phytopathogènes en raison de leur capacité de développement dans la rhizosphère et à leur production d'une large variété de métabolites tels que des sidérophores et des antibiotiques (Compant *et al.*, 2005 ; Garbeva *et al.*, 2004 ; Kloepper *et al.*, 1980). Les sidérophores sont des chélateurs de fer qui permettent à la bactérie de s'approvisionner en fer. Les bactéries productrices de sidérophores entrent ainsi en compétition avec les phytopathogènes présents dans la rhizosphère (Neilands, 1995).

Plusieurs études *in vitro* ont montré la capacité des *Pseudomonas* spp. à limiter le développement des symptômes de jambe noire et de pourriture molle (Kastelein *et al.*, 1999). Des souches de *Pseudomonas* fluorescents appliquées à des tubercules ont été capables de réduire l'incidence de la pourriture molle et de la jambe noire (Kloepper, 1983 ; Rhodes et Logan, 1987 ; Xu et Gross, 1986). Le pouvoir protecteur de ces souches est attribué à la production du 2,4-diacetylphloroglucinol (DAPG), un métabolite antimicrobien, et de sidérophores (Latour *et al.*, 2008). La souche *P. fluorescens* F113 a été utilisée comme antagoniste contre *P. atrosepticum*,

des essais *in vitro* ont montré que cette souche était capable d'inhiber la croissance de *P. atrosepticum* alors que la souche mutée, incapable de produire le DAPG, n'affecte pas la croissance de *P. atrosepticum* (Cronin *et al.*, 2006).

À partir d'un crible de 10000 bactéries isolées du sol, 96 isolats ont été caractérisés comme présentant un effet d'antibiose *in vitro* sur des espèces de *Pectobacterium* et *Dickeya* et parmi ces isolats, 5 souches de *Pseudomonas* et 1 souche de *Bacillus* ont montré, en serre, une capacité à inhiber le développement des *Pectobacterium* et *Dickeya* seuls ou en combinaison (Raoul des Essarts, 2015).

Un autre groupe de bactéries potentiellement utilisable dans la lutte biologique existe. Il s'agit des bactéries lactiques capables de produire plusieurs composés tels que des acides organiques, les sidérophores ou le peroxyde d'hydrogène qui permettent d'inhiber la croissance des *Pectobacterium* et *Dickeya* spp. (Czajkowski *et al.*, 2011). *Bacillus subtilis* BS107 a été testé *in vitro* sur une large gamme de pathogènes bactériens de l'homme tels que *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Staphylococcus aureus*. Il a montré un effet d'antagonisme important contre ces pathogènes. Il a également montré une activité antagoniste vis à vis de plusieurs phytopathogènes y compris les *Pectobacterium* et *Dickeya* (Sharga et Lyon, 1998).

L'implication du système QS dans la modulation de l'expression des facteurs de virulence chez *Pectobacterium* et *Dickeya* spp. est à la base du développement de plusieurs stratégies de lutte biologique. Ces méthodes de lutte ciblent ce mécanisme de communication bactérien afin d'affaiblir la virulence de ces pathogènes. Chez l'espèce non-pathogène ubiquiste *Rhodococcus erythropolis*, plusieurs voies de dégradation des NAHL ont été identifiées (Barbey *et al.*, 2012 ; Uroz *et al.*, 2005, 2008). La stimulation de la croissance de *R. erythropolis* grâce à l'apport de gamma caprolactone permet une dégradation accrue des NAHL impliquées dans la régulation des facteurs de virulence et ainsi une diminution des symptômes de jambe noire et de pourriture molle (Cirou *et al.*, 2011, 2012 ; Crépin *et al.*, 2012). Récemment une étude a montré l'activité

antagoniste de *R. erythropolis* R138 *in planta* contre *P. atrosepticum* en mettant en évidence la dégradation des signaux NAHL de type N-3-oxo-octanoyl-L-HSL par la voie de dégradation des gamma lactones (Barbey *et al.*, 2013).

Une stratégie de lutte, en cours de développement, consiste à utiliser des bactériophages s'attaquant spécifiquement aux espèces de *Pectobacterium* et de *Dickeya*. Plusieurs travaux ont réalisé l'isolement et la caractérisation de bactériophages capables de limiter le développement des pathogènes de pomme de terre appartenant aux genres *Dickeya* et *Pectobacterium* (Adriaenssens *et al.*, 2012 ; Czajkowski *et al.*, 2013b ; Czajkowski *et al.*, 2014). Des essais *in vitro* avec les bactériophages vB_DsoM_LIMEstone1 et 2 ont montré une réduction des symptômes de pourriture molle causés par *D. solani*. Ainsi, lors des expérimentations au champ infecté par *D. solani*, le traitement par ces bactériophages a permis un rendement plus élevé (Adriaenssens *et al.*, 2012).

I.4. Outils de la génomique appliqué aux bactéries pectinolytiques

I.4.1. Technologies de séquençage

Le séquençage d'ADN constitue un outil d'une grande importance dans plusieurs domaines allant de l'archéologie, la génétique, la biotechnologie jusqu'aux sciences forensiques. Durant la dernière décennie, des progrès spectaculaires ont été réalisés grâce au développement des technologies de séquençage d'ADN à haut débit donnant un accès massif à l'information biologique codée par cette molécule.

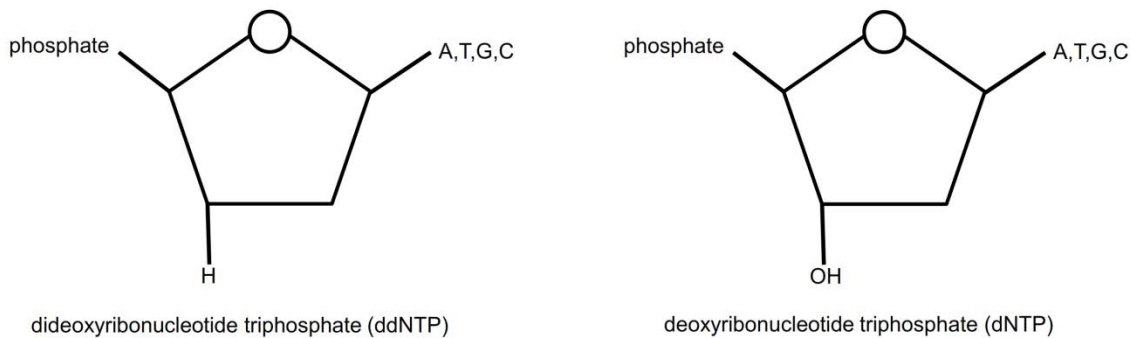
Dans cette partie, les technologies de séquençage les plus utilisées à l'heure actuelle et quelques exemples de leur apport en génomique des bactéries phytopathogènes seront détaillés.

I.4.1.1 Sanger (Première génération)

En 1975, Sanger et Coulson ont décrit la première méthode de séquençage d'ADN appelée "plus and minus". Cette méthode était basée sur l'utilisation d'une ADN polymérase I d'*E. coli* et d'une autre provenant du bactériophage T4 (Sanger et Coulson, 1975). Cette méthode étant peu

efficace, Sanger et ses collègues ont développé, 2 ans plus tard, une autre méthode basée sur le séquençage par polymérisation enzymatique (Sanger *et al.*, 1977). Elle consiste à initier la polymérisation de l'ADN à l'aide d'une amorce (courte séquence d'ADN) complémentaire à une région dans le brin à séquencer. L'ADN polymérase incorpore les dNTPs (désoxyribonucéotides triphosphate) correspondant au brin complémentaire jusqu'au moment de l'incorporation d'un terminateur d'élongation ddNTPs (didésoxyribonucéotides triphosphate) marqué par radioactivité.

Figure I.3 ddNTP et dNTP. L'absence du groupement hydroxyle sur le ddNTP bloque la progression de la synthèse du brin d'ADN.



Ces terminateurs d'élongation sont ajoutés au milieu réactionnel, ils sont différents des dNTPs par leur extrémité 3'OH. Sur la figure I.3, l'extrémité 3'OH des dNTPs est remplacée par une extrémité 3'H qui empêche la formation de la liaison phosphodiester entre le ddNTP incorporé dans la chaîne et le nucléotide suivant. L'allongement de la chaîne est alors interrompu et ce de façon aléatoire après chaque incorporation de ddNTP (França *et al.*, 2002).

À la fin de la réaction de séquençage, la migration des produits réactionnels sur gel de polyacrylamide à pouvoir résolutif élevé permet de différencier deux fragments d'ADN différents entre eux d'une seule paire de base. La séquence nucléotidique du brin séquencé se détermine par l'identification du ddNTP présent à l'extrémité 3' de chaque fragment.

Aujourd'hui, le principe de séquençage par méthode enzymatique est le même, mais la technologie a été automatisée en remplaçant les premières machines utilisant des gels de polyacrylamide par des séquenceurs à capillaires tel que ABI 3730 DNA Analyser de Life

technologies. Cette technologie a permis des progrès en génomique depuis sa commercialisation en 2002. En 2004, en utilisant le séquenceur à capillaire automatique ABI3700, dans le cadre d'un projet collaboratif, le génome de *P. atrosepticum* SCRI1043 a été séquencé, ce qui a permis la caractérisation des facteurs de virulence de cette espèce par comparaison aux pathogènes de plantes ou d'animaux de la famille des *Enterobacteriaceae* (Bell *et al.*, 2004). Sept ans plus tard, le génome de la souche modèle *Dickeya dadantii* 3937 a été séquencé par la même technologie (Glasner *et al.*, 2011).

I.4.1.2 Le séquençage de 2^{ème} génération

Les nouvelles technologies de séquençage dites NGS (*Next Generation Sequencing*), ont permis de séquencer de l'ADN plus rapidement avec une meilleure qualité tout en réduisant le coût qui est passé de plus de 1000 dollars pour un million de bases en 2001 à moins de 100 dollars en 2015 (<http://www.genome.gov/sequencingcosts/>).

I.4.1.2.1. Technologie 454

En 2005, Roche a mis sur le marché le premier séquenceur de deuxième génération appelé 454 Life technologie. La technologie 454 a été élaborée par Rothberg (Margulies *et al.*, 2005) et repose sur plusieurs techniques : PCR en émulsion, pyroséquençage et une technologie de pointe pour l'acquisition informatique des données de séquençage.

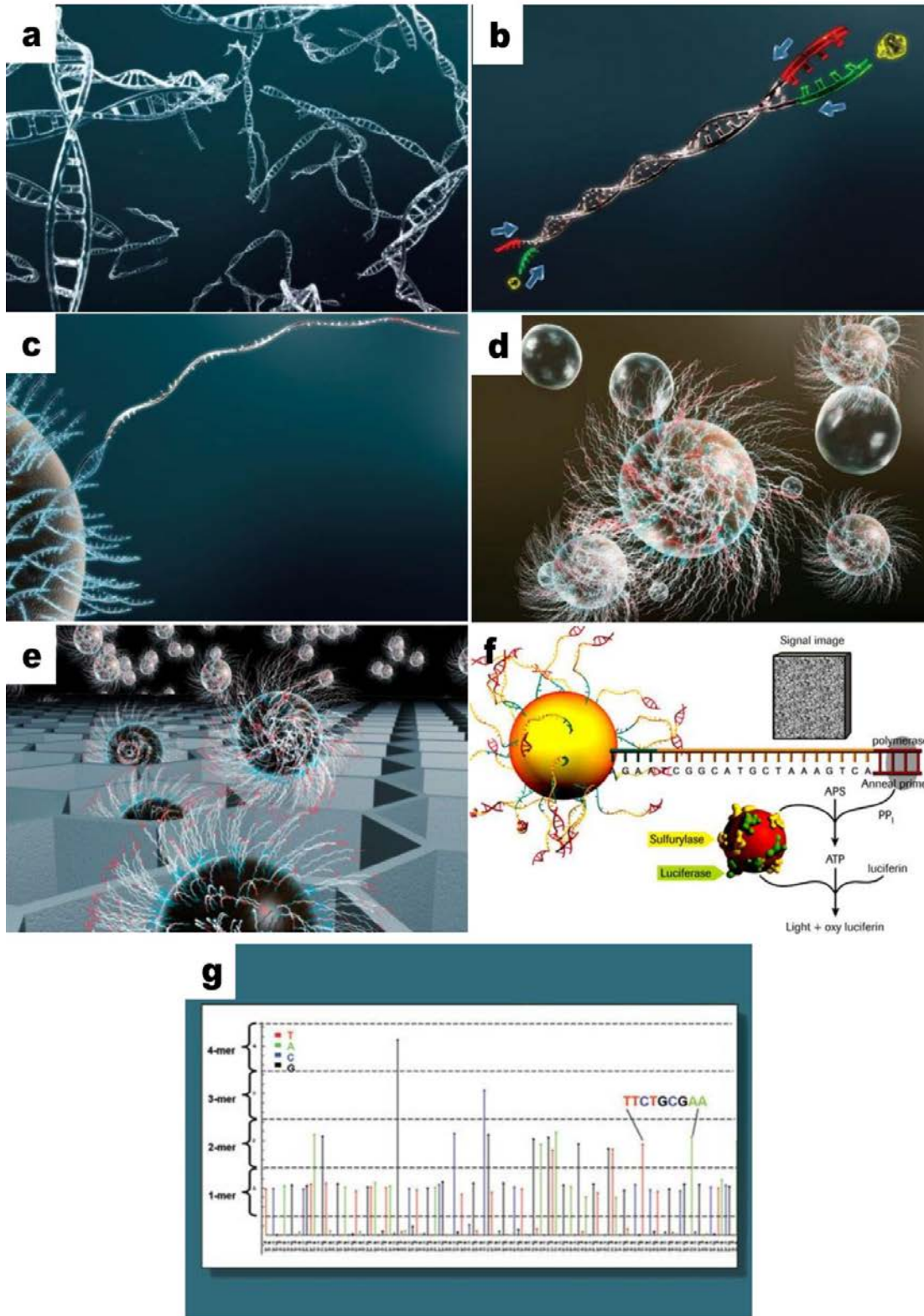
Les fragments d'ADN à séquencer sont d'abord amplifiés par PCR en émulsion dans des microgouttes renfermant une microbille d'agarose en phase aqueuse (Fig. I.4a, b, c, d). Ces microbilles sont ensuite déposées dans une plaque en fibre optique contenant 1.4 millions de puits qui possèdent un diamètre assurant le dépôt d'une microbille par puits (Fig. I.4e).

À l'échelle de la microplaque, le pyroséquençage se fait dans chaque puits en présence d'un mélange réactionnel contenant plusieurs enzymes et substrats. Les nucléotides sont ajoutés au milieu réactionnel l'un après l'autre. L'incorporation du nucléotide dans le brin par l'ADN polymérase (élongation) libère un pyrophosphate (PPi) : l'ATP sulfurylase transforme ce

pyrophosphate en ATP qui est alors couplé à une luciférine, par une luciférase. On a alors production d'oxyluciférine et d'un signal lumineux. Une autre enzyme du mélange réactionnel, l'apyrase, dégrade les nucléotides en surplus (Fig. 1.4f). Le signal lumineux libéré est capté par un capteur CCD (*Charge-Coupled Device*) qui le reproduit sous forme d'un pic sur le pyrogramme (El Fahime et Ennaji, 2012) (Fig. 1.4g). Après une dizaine d'heures de séquençage, on obtient environ 900 Mega bases de données avec une taille de lecture de 700 pb. La quantification des homopolymères (longues séquences d'ADN constituées par la succession de nucléotides identiques) est l'inconvénient principal de la technologie Roche 454.

En 2013, la société Roche a annoncé qu'elle prévoyait d'arrêter la commercialisation de la technologie 454 d'ici mi-2016 (www.genomeweb.com).

Figure 1.4 Présentation de la technologie 454. **a, b, c, d** : Fragmentation et amplification par PCR en émulsion des fragments d'ADN à séquencer. **e, f** : Pyroséquençage et acquisition de données. **g** : Pyrogramme. D'après le site de Roche (<http://www.454.com/>).



I.4.1.2.2. Technologie Illumina

La technologie Illumina de séquençage de l'ADN repose sur plusieurs étapes. Dans un premier temps, les librairies à séquencer sont construites par fragmentation de l'ADN et liaison d'adaptateurs à leurs extrémités. Ces adaptateurs vont permettre la fixation aléatoire de ces banques sur une surface de verre appelée *flow cell* (FC) par hybridation avec des amorces complémentaires fixées au FC. L'ADN polymérase vient ensuite pour synthétiser un nouveau brin complémentaire qui se libère par dénaturation thermique. L'ADN simple brin se lie à une amorce adjacente sur le FC pour former un pont. Le cycle d'amplification et dénaturation continue jusqu'à la formation d'un groupement d'ADN clonal dans une zone appelée cluster. L'étape suivante est le séquençage des centaines de milliers de clusters en parallèle (Sengenès, 2012). Le principe de séquençage utilisé dans cette technologie est basé sur l'incorporation réversible de nucléotides fluorescents suivie de la lecture optique de la fluorescence. Le terminateur contient un groupement (nitrophényle) qui permet la terminaison de la synthèse d'ADN. L'élimination de ce groupement terminateur attaché au nucléotide incorporé se fait par photoclivage utilisant la lumière ultraviolette. Ainsi le nucléotide incorporé devient fonctionnel permettant à l'ADN polymérase de continuer la synthèse (Li *et al.*, 2003). De plus, le séquençage se fait en temps réel en présence des 4 nucléotides en même temps, ce qui corrige le problème des homopolymères rencontré dans le cas de la technologie 454. Chaque cycle de séquençage comprend 3 étapes : l'incorporation, la détection de la fluorescence et la déprotection. Cette technologie de séquençage sur puce à ADN permet de séquencer plus de 100 000 paires de bases par seconde avec une taille de lecture allant de 34 jusqu'à 300 pb pour certaines machines (MiSeq Desktop Sequencer). Plusieurs appareils ont été mis sur le marché avec un ratio (coûts/rendement) inégalé, ce qui a donné lieu à une dominance de cette technologie dans les programmes et centres de séquençage du monde entier (Bentley, 2006).

I.4.1.3 Technologie PacBio (*Pacific biosciences*) : séquençage de 3^{ème} génération

Grâce à l'avancée des nanotechnologies, la troisième génération de séquençage est en cours d'émergence, elle est basée sur le séquençage de la molécule d'ADN de façon individuelle sans passer par l'étape d'amplification qui caractérise la deuxième génération de séquençage.

La technologie de séquençage PacBio, développée par l'entreprise Pacific Biosciences (Californie, USA), permet de séquencer une molécule d'ADN en temps réel (SMRT : *single molecule real time*) à l'aide d'un instrument, le PacBio RS. Le processus de séquençage se fait sur un support appelé "SMRT Cell" composé de dizaines de milliers de puits sous forme de structure détecteur appelé ZMW (*zero-mode waveguide*). Au fond de chaque puits, d'un diamètre de 100 nm, est fixée une seule polymérase. Ces structures ZMW permettent d'enregistrer en temps réel les dNTP incorporés où chacun des 4 nucléotides est lié à un fluochrome différent. Les signaux obtenus sont enregistrés et analysés par des méthodes informatiques de pointe (Eid *et al.*, 2009 ; Sengenès, 2012).

Pour l'assemblage des lectures issues du séquenceur PacBio, l'entreprise a développé plusieurs algorithmes d'assemblage adaptés selon la taille des lectures et la couverture du séquençage. Dans le cas des génomes bactériens, il s'agit de HGAP (*Hierarchical Genome Assembly Process*) qui comprend plusieurs étapes dans le processus d'assemblage (Chin *et al.*, 2013).

La vitesse de séquençage de cette technologie est de 10 bases par seconde avec une longueur moyenne de lecture de plus de 1000 pb. Malgré son succès, la technologie PacBio génère aussi des erreurs de séquençage liées à la quantification des homopolymères.

I.4.2. Assemblage des génomes

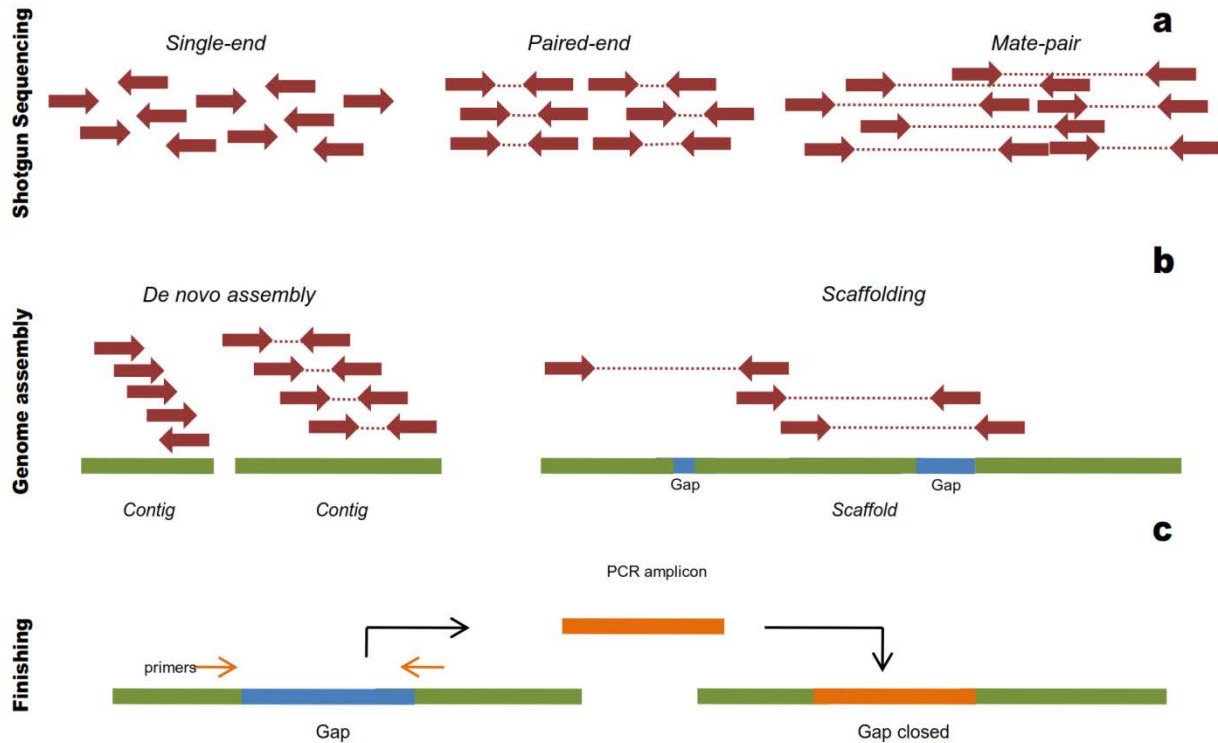
Le développement des technologies de séquençage et la chute vertigineuse des coûts ont augmenté le flux des données brutes de séquençage d'ADN et d'ADNc. Les plates-formes de séquençage telles que Roche 454 (Margulies *et al.*, 2005), Illumina (Bentley, 2006), ou ABI SOLID (Harris *et al.*, 2008) génèrent des millions de lectures sous forme de séquences de courte taille

(reads) allant de 50 jusqu'à 500 pb. Dans le cas du séquençage selon l'approche aléatoire globale dite *shotgun*, l'ADN génomique est fragmenté en petits morceaux de différentes tailles selon la librairie envisagée. Les librairies dites *paired-end* et *mate-pair* permettent le séquençage des deux extrémités des fragments d'ADN. Pour les premières, les fragments seront de taille inférieure à 1000 pb, tandis que pour les secondes les fragments auront une taille de plusieurs kilobases. Après la construction des librairies, le séquençage produit des lectures par paires des extrémités de fragments dont la taille des lectures obtenues sera dépendante de la technologie utilisée (Anderson, 1981 ; Bentley, 2006 ; Margulies *et al.*, 2005). La première étape dans l'assemblage du génome consiste à analyser les lectures de façon à reconstruire de longues séquences via les chevauchements possibles entre lectures (Fig. 1.5a). Dans ce contexte, plusieurs algorithmes et logiciels ont été développés pour réaliser ces assemblages génomiques. Les algorithmes d'assemblage envisagent des alignements entre les lectures pour trouver les chevauchements afin de créer de grandes séquences de plusieurs kilobases appelées *contigs* (Fig. 1.5b). Ce processus s'appelle l'assemblage *de novo* (Ekblom et Wolf, 2014). Par ailleurs, plusieurs stratégies d'assemblage *de novo* ont été optimisées selon le nombre de lectures et le taux de couverture ainsi que la taille du génome étudié (Miller *et al.*, 2010). Etant donné que le séquençage *shotgun* est aléatoire, l'orientation et l'ordre des *contigs* générés sont inconnus. La seconde étape dans l'assemblage génomique est le *scaffolding* (Fig. 1.5b) qui consiste à déterminer l'ordre et l'orientation des *contigs* ainsi que la taille de l'espace qui les sépare (*gaps*). Les *contigs* groupés dans le bon ordre et la bonne orientation forment ce que l'on appelle le *scaffold* qui peut contenir des *gaps* de taille approximativement déterminée par le logiciel de *scaffolding* utilisé.

Durant le processus d'assemblage des génomes, plusieurs critères et indices sont à prendre en considération pour évaluer la qualité de l'assemblage tels que le N50¹ des *contigs* et *scaffolds*, la couverture, le taux d'erreurs d'assemblage, le nombre de *contigs* et *scaffolds*, et la taille totale de l'assemblage (Ekblom et Wolf, 2014). L'assemblage ne permet jamais de reconstituer le génome d'un organisme à 100%, il reste toujours des *gaps* au sein des *scaffolds*, représentés par des "N". Ces *gaps* sont dus à des régions d'ADN répétées qui ne permettent pas aux logiciels d'assemblage de leur attribuer une localisation. Le *finishing* constitue la dernière étape dans l'assemblage génomique et permet d'éliminer ces *gaps*. L'amplification par PCR de ces *gaps* constitue une alternative pour le *finishing* du génome. Sur les jonctions des *contigs* adjacents formant le *gap*, on peut dessiner des amorces de façon à amplifier ces régions manquantes par PCR et ensuite les séquencer pour combler le *gap* (Fig. 1.5c). Un autre moyen est d'utiliser des logiciels développés pour réaliser le *finishing in silico* comme par exemple, GapFiller (BaseClear, Netherlands) qui utilise l'information de distance dans les lectures pairés pour le comblement des *gaps* au sein des *scaffolds* (Boetzer et Pirovano, 2012).

¹Le N50 permet d'évaluer la fragmentation du génome assemblé en contigs ou en scaffolds. Les contigs/scaffold sont triés par taille et on somme les contigs/scaffold un par un du plus grand au plus petit, le N50 est la longueur du dernier contig/scaffold obtenu lorsque tous les contigs/scaffolds déjà parcourus couvrent la moitié de la taille du génome. Plus le N50 est élevé, plus la qualité de l'assemblage est bonne.

Figure I.5 Résumé des étapes d'assemblage d'un génome. **a** : Les différents types de bibliothèques. **b** : Assemblage de novo et scaffolding. **c** : Etape de finishing par PCR.



I.5. Apport de la génomique comparative à l'étude des *Pectobacterium* et *Dickeya* spp.

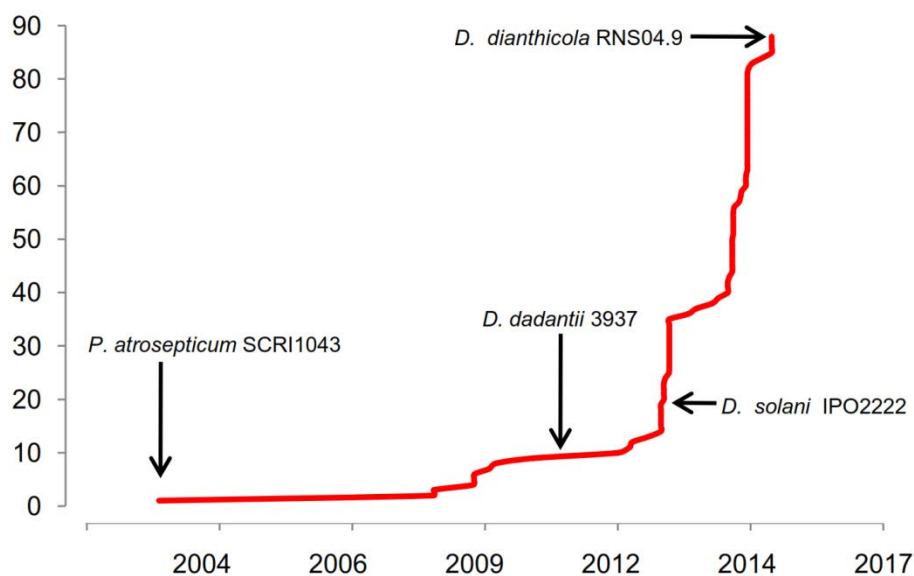
I.5.1. Génomique et taxinomie

En cinq décennies, les *Pectobacterium* et *Dickeya* spp. ont subi plusieurs changements de nomenclature (cf. sous-chapitre I.1.2). Ces phytopathogènes se distinguent entre eux principalement dans l'arbre phylogénétique fait à partir des séquences de gènes uniques tels que ARN 16S (Brown *et al.*, 2000 ; Charkowski, 2006 ; Gardan, 2003 ; Hauben *et al.*, 1998 ; Kwon *et al.*, 1997 ; Samson *et al.*, 2005). Avec l'émergence de la génomique, plusieurs outils alternatifs de classification taxinomique ont été développés, basés sur l'information contenue dans la séquence génomique. Une approche de taxinomie génomique pourrait contribuer à la caractérisation de marqueurs génétiques spécifiques aux espèces bactériennes des genres *Pectobacterium* et *Dickeya*. Par ailleurs, aujourd'hui on compte plus de 80 génomes d'espèces de *Dickeya* et *Pectobacterium* disponibles sur NCBI (Fig. I.6). Ces génomes constituent une base de données importante qui pourrait être utile pour la taxinomie en utilisant les outils tels que l'analyse de

séquences multilocus (MLSA : *Multilocus sequence analysis*), la moyenne d'identité nucléotidique (ANI : *average nucleotide identity*) l'identité des acides aminés (AAI : *amino acid identity*) et les signatures génomiques (Konstantinidis et Tiedje, 2005 ; Goris *et al.*, 2007). Ainsi, plusieurs études ont rapporté l'utilisation et la fiabilité de ces outils dans leurs analyses de ces phytopathogènes (Nykyri *et al.*, 2012 ; Pédrón *et al.*, 2014 ; van der Wolf *et al.*, 2014). Nykyri et ses collègues ont révisé la phylogénie de la souche *Pectobacterium* SCC3193 identifiée en premier comme *Pectobacterium carotovorum* (Pirhonen *et al.*, 1988) par utilisation d'une MLSA étendue sur 51 séquences protéiques et une analyse de protéomique comparative *in silico*. Ces analyses ont confirmé que la souche de *Pectobacterium* SCC3193 appartenait à l'espèce *Pectobacterium wasabiae* (Nykyri *et al.*, 2012). Par approche de taxinomie génomique, Naushad *et al.* (2014) ont identifié des signatures génomiques conservées permettant de distinguer les genres *Pectobacterium*, *Dickeya* et *Brenneria*. Ainsi, ces marqueurs moléculaires indiquent que ce groupe de bactéries constitue un clade à part au sein de l'ordre des *Entrobacterales* et incitent à une reclassification de ces genres dans une nouvelle famille des *Pectobacteriaceae* (Naushad *et al.*, 2014).

La génomique des *Dickeya* et *Pectobacterium* présente une grande importance dans la reconstruction des phylogénies de ces pathogènes ainsi que dans l'optique de développer des marqueurs genre spécifique et espèce spécifique pour le diagnostic et l'identification rapide de ces phytopathogènes de grande importance économique (Charkowski *et al.*, 2014).

Figure I.6 Evolution du nombre de génomes de *Pectobacterium* et *Dickeya* spp. déposés sur NCBI entre 2003 et 2015 (<http://www.ncbi.nlm.nih.gov/genome>).



I.5.2. Contribution de la génomique à l'identification et à la détection des *Pectobacterium* et *Dickeya* spp.

La transmission des *Pectobacterium* et *Dickeya* spp. via les semences limite les échanges de matériels biologiques en provoquant un impact sur les échanges économiques à travers le monde (Pérombelon, 2002). En effet, le développement des outils de diagnostic rapides et puissants constitue un enjeu majeur dans l'objectif de la lutte et le contrôle de la dissémination de ces pathogènes. Les méthodes classiques de détection et d'identification des *Pectobacterium* et *Dickeya* sont laborieuses, elles dépendent de l'isolement de cellules bactériennes viables sur des milieux gélosés semi-sélectifs suivi par des tests biochimiques, sérologiques et microscopiques permettant la révélation de ces pathogènes. Par ailleurs, le développement des outils moléculaires basés sur l'amplification d'ADN par PCR qualitative ou quantitative, a permis la caractérisation de plusieurs amorces spécifiques en se basant sur des gènes uniques tels que, *pelADE*, *fliC* ou 16S rDNA (Nassar *et al.*, 1996 ; Van Vaerenbergh *et al.*, 2012 ; Kwon *et al.*, 1997 ; Kettani-Halabi *et al.*, 2013). Ces amorces sont soit genre-spécifique soit espèce-spécifique (Tableau I.2).

En outre de leur intérêt dans la taxinomie et les analyses phylogénétiques, les séquences génomiques des *Pectobacterium* et *Dickeya* spp. revêtent une grande importance dans le sens de l'identification de marqueurs moléculaires plus spécifiques et plus fiables pour la détection de ces pathogènes et la certification des semences. Par utilisation d'approches bioinformatiques, Pritchard *et al.*, (2013) ont utilisé 17 *draft-genomes* et 4 génomes complets pour générer plus de 700 amorces discriminant les espèces de *Dickeya*. Les approches de génomique comparative permettent de comparer les génomes de *Pectobacterium* et *Dickeya* et d'identifier des clusters spécifiques aux espèces/genres ou encore à l'échelle de la souche (Garlant *et al.*, 2013 ; Khayi *et al.*, 2015 ; Pédrón *et al.*, 2014). Ces clusters spécifiques peuvent être utilisés pour déterminer des amorces spécifiques et obtenir une meilleure sensibilité de détection de ces pathogènes. Au sein de notre équipe, plus de 40 souches de *Pectobacterium* et *Dickeya* ont été séquencées (Khayi *et al.*, 2014 ; Kwasiborski *et al.*, 2013 ; Raoul des Essarts *et al.*, 2015). Par des approches de *mapping in silico*, des régions souche et espèce spécifiques ont été identifiées. Une dizaine d'amorces par souche/espèce ont été dessinées sur ces régions. Par PCR, ces primers ont été testés pour leur spécificité sur une trentaine de souches de *Pectobacterium* et *Dickeya*. Ensuite, les meilleurs candidats ont été évalués pour leur sensibilité par PCR en temps réel (J. Cigna, communication personnelle). La génomique des bactéries pectinolytiques a fortement contribué à l'identification de marqueurs moléculaires de diagnostic permettant une veille épidémiologique plus efficace contre ces phytopathogènes (Czajkowski *et al.*, 2015).

Tableau I.2 Amorces utilisées pour la détection et l'identification des différentes espèces de *Pectobacterium* et *Dickeya* spp. par PCR. Modifié d'après Czajkowski *et al.* (2015).

Espèce	PCR	Amorces	Références
<i>Dickeya</i> spp. et <i>Pectobacterium</i> spp.			
	PCR conventionnelle	SR3F GGTGCAAGCGTTAATCGGAATG SR1cR AGACTCTAGCCTGTCAGTTTT	Toth <i>et al.</i> (1999)
<i>Dickeya</i> spp.	PCR conventionnelle	5A GCGGTTGTTACCAGGTGTTTT 5B ATGCACGCTACCTGGAAGTAT ADE1 GATCAGAAAGCCCGCAGCCAGAT ADE2 CTGTGGCCGATCAGGATGGTTTTGTCGTGC	Chao <i>et al.</i> (2006) Nassar <i>et al.</i> (1996)
	PCR en temps réel	Df AGAGTCAAAGCGTCTTG Dr TTTCACCCACCGTCAGTC	Laurila <i>et al.</i> (2010)
	PCR en temps réel (TaqMan)	EChf GAGTCAAAAAGCGTCTTGCGAA EChr CCCTGTTACCGCCGTGAA Probe ECH CTGACAAGTGATGCCCTTCGTCTAGAGG	Pritchard <i>et al.</i> (2013)
	PCR en temps réel (TaqMan)	DIA-Af GGCCGCCTGAATACTACATT DIA-Ar TGGTATCTCTACGCCATCA Probe ATTAACGGCGTCAACCCGGC DIA-Cf CCAACGATTAGTCGGATCT DIA-Cr TAGTTGGTGCCAGGTTGGTA Probe DIA-C TCGACGTATGGGACGGTCCG	Pritchard <i>et al.</i> (2012)
<i>Dickeya solani</i>	PCR en temps réel (TaqMan)	SOL-Cf GCCTACACCATCAGGGCTAT SOL-Cr AACTACAGCGCGCATAAAC Probe Sol-C CCAGGCCGTGCTCGAAATCC SOL-Df GCCTACACCATCAGGGCTAT SOL-Dr CACTACAGCGCGCATAACT Probe SOL-D CCAGGCCGTGCTCGAAATCC	Pritchard <i>et al.</i> (2012)
	PCR en temps réel (TaqMan)	dsf GCGAACTTCAACGGTAAA dsr CAGAGCTACCAACAGAGA Probe CTCTGCTGGACGGTTC	Van Vaerenbergh <i>et al.</i> (2012)
	PCR conventionnelle	Y1 TTACCGGACGCCGAGCTGTGGCGT Y2 CAGGAAGATGTCGTTATCGCGAGT	Darrasse <i>et al.</i> (1994)
	PCR conventionnelle	Y45 TCACCGGACGCCAACTGTGGCGT Y46 TCGCCAACGTTACGAGAACAAGT Eca1 CGGCATCATAAAAACAGC Eca2 GCACACTTCATCCAGCGA PEAF CCGGTAATCAAGCTAATCC PEAR CCTTACCTATCGCTTCTCCT	Frechon <i>et al.</i> (1998) (De Boer et Ward, 1995) Park <i>et al.</i> (2006)
<i>Pectobacterium</i> c. subsp. <i>brasiliensis</i>	PCR conventionnelle	BR1f GCGTGCCGGGTTTATGACCT L1r CA(A/G)GGCATCCACCGT	Duarte <i>et al.</i> (2004)
Pcc et Pwa	PCR conventionnelle Nested PCR	Expccf GAACTTCGCACCCGCCGACCTTCTA Expccr GCCGTAATTGCCCTACCTGCTTAAG INPCCF GGCCAAGCAGTGCCTGTATATCC INPCCR TTCGATCAGCAACCTGCATTACT	Kang <i>et al.</i> (2003)
Pwa	PCR conventionnelle	Contig 1F CCTGCTGGCGTGGGGTATCG Contig 1R TTGCGGAAGATGTCGTGAGTGCG Contig 3F GCATTGACCAGTTTCGCCAGTTAC Contig 3R CTTTTTGAGCAGCGGGGTTGTG	de Haan <i>et al.</i> (2008)
	PCR conventionnelle	PW7011F CTATGACGCTCGCGGTTGCTGTT	Kim <i>et al.</i> (2012)
	PCR en temps réel	PW7011R CGCGCGCTCGTAGTGGAAAGTC	
	PCR conventionnelle multiplex	ERWFOR ACGCATGAAATCGGCCATGC ATROREV ATCGATAATTGATTGTCCT CHRREV AGTGCTGCCGTACAGCACGT	Smid <i>et al.</i> (1995)
Pa, <i>Dickeya</i> spp.	Multiplex	Y45 TCACCGGACGCCAACTGTGGCGT Y46 TCGCCAACGTTACGAGAACAAGT Ech1 TGGCGCGTCAGGAAGTTTAT Ech1' TCACCGGTCAGGGTGAAGTT	Diallo <i>et al.</i> (2009)
	Multiplex	Pca for GATCGGCATCATAAAAACACG Pca rev CGCACACTTCATCCAGCGAG Dcd Forw GAAAGCCCGCAGCCAGATC Dcd Rev TCAGGATGGTTTTGTCATGC	Peters <i>et al.</i> (2007)

I.5.3. Génomique et évolution des *Pectobacterium* et *Dickeya* spp.

L'adaptation des bactéries à leur niche écologique constitue un élément majeur qui confère à l'évolution de leur génomes (Altermann, 2012). L'évolution naturelle des bactéries est un processus complexe qui se fait par plusieurs mécanismes tels que les mutations, les réarrangements génétiques ou par les événements de transferts horizontaux de gènes (Juhas *et al.*, 2009). Le transfert horizontal de gènes joue un rôle important dans l'évolution des bactéries, il permet l'acquisition de nouvelles fonctions ou de nouveaux caractères phénotypiques (Smith *et al.*, 1992, 1993 ; Koonin *et al.*, 2001). Il est lié plus à la coexistence des espèces au sein d'une même niche écologique qu'au rapprochement taxinomique entre les bactéries (Brochier *et al.*, 2004). Dans le cas de *Pectobacterium* et *Dickeya*, la connaissance et la description de l'évolution génétique au sein de ce groupe bactérien s'avère d'une grande importance sur le plan de la compréhension de leurs mécanismes d'émergence et de pathogénie et de la lutte contre ces phytopathogènes. Les approches de génomique comparative appliquée aux *Pectobacterium* et *Dickeya* permettraient la description de l'ensemble des gènes propres à chacune des espèces qu'on appelle le core-génome (Medini *et al.*, 2005), l'identification d'acquisitions et de pertes de gènes ou de clusters de gènes ainsi que les événements de transfert horizontal qui peuvent se produire entre les espèces bactériennes. De plus, la comparaison des séquences génomiques deux à deux permet la révélation des variants nucléotidiques (SNP/InDel) affectant la séquence génomique au cours de l'évolution, ces variations nucléotidiques peuvent être associées à des variations phénotypiques en réponse au changement des conditions qui entourent la niche écologique (Lawrence et Hendrickson, 2005).

Les études préliminaires ont rapporté que le génome des entérobactéries *Dickeya* et *Pectobacterium*, ne dépasse pas 5 Mega bases et que le core-génome est largement conservé au sein de chaque espèce avec des différences enregistrées dans la contenance en îlots génomiques. Par ailleurs, peu de génomes de *Pectobacterium* contiennent des éléments

plasmiques (Kwasiborski *et al.*, 2013), quoique Nomura *et al* (1996) ont indiqué la présence de petits plasmides et éléments extra-chromosomiques dans certains *Pectobacterium*.

La souche SCC3193, a été reclassifiée en *P. wasabiae* après une étude génomique qui a révélé la présence d'îlots génomiques qui pourraient avoir été acquis par transfert horizontal. Parmi ces îlots génomiques, 15 étaient spécifiques à l'espèce *P. wasabiae*, dont des clusters de gènes qui pourraient être impliqués dans la virulence de ce pathogène (Nykyri *et al.*, 2012). L'évolution des *Pectobacterium* et *Dickeya* pourrait être liée au mode de vie de ces bactéries comprenant les différentes plantes hôtes et niches écologiques (Koonin *et al.*, 2001). Ainsi Glasner *et al.* (2008) ont effectué une comparaison génomique de 3 *Pectobacterium* ayant des variations en termes de distribution géographique et de gammes d'hôtes. Ils ont mis en évidence la conservation du core-génome, qui représente environ 80% de la taille du génome, entre les 3 espèces *P. atrosepticum* SCRI1043, *P. carotovorum* WPP14, et *P. brasiliensis* 1692. Ainsi, le T3SS était conservé dans les 3 espèces. L'analyse a montré que la variabilité liée à la gamme d'hôtes est associée aux gènes du système de sécrétion de type 4 (T4SS), de phytotoxines, de gènes de mobilité et de protéines de surface (Glasner *et al.*, 2008).

1.5.4. Génomique et virulence des *Pectobacterium* et *Dickeya* spp.

La disponibilité de plusieurs génomes de *Pectobacterium* et *Dickeya* dans les bases de données a contribué à l'amélioration des connaissances à propos des processus de pathogénie chez ce groupe de pathogènes (Charkowski *et al.*, 2014). Suite à la publication du premier génome de *P. atrosepticum* en 2004, une analyse de génomique comparative extensive avec une dizaine de séquences génomiques de bactéries pathogènes des *Enterobacteriaceae* a été effectuée. Il a été montré pour la première fois que la structure génomique de *P. atrosepticum* était la même que celle des bactéries pathogènes de la même famille, mais que ce pathogène avait également acquis, par transfert horizontal, 17 îlots génomiques qui comprennent des gènes impliqués dans l'interaction hôte-pathogène. Cette analyse a permis de caractériser les déterminants génétiques

de virulence chez *P. atrosepticum* tels que les systèmes de sécrétion, les phytotoxines et les enzymes de macération (Fig. I.7) (Bell *et al.*, 2004 ; Toth *et al.*, 2006). Ensuite, plusieurs études ont ciblé les mécanismes d'adhésion (Jahn *et al.*, 2011 ; Pérez-Mendoza *et al.*, 2011) ainsi que la synthèse des pigments (Williamson *et al.*, 2010). De plus, par approches de génomique et transcriptomique, plusieurs études ont rapporté l'implication de gènes, conservés au sein des *Dickeya* et *Pectobacterium* spp., dans l'interaction pathogènes-insectes vecteurs (Acosta Muniz *et al.*, 2007 ; Basset *et al.*, 2003 ; Charkowski *et al.*, 2014 ; Costechareyre *et al.*, 2010, 2013 ; Quevillon-Cheruel *et al.*, 2009).

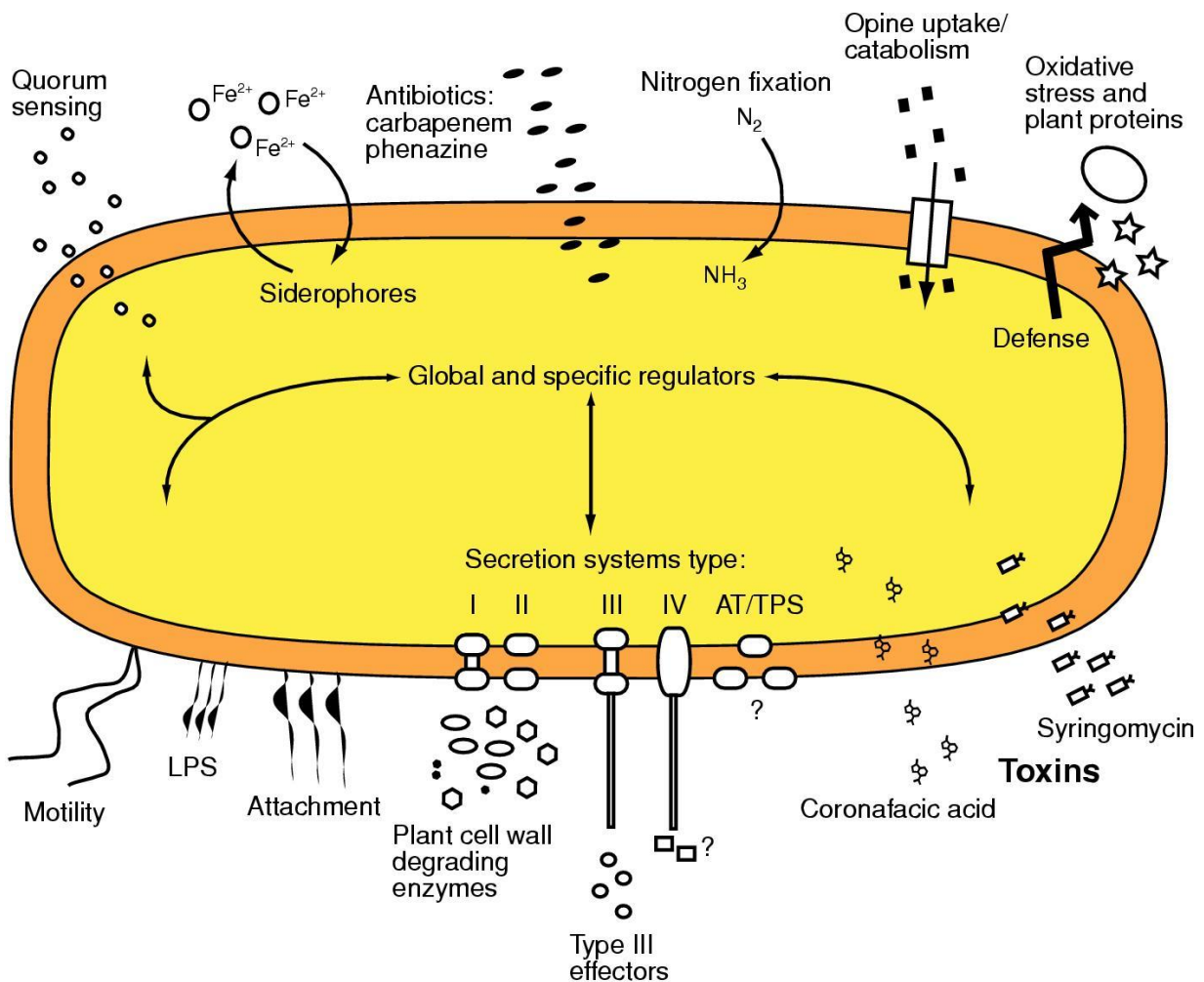
Aujourd'hui, il a été démontré que les *Pectobacterium* et *Dickeya* possèdent les 6 systèmes de sécrétion décrits chez les *Enterobacteriaceae* (Bell *et al.*, 2004 ; Charkowski *et al.*, 2012 ; Glasner *et al.*, 2008, 2011) à l'exception de *P. wasabiae* qui n'a pas de système de sécrétion de type 3 (Kim *et al.*, 2009a ; Nykyri *et al.*, 2012 ; Pitman *et al.*, 2010). Ces systèmes de sécrétion ont été largement étudiés pour leur implication dans les processus de pathogénie et d'interactions hôtes-pathogènes (Charkowski *et al.*, 2012).

En outre, l'utilisation des séquences génomiques a facilité l'élaboration et le développement des expériences s'intéressant à l'étude d'expression des gènes de virulence chez ces pathogènes (Rodionov *et al.*, 2004 ; Venkatesh *et al.*, 2006 ; Hommais *et al.*, 2008 ; Liu *et al.*, 2008 ; Yap *et al.*, 2008 ; Kõiv *et al.*, 2013 ; Monson *et al.*, 2013).

En revanche, très peu d'études de transcriptomique ont été menées chez *Pectobacterium* et *Dickeya* pour évaluer les variations d'expression de gènes dans l'objectif d'élucider les mécanismes de réponses des bactéries à une condition de stress donnée (Chapelle *et al.*, 2015 ; Jiang *et al.*, 2015). Dans une étude transcriptomique s'intéressant à l'interaction entre *D. dandantii* et l'aphide *Acyrtosiphon pisum*, on observe une surexpression des gènes impliqués dans la détoxification des éléments antimicrobiens qui apparaît utile pour le système immunitaire des *Acyrtosiphon pisum* (Costechareyre *et al.*, 2013).

La comparaison des transcriptomes de *D. solani* et *D. dianthicola* en conditions de macération ou de culture pure a révélé une différence, en termes de gènes surexprimés et réprimés, dans les deux conditions pour ces deux espèces. Certains gènes sont connus comme étant impliqués dans la virulence, tels que les gènes codant pour les enzymes de macération, de transporteurs membranaires et les gènes du métabolisme des sucres (Raoul des Essarts, 2015).

Figure I.7 Schéma simplifié des différents déterminants de virulence connus chez *Pectobacterium atrosepticum*. D'après Toth et al. (2006).



Chapitre II. Genome sequence of the emerging plant pathogen *Dickeya solani* strain RNS 08.23.3.1A

Dickeya solani est une nouvelle espèce émergente dans plusieurs pays européens responsable de la maladie de la jambe noire et de la pourriture molle de pomme de terre, en culture ou lors du stockage des tubercules (Laurila *et al.*, 2008 ; Sławiak *et al.*, 2009 ; Tsrör *et al.*, 2009 ; van der Wolf *et al.*, 2014). Elle a été également isolée à partir de plantes ornementales (Toth *et al.*, 2011). C'est une espèce capable d'envahir la plante hôte en utilisant le système vasculaire et de causer la maladie à partir d'un faible inoculum (Czajkowski *et al.*, 2010).

Au début de mes travaux, aucun génome complet de cette espèce n'était disponible dans les bases de données. Dans ce contexte, nous avons envisagé de séquencer le génome de la souche *D. solani* 3337 isolée en France en 2008. L'objectif était d'avoir un génome de bonne qualité permettant de mieux caractériser cette espèce émergente. Pour ce faire, nous avons utilisé deux technologies de séquençage Illumina et PacBio. Le séquençage Illumina en utilisant deux librairies *mate-pair* et *paired-end* nous a permis d'assembler un premier *draft genome* qui a été déposé dans GenBank (AMY101000000). Le séquençage PacBio nous a permis de confirmer le *scaffolding* qui a été fait par le séquençage illumina et aussi d'améliorer la séquence du *draft genome* en fermant les *gaps* au sein des *scaffolds*. Au final, nous avons pu assembler la séquence de *D. solani* 3337, sous forme d'un seul chromosome circulaire de 4 922 460 bp, qui a été annoté sous le serveur automatique RAST.

Ce chapitre est présenté sous la forme d'un article publié dans *Genome Annoucement* (Khayati *et al.*, 2014), cette version modifiée inclut la méthodologie et le résultat du séquençage PacBio de la souche *D. solani* 3337, publié dans BMC genomics (Khayati *et al.*, 2015). Dans cet article, j'ai contribué à l'assemblage génomique en utilisant les différentes méthodologies bioinformatiques et outils (logiciels et machines) disponibles au laboratoire avec l'aide de S. Mondy.

II.1. Abstract

Here we present the genome of *Dickeya solani* strain RNS08.23.3.1A isolated from *Solanum tuberosum*. *Dickeya solani*, recently described on potato cultures in Europe, is a proposed new taxon closely related to the *Dickeya dianthicola* and *D. dadantii* species

II.2. Introduction

Bacteria belonging to the *Dickeya* Genus cause soft rot disease on a wide range of plants, particularly on ornamentals (Dianthus, Dahlia, Kalanchoe, Pelargonium, Chrysanthemum, Parthenium, Dieffenbachiae, Saintpaulia) and crops (chicory, banana, sunflower, rice, artichoke, pineapple, tomato, banana, maize, and potato) (1,2). Infection is characterized by the maceration of plant tissues due to degradation of pectin, the major component of primary cell walls in plants (3). The resulting symptoms on potato stems are named 'blackleg' and 'soft rot' on tubers. On this plant host, disease symptoms caused by *Dickeya* spp. are similar to those caused by *Pectobacterium atrosepticum*. Three main *Dickeya* species, *D. dadantii*, *D. zea* and *D. dianthicola* were previously described on potato. A new taxon, tentatively named *D. solani*, had been recently identified largely in Europe and beyond (4). The *D. solani* strain RNS 08.23.3.1A (= PRI3337) has been isolated from potato in France in 2008 (5).

This genome sequence was sequence first by Illumina technology then, in order to upgrade the complete sequence, a PacBio sequencing technology was performed on the genomic DNA of this strain.

II.3. Illumina sequencing and assembly

Here we report the de novo genome assembly of *D. solani* strain RNS 08.23.3.1A using Illumina HiSeq 2000 v3 technology. Two libraries were constructed using the TruSeq (TM) SBS v3 sequencing kit: a shotgun (SG) paired-end library with a fragment size of 150-500 bp and a long jumping distance (LJD) mate-pair library with an insert size of 6,000 bp. The two libraries were sequenced using 2x100 bp paired-end read module by Eurofins Genomics (France). Assembly of

the sequences was carried out by the software CLC Genomics Workbench (V5.1) from CLC bio. Sequence reads with low quality (limit 0.05), ambiguous nucleotides ($n \leq 2$), and sequence length less than 30 nucleotides were dropped for assembly. In total 5,682,625 mate-pair reads and 52 496 166 paired-end reads were obtained corresponding to 419,377,725 bases with an average length of 73.8 bp for the mate-pairs and 4,903,141,904 bases with an average length of 93.4bp for the paired reads. The de novo assembly (length fraction = 0.5 and similarity = 0.8) generated 42 contigs (>2,000bp) with an average coverage of 1,075 fold. The average length of the contig was 121,605 bp, and the largest contig was 483,425 bp, with a N50 contig size at 299,659 bp. Scaffolding of the contigs was processed using SSPACE basic v2.0 (6). Two scaffolds with gaps from 367 bp to 4,771 bp in size were obtained. For finishing, the gaps were closed by the mapping of mate-pairs using as reference the 5 kbp from each of the contig ends (read length = 0.9, identity = 0.95). Then, using homemade script and fastqselect.tcl from MIRA3 package, the mapped reads for both orientations (R1 and R2) were retrieved and de novo assembled (read length = 0.5, identity = 0.8). The remaining gaps were resolved by PCR amplification followed by Sanger sequencing.

II.4. PacBio sequencing and assembly

II.4.1. Single Molecule Real Time (SMRT) Sequencing

An aliquot of about 4 μ g genomic DNA of strain 3337 was sheared using g-TUBETM (Covaris®) into fragments size targeted at 10 kb. SMRT cell template libraries were then prepared from the sheared DNA using the commercial Template Preparation Kit from Pacific Biosciences Inc. In brief, the DNA was end repaired, adapters ligated followed by exonuclease digestion of incompletely ligated products. Using the provided Template Binding Kit from Pacific Biosciences, the libraries were then annealed with sequencing primers and bound to P4 DNA polymerase. For enhanced loading efficiency into the sequencing zero-mode waveguides (ZMWs), the bound complexes were immobilized into Magbeads (Pacific Biosciences Inc.) according to the

accompanied protocols. Sequence collection was carried out in 3 SMRT cells using P4/C2 chemistry with duration set at 180 minutes with stage start option for each cells.

II.4.2. De novo Assembly using Hierarchical Genome Assembly Process (Hgap)

Upon acquisition of the sequencing data, the reads were filtered and assembled using RS_HGAP_Assembly version 3.0, an analysis pipeline module from Pacific Biosciences's SMRT portal analysis. The assembly protocol incorporates Celera Assembler, BLASR mapper and Quiver consensus caller algorithm for generation of polished assemblies.

Prior to assembly, short reads that are less than 500 bp were filtered off and minimum polymerase read quality was set at 0.75. For the assembly process, the length cutoff of seeding reads to serve as reference to recruit shorter reads for preassembly was set at 3,606 bp.

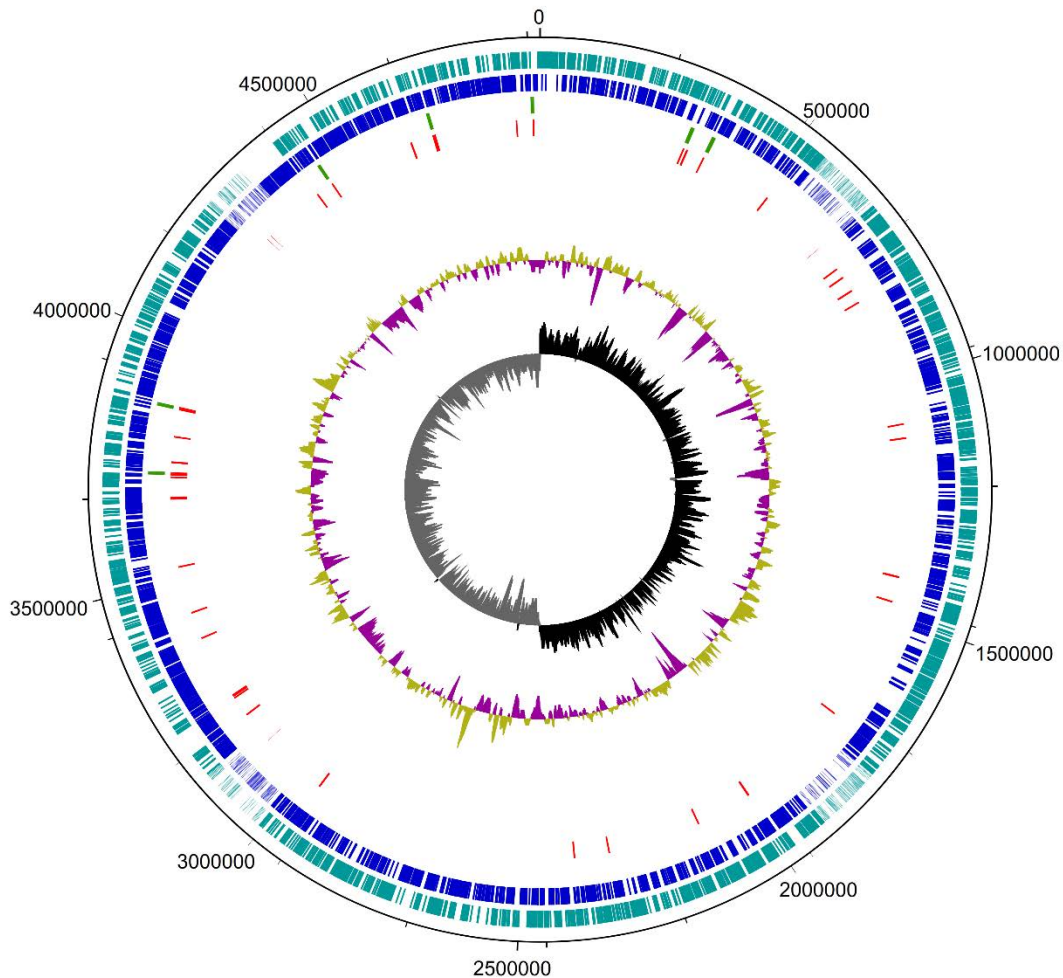
A total of 112,228 polymerase reads with average read length of 7,257 bp, amounting to 814,445,948 bases were generated after quality filtering process. Polished assembly size of the draft genome of strain 3337 was at 4,945,173 bp. The assembly has yielded 6 polished contigs with average sequencing depth of 144 X and the largest contig size was at 2,473,627 bp.

II.5. Complete genome sequence

Combining these two sets of sequences, the Illumina scaffolding was confirmed and the remaining gaps were filled using the PacBio contigs. Hence, we obtained a complete sequence of the unique circular chromosome (4,922,460 bp). The RAST annotation generated 4,536 CDS and 97 RNAs (Fig. II.1).

Figure II.1 Map of *Dickeya solani* 3337 genome.

The innermost circle highlights GC-skew variations, the second indicate GC-content (higher and lower than the mean value 56 %) ; the last circles in aqua (forward strand) and blue (reverse strand) colors shows genes of the *D. solani* 3337. In red are indicated tRNAs and in green rRNAs.



II.6. Nucleotide sequence accession numbers

The draft genome sequence generated by Illumina was deposited at GenBank. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AMY100000000. The version described in this paper is the first version, AMY101000000.

II.7. Acknowledgments

This work was supported by a cooperative project between France and Morocco (PRAD 14-02, Campus France no. 30229 ZK) and a collaborative project between CNRS, FN3PT, and CNPPT-SIPRE. CNPPT-SIPRE.

II.8. References

1. Burkholder WD, McFadden LA, Dimock AW. 1953. A bacterial blight of chrysanthemums. *Phytopathology* 43: 522-526.
2. Samson R, Legendre JB, Christen R, Fischer-Le Saux M, Achouak W, Gardan L. 2005. Transfer of *Pectobacterium chrysanthemi* (Burkholder *et al.* 1953) Brenner *et al.* 1973 and *Brenneria paradisiaca* to the genus *Dickeya* gen. nov. as *Dickeya chrysanthemi* comb. nov. and *Dickeya paradisiaca* comb. nov. and delineation of four novel species, *Dickeya dadantii* sp. nov., *Dickeya dianthicola* sp. nov., *Dickeya dieffenbachiae* sp. nov. and *Dickeya zaeae* sp. nov. *Int. J. Syst. Evol. Microbiol.* 55: 1415-1427.
3. Hugouvieux-Cotte-Pattat N, Condemine G, Nasser W, Reverchon S. 1996. Regulation of pectinolysis in *Erwinia chrysanthemi*. *Ann. Rev. Microbiol.* 50: 213-257.
4. Toth IK, van der Wolf JM, Saddler G, Lojkowska E, Pirhonen M, Tsrer L, Elphinstone JG, Hélias V. 2011. *Dickeya* species: an emerging problem for potato production in Europe. *Rev. Plant Pathology* 60: 385-399.
5. Hélias V. 2012. Jambe noire: Evolution des souches et risques associés. *La pomme de terre française*. 580 : 48-49.
6. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578-579.
7. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman A. L, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9: 75.

Chapitre III. Population genomics reveals additive and replacing horizontal gene transfers in the emerging pathogen *Dickeya solani*

L'émergence de *Dickeya solani* et son essor comme pathogène de pomme de terre en Europe a accentué les travaux de recherche pour comprendre l'écologie et les traits génétiques à l'origine de son succès écologique (Toth *et al.*, 2011). Plusieurs études ont été conduites dans ce sens révélant le caractère clonal de l'espèce ainsi que la forte synténie avec le génome de *D. dandantii*. Des divergences sont cependant observées concernant les complexes de toxine-antitoxine liés aux systèmes de sécrétion de type 5/6 (Garlant *et al.*, 2013 ; Laurila *et al.*, 2008 ; Pédrón *et al.*, 2014 ; Sławiak *et al.*, 2009). Potrykus *et al.* (2014) a mis en évidence l'implication des régulateurs de type *pecS*, *pecT* et *kdgR* dans la virulence de *D. solani*.

Dans ce contexte, nous avons envisagé une analyse de génomique de population sur une vingtaine de souches de *D. solani* en prenant comme référence la séquence complète de *D. solani* 3337 (Chapitre II). La génomique de population va nous permettre d'avoir une idée globale sur la diversité et la dynamique génomique de l'espèce *D. solani* en faisant un échantillonnage étendu dans le temps et dans l'espace (lieu, année, plantes hôtes). Dix-neuf souches ont été séquencées en *paired-end* (Illumina) et les variations génétiques (SNP/InDels) ont été recherchées en alignant les lectures de chaque souche sur le génome de référence. Ce qui était attendu de cette analyse c'est une forte homogénéité génétique au sein de la population de *D. solani* dont certains isolats ont peu de divergence (de 45 à 85 SNPs/InDels). En revanche, ce qui s'est avéré inattendu, est l'observation dans certains isolats de variations de quelques centaines à milliers de SNPs/InDels qui sont regroupés dans des îlots génomiques. Leur analyse phylogénétique révèle qu'ils ont été acquis par transferts horizontaux d'autres pathogènes. Par ailleurs, l'analyse des fonctions affectées par les SNPs/InDels a permis de prédire, puis de vérifier sur pomme de terre qu'un des isolats était peu virulent. Cette étude représente une première description de transferts horizontaux inter et intra-espèces chez *D. solani* contribuant ainsi à sa rapide diversification génomique. Elle apporte des éléments novateurs sur la diversité et l'évolution de *D. solani*.

Dans ce travail, j'ai réalisé les analyses de génomique comparative, les essais de virulence ont été effectués par Pauline Blin (Doctorante, Paris sud). Les souches de *D. solani* ont été fournies par V. Hélias et F. Van Gijsegem, le séquençage PacBio a été fait par T. Chong, K. Chan. Le manuscrit a bénéficié de la relecture critique de D. Faure, F. Van Gijsegem, J. Pédrón, V. Hélias et Y. Dessaux. Ce chapitre est présenté sous forme d'une version modifiée de l'article publié dans *BMC genomics* (Khayal *et al.*, 2015), la partie qui concerne le séquençage PacBio de *D. solani* 3337 a été développée dans le chapitre précédent (Chapitre II).

III.1. Abstract

Background

Dickeya solani is an emerging pathogen that causes soft rot and blackleg diseases in several crops including *Solanum tuberosum*, but little is known about its genomic diversity and evolution.

Results

We combined Illumina and PacBio technologies to complete the genome sequence of *D. solani* strain 3337 that was used as a reference to compare with 19 other genomes (including that of the type strain IPO2222^T) which were generated by Illumina technology. This population genomic analysis highlighted an unexpected variability among *D. solani* isolates since it led to the characterization of two distinct sub-groups within the *D. solani* species. This approach also revealed different types of variations such as scattered SNP/InDel variations as well as replacing and additive horizontal gene transfers (HGT). Infra-species (between the two *D. solani* sub-groups) and inter-species (between *D. solani* and *D. dianthicola*) replacing HGTs were observed. Finally, this work pointed that genetic and functional variation in the motility trait could contribute to aggressiveness variability in *D. solani*.

Conclusions

This work revealed that *D. solani* genomic variability may be caused by SNPs/InDels as well as replacing and additive HGT events, including plasmid acquisition; hence the *D. solani* genomes are more dynamic than that were previously proposed. This work alerts on precautions in molecular diagnosis of this emerging pathogen.

III.2. Background

Pectinolytic bacteria belonging to the *Dickeya* and *Pectobacterium* genera are pathogens that cause soft rot and blackleg diseases in a wide range of plants and crops including *Solanum tuberosum* [1],[2]. These phytopathogens produce plant cell-wall degrading enzymes that are able to macerate the tuber and stem tissues, thus provoking the disease symptoms [3]. Since 2000s,

the emerging *D. solani* species has been proposed as a contributor to the increased incidence of blackleg and soft rot diseases on potato crop in Europe and the Mediterranean basin [4]. The *D. solani* species has been officially described recently [5].

Little is known about the ecological and genetic traits that may support the relative success of *D. solani* in invading potato fields [6], [7]. *D. solani* can initiate disease from a low inoculum level in warm climates and was described in some studies to spread more easily through vascular tissues than other *Dickeya* species [4], [8]. Besides classical intergenic spacers 16S-23S rDNA [9], several molecular studies have proposed different marker genes for the identification of *D. solani* strains collected from potato and ornamental plants, such as *dnaX* [10], *recA* [11] and *fliC* [12]. At whole genome level, genomic and metabolic comparisons of two *D. solani* strains Ds0432-1 (isolated in Finland) and 3337 (isolated in France) vs. *D. dadantii* 3937 indicated a conserved synteny between the two species, but also the presence of distinctive traits [13], [14]. *D. solani* and *D. dadantii* diverged in their battery of non-ribosomal peptide/polyketide synthase clusters, T5SS/T6SS-related toxin-antitoxin systems and several metabolic abilities. Some of these traits would contribute to the successful invasion of this pathogen [13], [14]. More recently, a reverse genetic approach revealed that the virulence master regulators are quite the same in *D. solani* and *D. dadantii* [7].

The analysis of population genome structure and dynamics, including additive or replacing horizontal gene transfer (HGT) may bring valuable clues on the mechanisms of emergence of *D. solani*. While additive HGT allows the acquisition of novel genes by a population [15–20] replacing HGT provokes the replacement of an allele by another from close relatives [21]. HGT events inform about the genome diversification and adaptation processes, but also on the companion populations that the pathogens met during the emergence and dissemination steps. Replacing HGT is also of a major stake in pathogen diagnostic, as it may provoke false identification when the alleles exchanged by replacing HGT are used as molecular taxonomic markers.

Here, we analyzed the whole genome polymorphism of 20 *D. solani* isolates, including the type strain IPO2222^T, collected from different geographic locations, dates of isolation and plant hosts. We combined Illumina and PacBio technologies to complete the 3337 *D. solani* strain genome that we used as a reference in the comparative genomics. While most strains belonged to a core-population that exhibited less than one hundred variant positions between two given genomes, some other genomes revealed massive replacing HGT from the companion pathogen *D. dianthicola* and a plasmid acquisition from *Burkholderia ambifaria*. Moreover, we were able to correlate SNPs in virulence genes with a decrease in aggressiveness, highlighting the power of genomics as a tool to reveal functional variability in *D. solani* population. To our knowledge, this is the first study that reports whole genome analysis of a *D. solani* population and describes its diversity.

III.3. Results

III.3.1. Complete genome of the *D. solani* 3337

The *D. solani* 3337 genome was previously sequenced by Illumina technology using two libraries (*mate-pair* and *paired-end*) and de novo assembled in a high quality draft genome deposited at NCBI [22]. In this work, the 3337 *D. solani* genome was re-sequenced using PacBio technology. The PacBio sequencing generated six contigs (2 473 62pb, 1 512 701 bp, 894 591 bp, 49 337 bp, 10 627 bp, and 4 290 bp) with an average 150 fold coverage. The published Illumina-scaffolding was confirmed and the remaining gaps were filled using the PacBio contigs. Hence, combining the Illumina and PacBio sets of sequences, we obtained a complete sequence of the unique circular chromosome (4 922 460 bp). The RAST annotation generated 4536 CDS and 97 RNAs. The *D. solani* 3337 complete genome was used as a reference for comparative genomics.

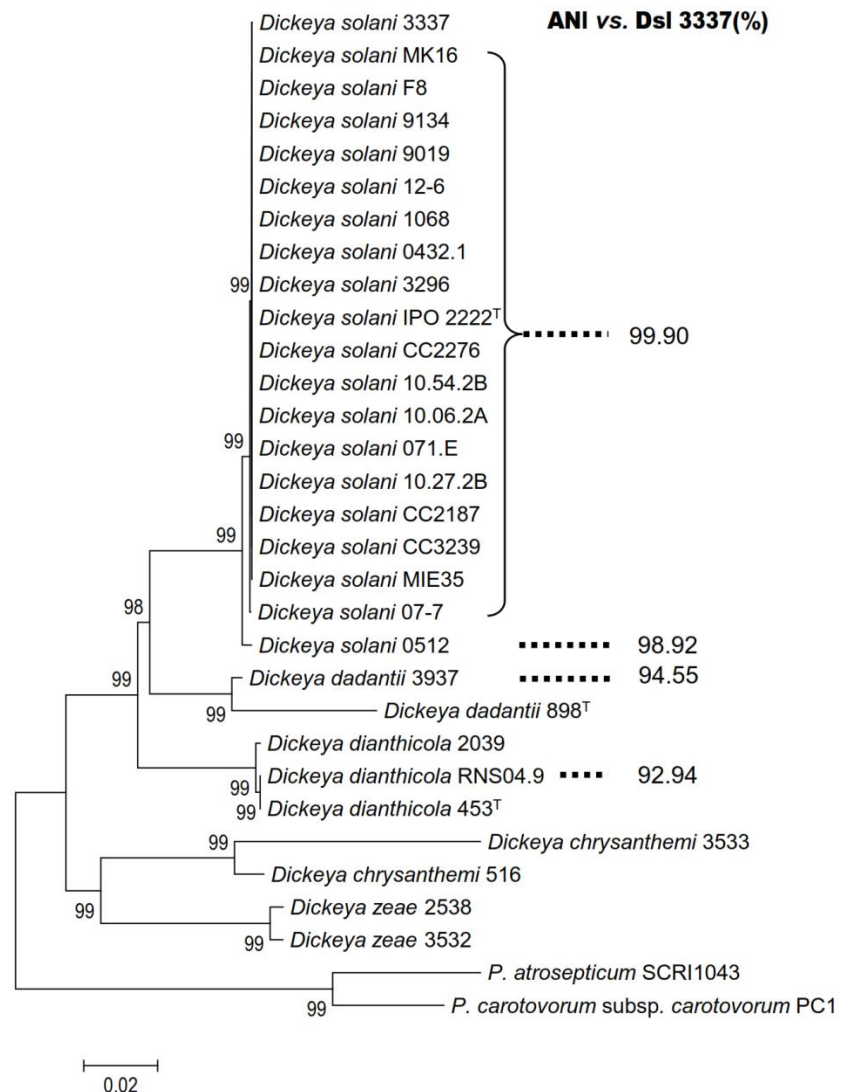
III.3.2. Positioning the sequenced *D. solani* strains within the *Dickeya* genus

In addition to *D. solani* 3337, 19 *D. solani* strains including the type strain IPO2222^T were collected at different years and geographical locations (Table III.S1) and their genomes sequenced by Illumina technology. All these draft and complete genomes were used in multi-locus sequence analysis (MLSA) and average nucleotide identity (ANI) calculation. For MLSA, the eleven concatenated *rpoD*, *gyrB*, *recA*, *rpoS*, *dnaX*, *dnaA*, *gapA*, *fusA*, *rplB*, *purA*, *gyrA* housekeeping genes (17298 bp) were aligned to construct a relation-tree using Neighbor-Joining method, the evolutionary distances were computed using the Maximum Composite Likelihood method [23]. All the *D. solani* (Dsl) isolates were grouped in a same cluster that was separated from the other pectinolytic enterobacteria (Fig.

III.1).

Figure III.1 MLSA and ANIs of *D. solani* strains.

In MLSA the sequences of the genes (*rpoD*, *gyrB*, *recA*, *rpoS*, *dnaX*, *dnaA*, *gapA*, *fusA*, *rplB*, *purA*, *gyrA*) were aligned with ClustalW, and a Neighbour-joining tree was created by Bootstrap method with 1000 bootstrap replications. ANI values were calculated using Dsl 3337 as a reference.



Noticeably, the strain Dsl 0512 was the unique strain that was consistently distant from the other *D. solani* strains. As previously reported [5], within the genus *Dickeya* the most related species to *D. solani* were *D. dadantii* and *D. dianthicola*. ANI values which were calculated using the strain Dsl 3337 as a reference were in accordance with the MLSA clustering. All the *D. solani* strains exhibited an ANI value equal to or above 99.9%, but that of Dsl 0512 was below 99%. Among strains of the closest species, *D. dadantii* and *D. dianthicola* ANI values dropped to 94% and 92%, respectively.

III.3.3. Overview of the SNP and InDel variations in *D. solani* genomes

Illumina reads of the *D. solani* strains were mapped on the complete genome sequence of Dsl 3337. The percentage of mapped reads was above 99% for all strains with the exception of Dsl 9019 (98.08%) and 0512 (92.34%) (Table III.S2). The mapping vs. Dsl 3337, which reached a high mean coverage value (between 400 and 900), allowed us to identify variations (SNPs and InDels) in each of the genomes (Table III.S3). According to the number of variations, the *D. solani* strains could be clustered into three groups. The first group, which we thereafter term as the core-population, encompassed most of the strains (including IPO2222^T and the reference Dsl 3337) with a variation number ranging from 43 to 85. In the second group were the strains Dsl 07-7, 9019 and 9134 with a variation number between 1454 and 3433. The third group consisted in the only strain Dsl 0512 with a very high variation number that reached 37493. RAST annotation of the strain Dsl 3337 was used to position the variations in or out coding DNA sequences (CDSs), as well as to identify non-synonymous variations in CDSs (Table III.S3). Non-synonymous variations ranged between 14 to 21% of the total number of variations, hence only 6 to 18 non-synonymous variations were identified in strains of the *D. solani* core-population (Table III.S3).

III.3.4. Heterogeneous distribution of the SNP and InDel variations in *D. solani* genes

We refined our analysis by calculating the number of genes (CDSs) that were affected by SNPs and InDels as well as non-synonymous variations (Fig. III.2a-b). In the core-population, 9 to 17

genes exhibited variations and about one half of them (4 to 10) harbored non-synonymous variations. In Dsl 07-7, 9019 and 9134, 56 to 144 genes were affected ; and among them, 46 to 81 contained non-synonymous variations. In Dsl 0512, 2760 genes, hence half of the genome showed variations. To compare variation abundance in genes, a mean value of the number of all variations (synonymous and non-synonymous) per affected gene was calculated (Fig. III.2c). In the core-population, this value ranged from 2 to 5. In Dsl 0512 and 07-7, it was similar (11 and 9, respectively) while the highest value was observed in Dsl 9134 and 9019 (45 and 42, respectively). Overall, these analyses revealed that Dsl 0512, 07-7, 9134 and 9019 harbored genes with different numbers of variations as compared to those in the core-population, suggesting putative HGTs from distinct sources.

Finally, all these different variations were positioned along the Dsl chromosome (Fig. III.3). In the core-population, the rare variations appeared to be scattered with a mean distribution of 0.015 variations per kbp (when all SNPs and InDels were counted) and 0.012 variations per gene (when only SNPs and InDels in CDSs were counted). In Dsl 9134, 9019 and 07-7, most of the variations affected several tens of genes that are clustered in distinct regions, while only a few variations remain scattered. In Dsl 0512, variations exhibited a genome wide distribution. In the next part of the work, the three types of SNP/InDel distribution (scattered, clustered and wide genome distribution) have been analyzed in details.

Figure III.2 Number of genes affected by variations (SNPs and Indels).

a. Total number of genes affected by the variations for each strain. **b.** Number of genes revealing amino acids change further variation. **c.** The average number of variations affecting the genes in each strain.

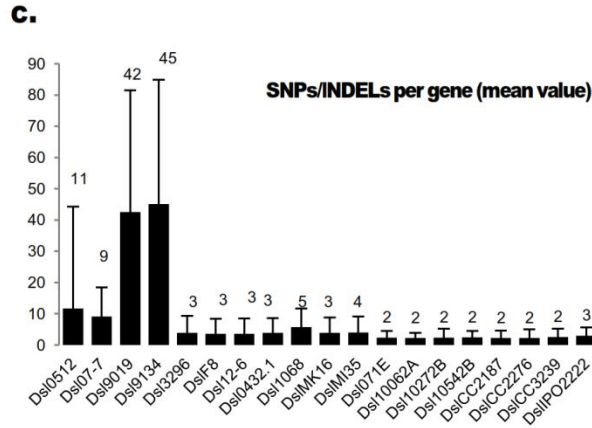
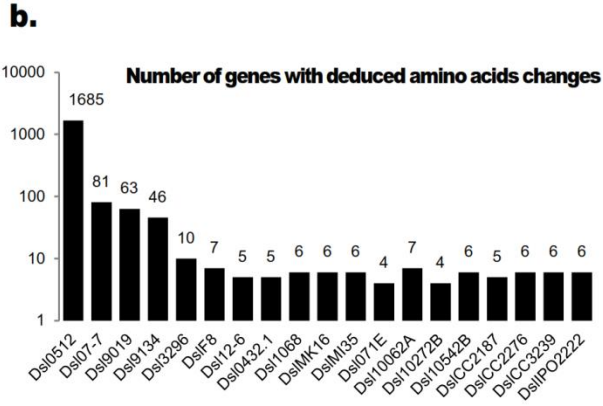
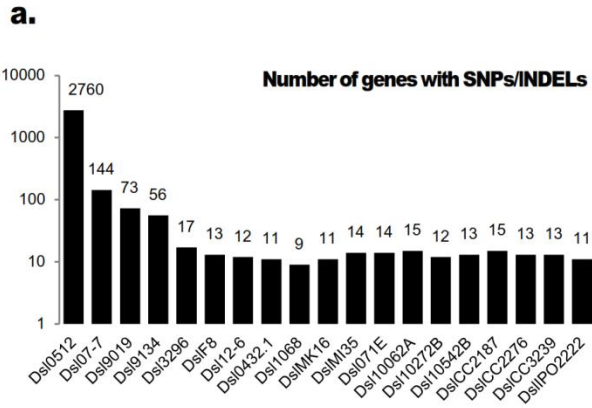
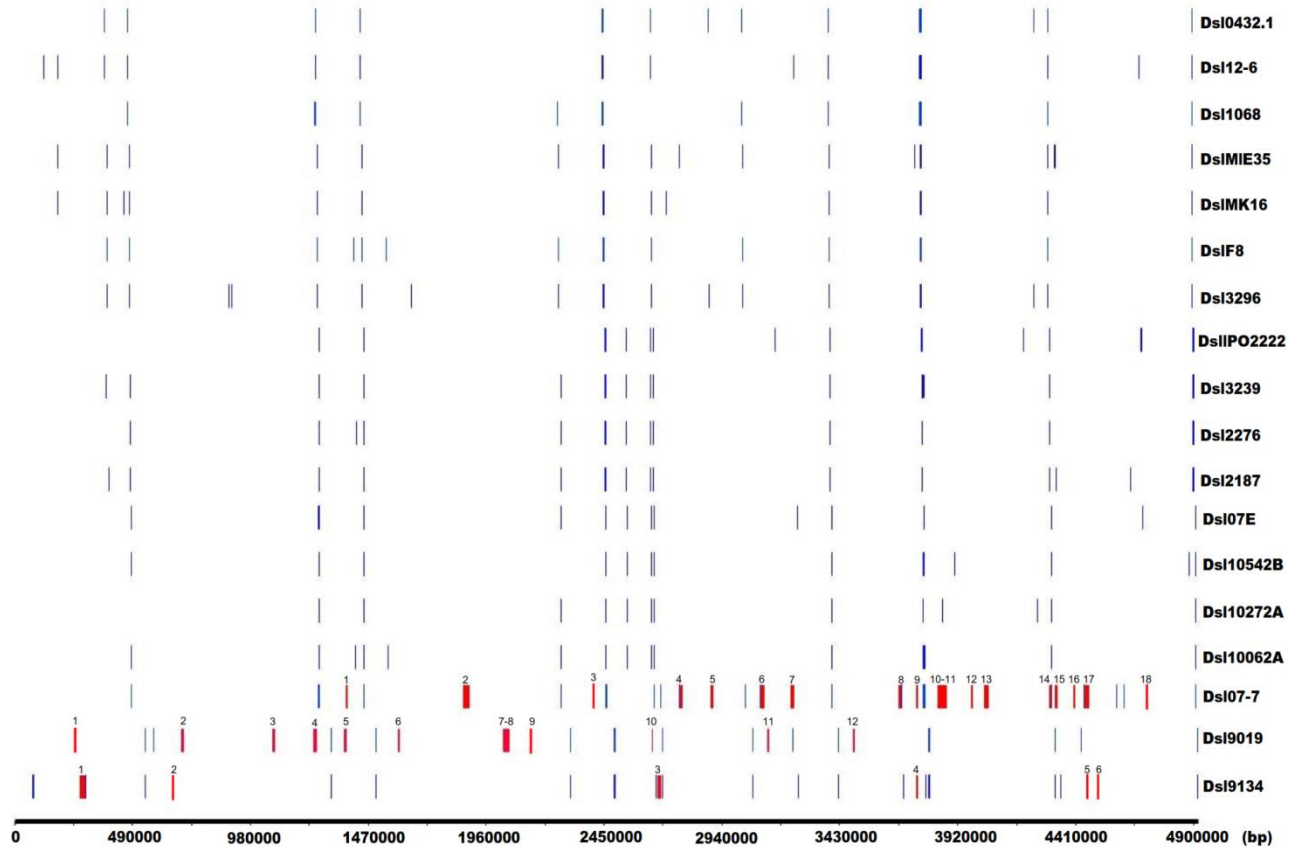


Figure III.3 Mapping of the clustered and scattered SNP/InDel variations using Dsl 3337 as a reference genome. Small colored sticks indicate variations positions: the scattered SNP/InDels are in blue color, while the clustered SNP/InDels (RGTs) are in red color and are numbered according to their successive position along the chromosome. Dsl 0512 is excluded from this figure due to high number of variations.



III.3.5. The mosaic genome of *D. solani* 0512 might define a novel *D. solani* sub-group

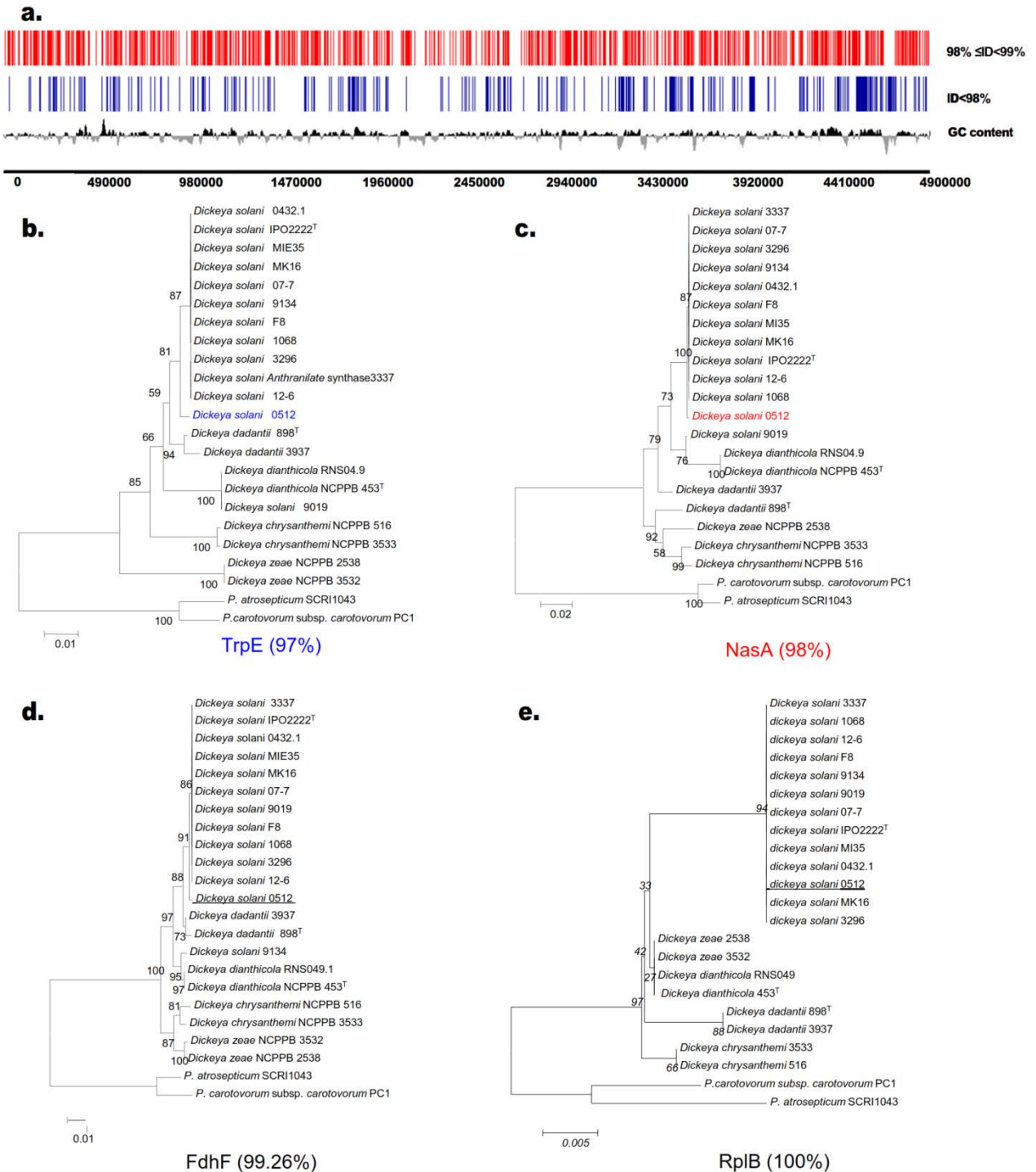
Dsl 0512 differed from the other *D. solani* by the high number and wide distribution of variations (Table III.S3, Fig. III.2 and III.3), a unique phylogenetic position in MLSA (Fig. III.1), and a high percentage (7.66%) of unmapped reads against Dsl 3337 genome (Table III.S2). Unmapped reads were used for a *de novo* assembly which generated six contigs with a size ranging from 13248 bpto 36630bp. All these six contigs were absent from the other *D. solani* strains. Using MAUVE [24], these six sequences were positioned on the draft genome of Dsl 0512 that was constructed using the strain Dsl 3337 as a reference (Fig. III.S1). RAST annotation indicated that most of the genes belonging to these 6 contigs coded for phage elements and hypothetical or unknown

proteins, with the exception of some genes coding for two putative ABC transporters, two putative virulence factors and one methyl-accepting chemotaxis protein, all being carried by the contig4. The similarity scores were too weak to assign a more precise function and phylogenetic origin to these putative genes/proteins.

Another characteristic of Dsl 0512 was a high number of genes (half of the genome) that exhibited variations. These genes were distributed along the genome without any clustering in specific regions. Constructed phylogenetic trees revealed that the analyzed genes exhibiting a nucleotide identity below 98% (compared to Dsl 3337 genes) did not belong to the core population gene cluster (Fig. III.4, III.S2, III.S3, III.S4, III.S5). All these features supported the existence of a novel *D. solani* sub-group. The strain Dsl 0512 could be proposed as the eponym of the *D. solani* 0512 sub-group.

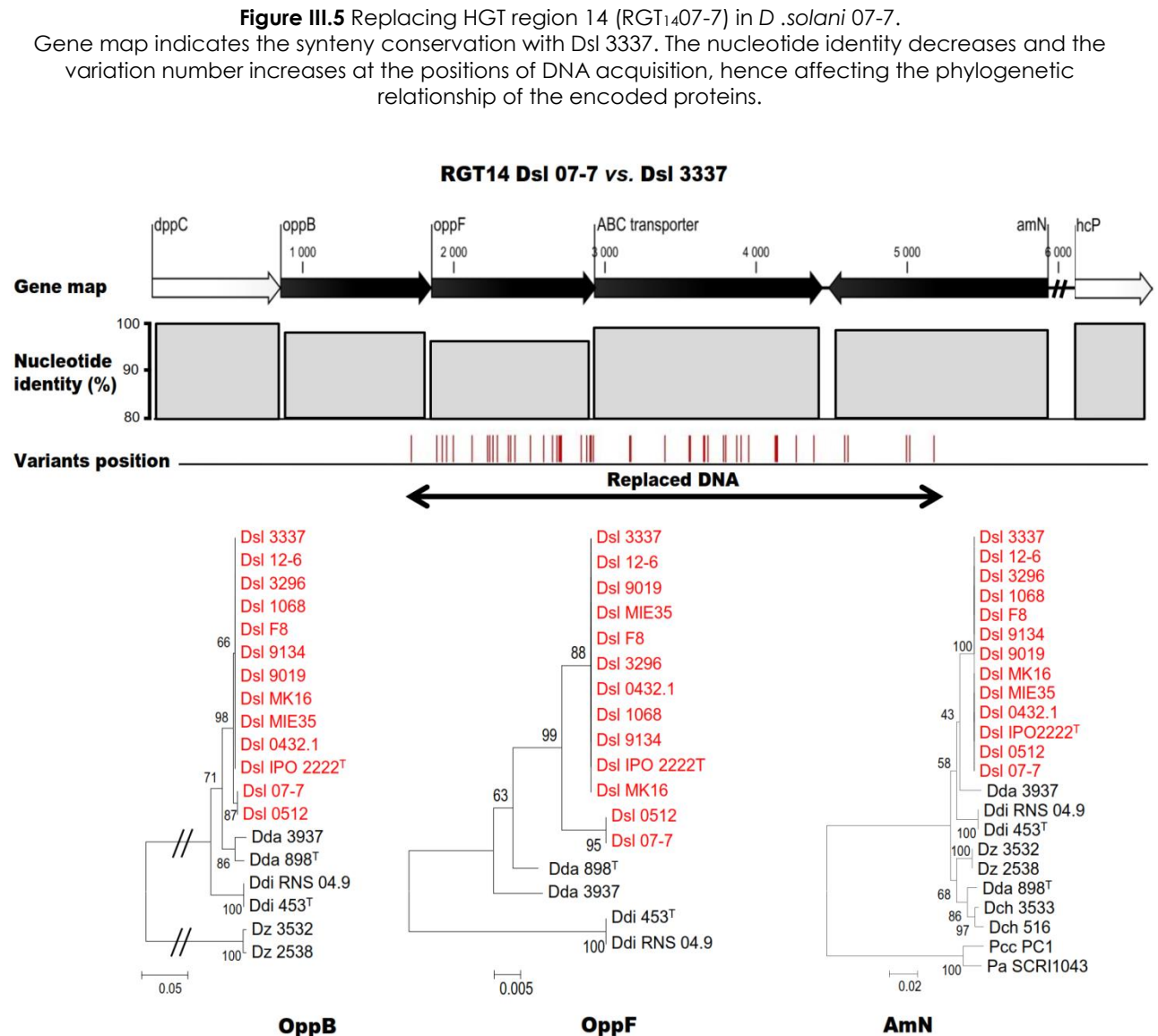
Figure III.4 Mapping and phylogeny of the Dsl 0512 variant genes.

In panel **a**, Position of variant genes of Dsl 0512 on the reference strains Dsl 3337. The panels **b**, **c**, **d** and **e** exemplify phylogenetic trees of selected proteins with different percentage of nucleotide identity.



III.3.6. Intra-species replacing HGT in Dsl 07-7

The 144 variant genes of Dsl 07-7 showed a non-uniform distribution along the chromosome, since most of them were clustered in 18 separate regions. One of these regions, which is presented in figure III.5, contained four genes: *oppB*, *oppF*, an ABC transporter gene and *amN*. These genes contained in total 47 variations leading to a decrease of their nucleotide identity as compared to the corresponding genes in the Dsl 3337 genome.



Moreover, the phylogenetic analysis of the protein sequence coded by the *oppB* and *oppF* genes, which were the most affected by variations, revealed a replacing HGT from a strain belonging to

the *D. solani* 0512 sub-group. The 17 other regions exhibited a similar gene organization and a phylogenetic clustering with Dsl 0512 genes. Hence, all these 18 regions were called as RGT (replacing HGT) regions. They were numbered according to their successive position along the chromosome with the strain name in subscript position: RGT₁07-7, RGT₂07-7, RGT₃07-7 ... (Fig. III.3). This analysis suggested that Dsl 07-7 acquired a dozen of gene fragments during massive replacing HGT(s) from strain(s) belonging to the Dsl 0512 sub-group. Hence, Dsl 07-7 exemplified the occurrence of an infra-specific gene exchange among the *D. solani* population, and also supported the possible co-existence of strains of the *D. solani* 0512 sub-group with those of the core-population.

III.3.7. Inter-species replacing HGT in *D. solani* strains 9134 and 9019

In Dsl 9134, 39 among the 56 genes with variations were clustered in 6 RGT regions, the other genes with variations being scattered along the chromosome. In Dsl 9019, 63 among the 73 genes with variations were clustered in 12 RGT regions. In both strains, the RGT regions were named according to the same nomenclature as in Dsl 07-7 (Fig. III.3).

The RGT₄9134 illustrated the typical organization of these RGTs in Dsl 9134 (Fig. III.6). RGT₄9134 (4860 bp) exhibited 229 positions of variations that were distributed in three genes: *norF*, *norR*, and *fumA*. These genes were related to the nitric oxide metabolism. Because of the high number of variations, the gene identity with *D. solani* strain 3337 decreased in RGT₄9134, especially in the *norR* gene that was located in the central part of the RGT region. Protein phylogeny revealed that the three proteins encoded by the RGT₄9134 genes did not branch with their *D. solani* counterparts but were most closely related to those of *D. dianthicola*. The variation positions suggested that replacing HGT occurred in the middle of the genes *norF* and *fumA*, and hence generated proteins with an intermediate position between the *D. solani* and *D. dianthicola* proteins in the phylogenetic trees. A second example of inter-species replacing HGT is given with the RGT₇9019 (6248bp) of Dsl 9019, which contained five genes *dnaJ*, *dnaK*, *yaaH*, a MFS transporter

gene and *mogA* (Fig. III.7). This example highlighted that replacing HGT might also affect genes such as *dnaK* and *dnaJ* which are used for MLSA and taxonomic identification [25]. In RGT₇₉₀₁₉, 269 variants were detected. Discrepancies within DnaJ, DnaK and MogA phylogenies suggested the occurrence of a replacing HGT from *D. dianthicola*. In all the other RGTs of Dsl 9134 and 9019, a phylogeny approach (Fig. III.S4, III.S5) also supported the occurrence of a replacing HGT using *D. dianthicola* population as the unique source.

Figure III.6 Replacing HGT region 4 (RGT₄₉₁₃₄) in *D. solani* 9134.

Gene map indicates the synteny conservation with Dsl 3337. The nucleotide identity decreases and the variation number increases at the position of DNA acquisition, hence affecting the phylogenetic relationship of the encoded proteins.

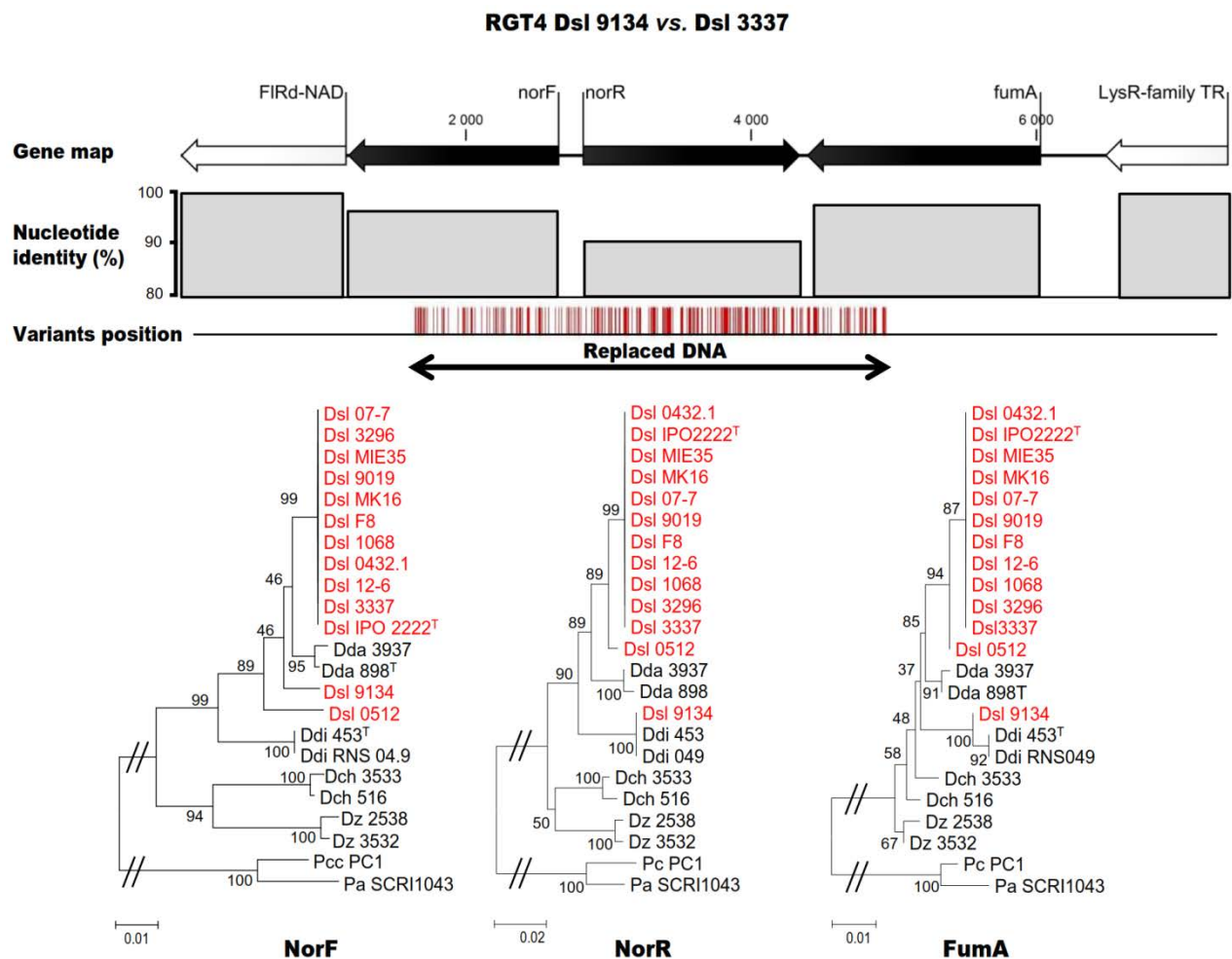
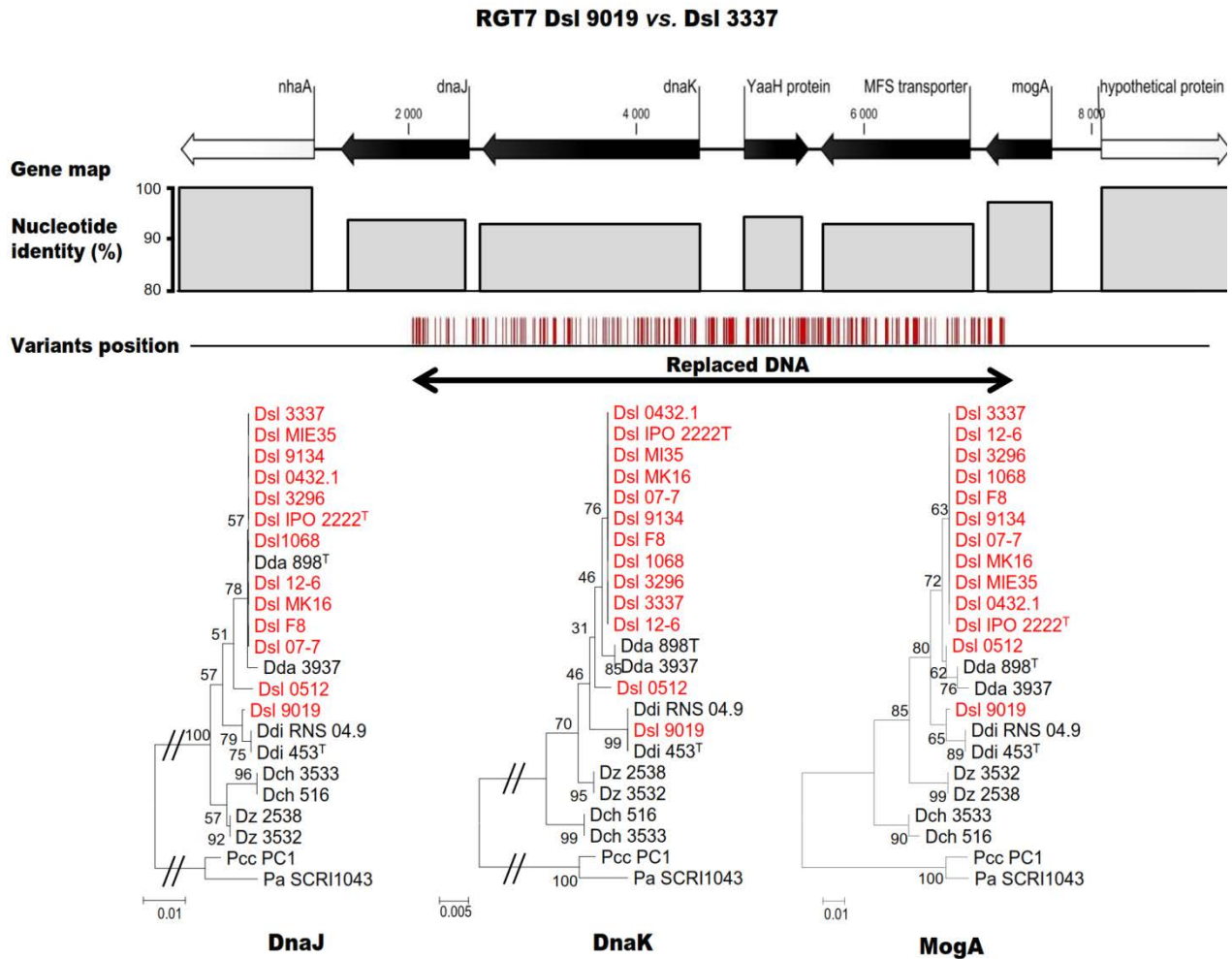


Figure III.7 Replacing HGT region 7 (RGT79019) in *D. solani* 9019.

Gene map indicates the synteny conservation with Dsl 3337. The nucleotide identity decreases and the variant number increase at the position of DNA acquisition, hence affecting the phylogenetic relationship of the encoded proteins.



III.3.8. Plasmid acquisition in *D. solani* strain 9019 from *Burkholderia*

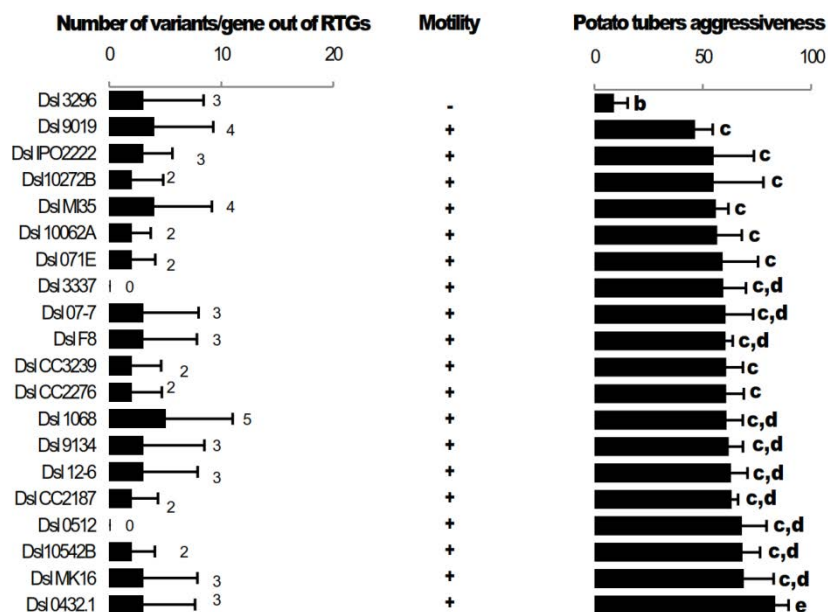
In addition to replacing HGT, an additive HGT event that consisted in a plasmid acquisition occurred in Dsl 9019. The Dsl 9019 unmapped reads, which represented 1.9% of the total read number (Table III.S2), allowed the generation of a single contig (43564 bp) by *de novo* assembly. This plasmid exhibited a complete identity (100%) with a plasmid of *Burkholderia ambifaria* AMMD (CP000443.1). The stable replication of this plasmid in Dsl 9019 was verified in sub-cultures using plasmid-specific primers (pF1:cagcgaagagcaagacaa, pR1:tcattggaagcgatctcgg and pF2:ttaccggacgccgagctgtggcgt, pR2:caggaagatgtcgttatcgcgagt).

III.3.9. In *D. solani* 3296, variations in flagellar genes correlated motility and virulence decrease

All the non-synonymous variations of the core-population were listed in Table III.S4. Remarkably, two unique non-synonymous variations that affected the *fliC* and *fliN* flagellar genes were present in Dsl 3296. The substitution C to T at the position 952985 lead to conversion of Ala207 to Thr in FliC, while deletion of the GTC codon starting at the position 966 038 provoke the loss of the Val112 in FliN. The nucleotide variations were verified by Sanger sequencing. These two variations were unique among the sequenced *D. solani* strains, as well as the known *Dickeya* genomes (Fig. III.S6). These genes retained our attention as *fli* genes are required for aggressiveness in *Dickeya* and in other *Enterobacteriaceae* [26],[27],[28],[29]. We hypothesized that Dsl 3296 could be impaired in motility, hence also exhibited a reduced aggressiveness on potato host plants. We compared motility and virulence of all the 20 Dsl analyzed in this study (Fig. III.8).

Figure III.8 Motility and aggressiveness assays performed on potato tubers.

The average of variants per gene was calculated for each strain (the RGT regions of the strain Dsl 9019, 9134 and 07-7 were omitted for calculation). The signs + and - indicate that the strain is motile or not. The letters b, c, d and e indicate statistical significance at $p < 0.05$ (Kruskal-walis and Tukey tests) of the aggressiveness which was measured by infecting 30 potato tubers by each of the Dsl strains.



All strains except Dsl 3296 were motile. Moreover, a weak aggressiveness of the strain 3296 was observed in virulence assay on potato tuber, hence correlating genomic variants in fli genes with motility and virulence deficiency. As a consequence, even if SNP/InDel variations are scarce, some of them may affect virulence functions in Dsl strains.

III.4. Discussion

This work provided new insights into the analysis of the emerging plant-pathogen *D. solani*. We combined Illumina and PacBio technologies to determine a high quality genome sequence of *D. solani* 3337 that we used as a reference to compare 19 other genome sequences generated by Illumina technology. While previous studies reported pairwise comparison between a single *D. solani* genome with that of other *Dickeya* and *Pectobacterium* species [13], [14], this work was also based on a population genomic approach. This approach revealed the unexpected diversity of the *D. solani* genomes that resulted from a combination of scattered SNP/InDel variations as well as replacing and additive HGT events.

The majority of analyzed *D. solani* strains (16 among 20) that we called the core-population contained only 43 to 85 variants (SNPs and InDels). This result is in accordance with the high ANI values (>99.9%) that were calculated between each strain against the reference 3337. Other studies have pointed the high homogeneity within genetic equipment of *D. solani* population [12],[14],[9],[30]. All these molecular analyses support the clonal hypothesis of the *D. solani* population. In spite of this high homogeneity, Dsl strains may exhibit some variability in aggressiveness on potato tubers [7]. A previous pairwise comparative study of two Dsl strains did not succeed in the identification of the genes and functions that could explain the different aggressiveness trait [7]. However, using a population comparative approach, we pointed out that genetic and functional variations in the motility trait could contribute to an aggressiveness decrease. This observation exemplifies the powerfulness of genomic diversity analyses on field isolates for the identification of genes that modulate aggressiveness.

Another important result was the characterization of a sub-group within the *D. solani* species, highlighting that *D. solani* population structure was more complex than described previously. The prototypic strain of this sub-group was Dsl 0512 (RNS 05.1.2A) that has been isolated from potato plant showing blackleg and soft rot symptoms in France (in 2005). The existence of the 0512 sub-group was supported by ANI value, MLSA, genomic architecture (presence of specific regions) and SNP/InDel abundance and distribution. The Dsl 0512 genome appeared as a mosaic of genes with a phylogenetic position inferred to either the *D. solani* core population or the Dsl 0512 sub-group. Remarkably, genes that belong to the Dsl 0512 phylogenetic sub-cluster have also been discovered in the 18 RGTs (143 genes) of the strain Dsl 07-7 that was also isolated in France. The involvement of the 0512 sub-group as a gene resource in replacing HGT reinforced its importance in the generation of variability in *D. solani* isolates. The strains Dsl 0512 and 07-7 showed aggressiveness level similar to that of most of the studied *D. solani* strains, suggesting that the 0512 sub-group is not associated to any particular aggressiveness behavior, at least on potato tubers.

This study also highlighted that additive and replacing HGT occurred in inter-species exchanges. Additive HGT was observed in the strain 9019 which acquired a plasmid from *B. ambifaria* AMMD. *B. ambifaria* AMMD was isolated from the rhizosphere of healthy pea plants in Wisconsin (USA) in 1985 [31] and it has been reported as very effective in controlling phytopathogenic *Pythium* species [32]. Moreover we discovered replacing HGT events that recruited *D. dianthicola* genes in the two strains Dsl 9019 (63 genes distributed in 12 RGT regions) and 9134 (39 genes in 6 RGT regions) isolated from ornamental plants (respectively Muscari and Hyacinth). These exchange events between *D. solani* and *D. dianthicola* suggested that these two pathogens could coexist in the same ecological niche. In the case of *Pectobacterium* spp., multiple species isolations from the same infected plants have been reported [33],[34]. Importantly, the replacing HGT events did not correlate with an aggressiveness increase in Dsl 9019 and 9134 at least in potato tubers.

However, replacing HGT generates major impact on phylogenetic inference by generating incongruities that could impair pathogen molecular diagnostics which are based on housekeeping genes. Importantly, it has been reported that succeeded HGT events between distantly related bacteria mostly implicate housekeeping genes that are also the most conserved between different species [35], [36]. Our work revealed that in the Dsl 9019 strain, the *dnaJ* and *dnaK* genes, which are usually used in phylogenetic classifications [37],[25], have been recruited from *D. dianthicola*. Since the *Dickeya* pathogens are genetically very close (ANI \geq 93%), replacing HGT could be predicted to interfere recurrently with taxonomical diagnostics, hence provoking assignation errors. The impact of HGT on taxonomy has been discussed in different *Enterobacteriaceae* [38], [39]. An immediate applied recommendation from our work is that even though *D. solani* is mainly described as a homogeneous population, the existence of HGT events should encourage the use of multiple taxonomical markers.

III.5. Conclusions

As a conclusion, this work revealed that *D. solani* genomic variability may be caused by SNPs/InDels as well as replacing and additive HGT events, including plasmid acquisition. From this work, the question arises about the dynamics of the *D. solani* diversity in the course of its emergence and spreading in crop cultures. This might be further investigated by a larger scale sampling and genomic analysis.

III.6. Methods

III.6.1. Bacterial strains and growth conditions

D. solani strains were collected from different geographical locations and dates of isolation and also from different hosts or environments (Table III.S1). All the strains were routinely cultured in TY medium (tryptone 5 g/L, yeast extract 3 g/L and agar 1.5%) at 28°C.

III.6.2. DNA extraction and sequencing

Genomic DNA from each strain was extracted from overnight culture using a phenol-chloroform purification method followed by an ethanol precipitation as described by Wilson [40]. Quantity and quality control of the DNA was completed using a NanoDrop (ND 1000) device and agarose gel electrophoresis at 1.0%

Paired-end libraries with an insert size of 270 to 390 bp were constructed for each strain, and DNA sequencing was performed by Illumina HiSeq 2000 v3 technology. Sequencing of the library was carried out using 2x100 or 2x150 bp paired-end read module. Illumina sequencing was performed at the CNRS IMAGIF platform (Gif-sur-Yvette).

III.6.3. Assembly, variants calling and genome sequence analysis

Assembly of the sequences was performed using the CLC Genomics Workbench v7.0.0 software (CLC Inc, Aarhus, Denmark). After quality (quality score threshold 0.05) and length (above 40 nucleotides) trimming of the sequences, contigs were generated by *de novo* assembly (CLC parameters: automatic determination of the word and bubble sizes with no scaffolding) for each strain.

Paired end reads for each strain were mapped against the reference sequence of the strain *D. solani* 3337 at mild stringency threshold (0.8 of identity on 0.5 of read length) using CLC Genomics Workbench version 7.0.0 software. The unmapped reads for each strain were collected. The mappings were used for detection of variations (SNPs and InDels) using basic variant calling tool from CLC genomic workbench version 7.0.0. Draft genome sequences composed of the contigs of each strain were used to search and analyze the variations detected. Variations with an occurrence below 99% in the mapping step were discarded from the study.

The nucleotide identity (ANI) values were calculated as previously proposed [41] using the ANI calculator from the Kostas lab with default settings (<http://enve-omics.ce.gatech.edu/ani/>). Phylogenetic and molecular evolutionary analyses were conducted using MEGA, version 6 [42].

An MLSA (Multi-locus sequence analysis) was performed using eleven housekeeping genes (*rpoD*, *gyrB*, *recA*, *rpoS*, *dnaX*, *dnaA*, *gapA*, *fusA*, *rplB*, *purA*, *gyrA*) retrieved from the twenty *D. solani* strains in order to confirm their phylogenetic position within known pectinolytic *Dickeya* and *Pectobacterium* strains.

III.6.4. Nucleotide sequence accession number

Draft genome sequences of *Dickeya solani* strains 9109, 0512, 9134, 07-7 have been deposited at DDBJ/EMBL/GenBank under the following accession numbers: (JWLS00000000) *D. solani* 9019, (JWMJ00000000) *D. solani* 0512, (JWLT00000000) *D. solani* 9134, (JWLR00000000) *D. solani* 07-7. The versions described in this paper are versions (JWLS01000000) *D. solani* 9019, (JWMJ01000000) *D. solani* 0512, (JWLT01000000) *D. solani* 9134, (JWLR01000000) *D. solani* 07-7. Genomes of other *Dickeya* and *Pectobacterium* species were collected from public database (Table III.S5).

III.6.5. Aggressiveness and motility assays

Motility assays were conducted on semi-solid SM medium (beef extract at 3 g/L, peptone at 5 g/L, and 25 ml/L of 20% glucose) with 0.5% of agar. Two μL of an overnight bacterial suspension of each strain were used to inoculate agar plates which were incubated 16 hours at 28°C. The experiment was performed twice with 2 replicates each time.

Assessment of the aggressiveness of the strains was performed on potato tubers (cv. Binjite). To this end, 10^6 CFU were used to infect 10 potato tubers for each strain. After 24 h of incubation at 25°C, five aggressiveness categories were considered and attributed to tuber samples to assess the virulence of the strains. The experiments were performed three times, hence 600 tubers were infected and analyzed. The results were represented as normalized values.

Virulence assays were statistically analyzed to infer the aggressiveness variability within strain on potato tubers. Heterogeneity of strains was assessed using a *Kruskal-Wallis* test with $p < 0.05$.

Statistical significance of the pairwise comparisons between strains was calculated using a post hoc *Tukey* test with $p < 0.05$.

III.7. Supplementary data

Figure III.S1 Synteny between the strain *D. solani* 3337 and the draft genome *D. solani* 0512. The alignment was performed using MAUVE software, underlining a high conservation of the synteny. The numbers indicate the positions of the strain-specific genomic regions generated by *de novo* assembly of the unmapped reads.

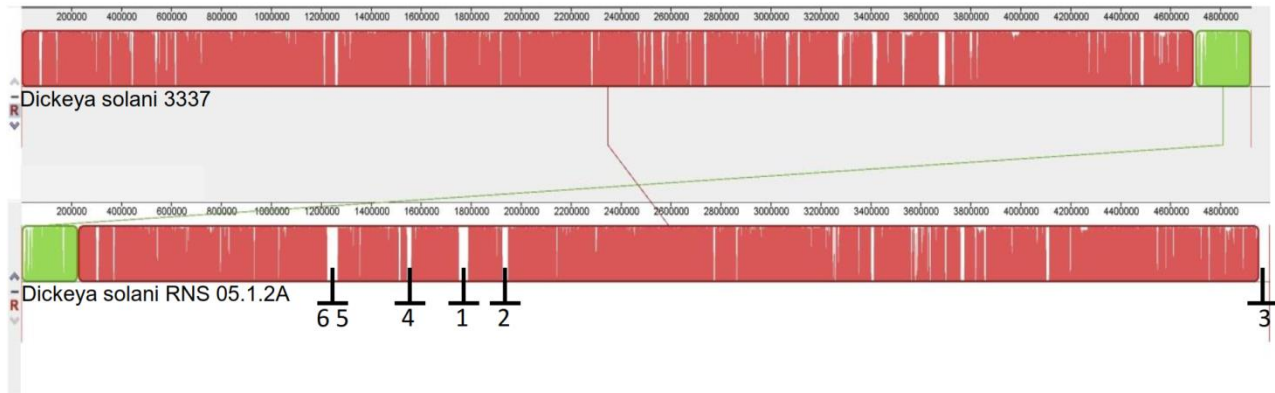


Figure III.S2 Protein-based phylogenetic trees revealing Dsl 0512 as a member of in distinct sub-cluster within the *D. solani* species.

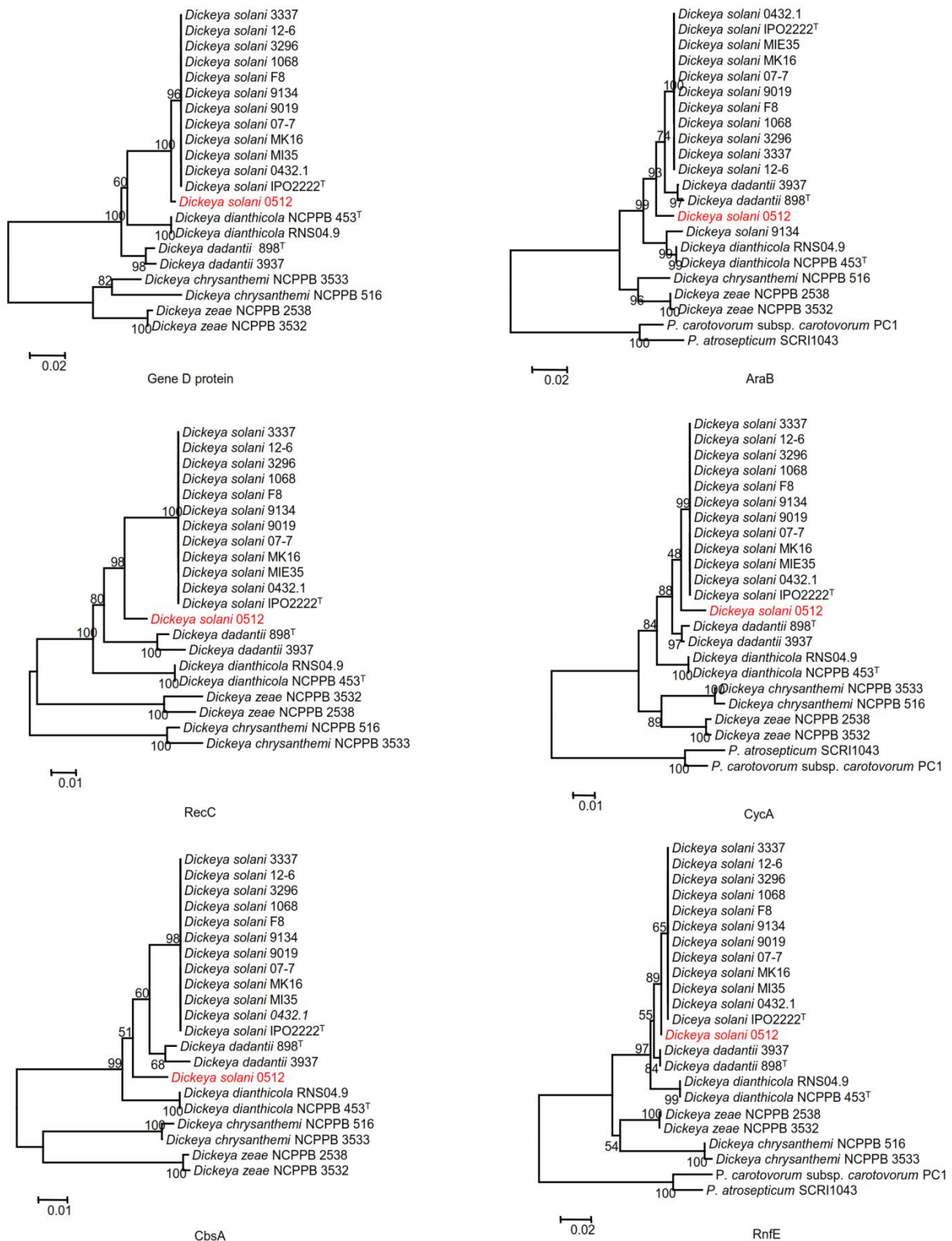


Figure III.S3 Protein-based phylogenetic trees of different RGTs in Dsl 07-7.

The genes were retrieved from RGT1, RGT2, RGT3, RGT6, RGT10 and RGT12 of Dsl 07-7. The phylogenetic positions indicate replacing HGT events from the *D. solani* 0512 sub-group.

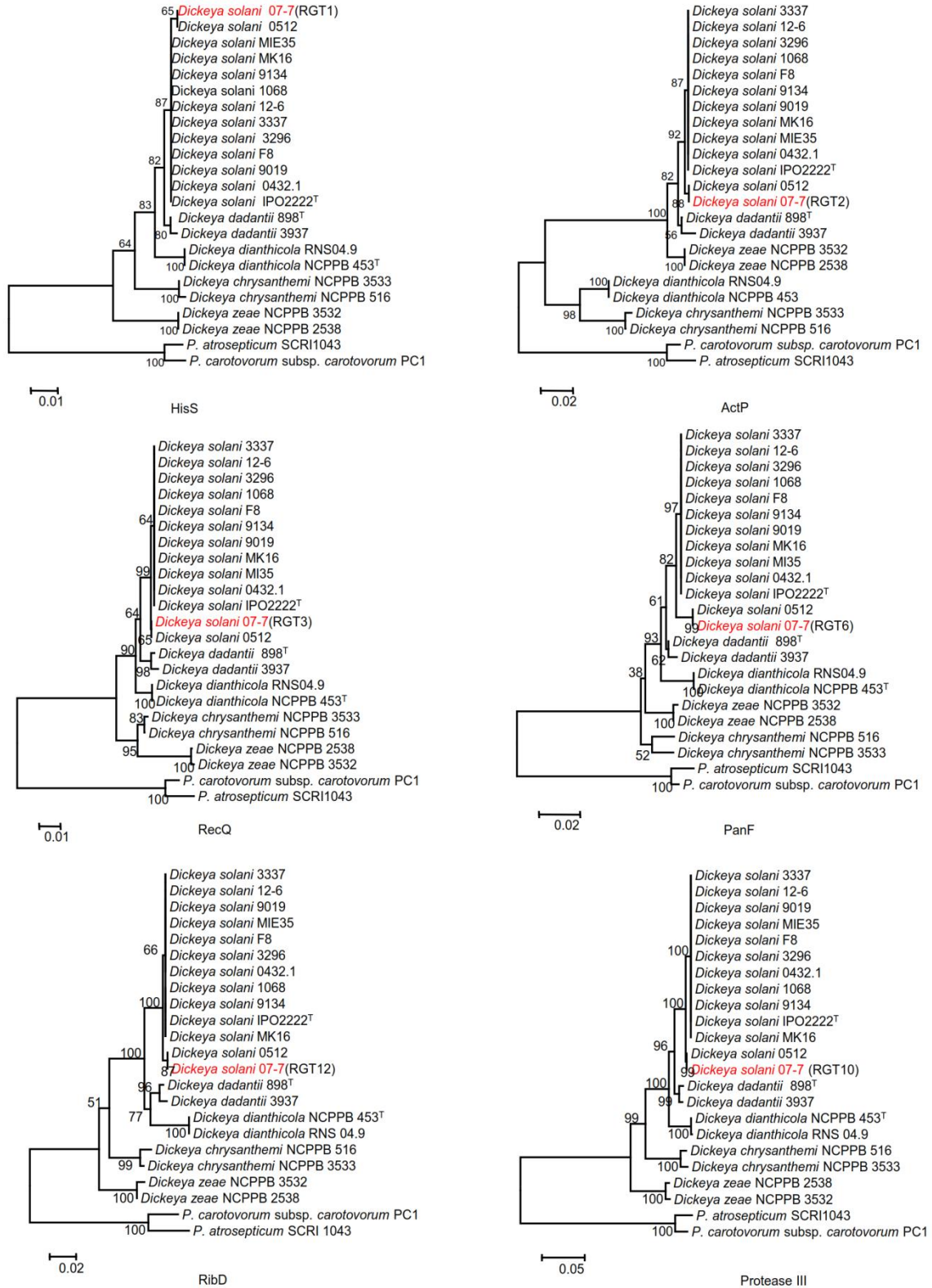


Figure III.S4 Protein-based phylogenetic trees of different RGTs in Dsl 9134. The genes were retrieved from RGT1, RGT3, RGT5 and RGT6 of Dsl 9134. The phylogenetic positions highlight replacing HGT events from the *D. dadantii* species.

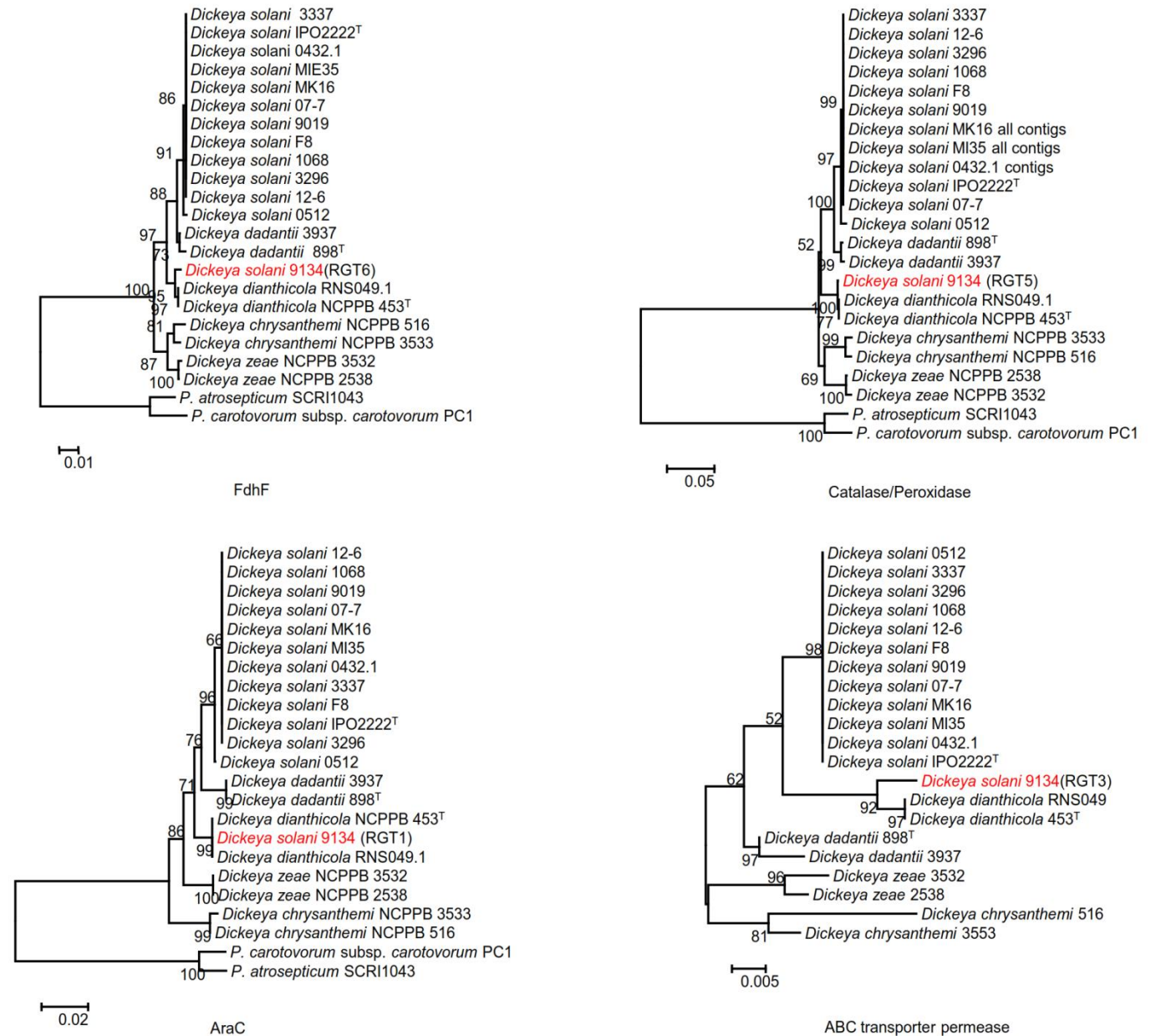


Figure III.S5 Protein-based phylogenetic trees of different RGTs in Dsl 9019.

The genes were retrieved from RGT1, RGT2, RGT3, RGT4, RGT5 and RGT10 of Dsl9019. The phylogenetic positions highlight replacing HGT events from the *D. dadantii* species.

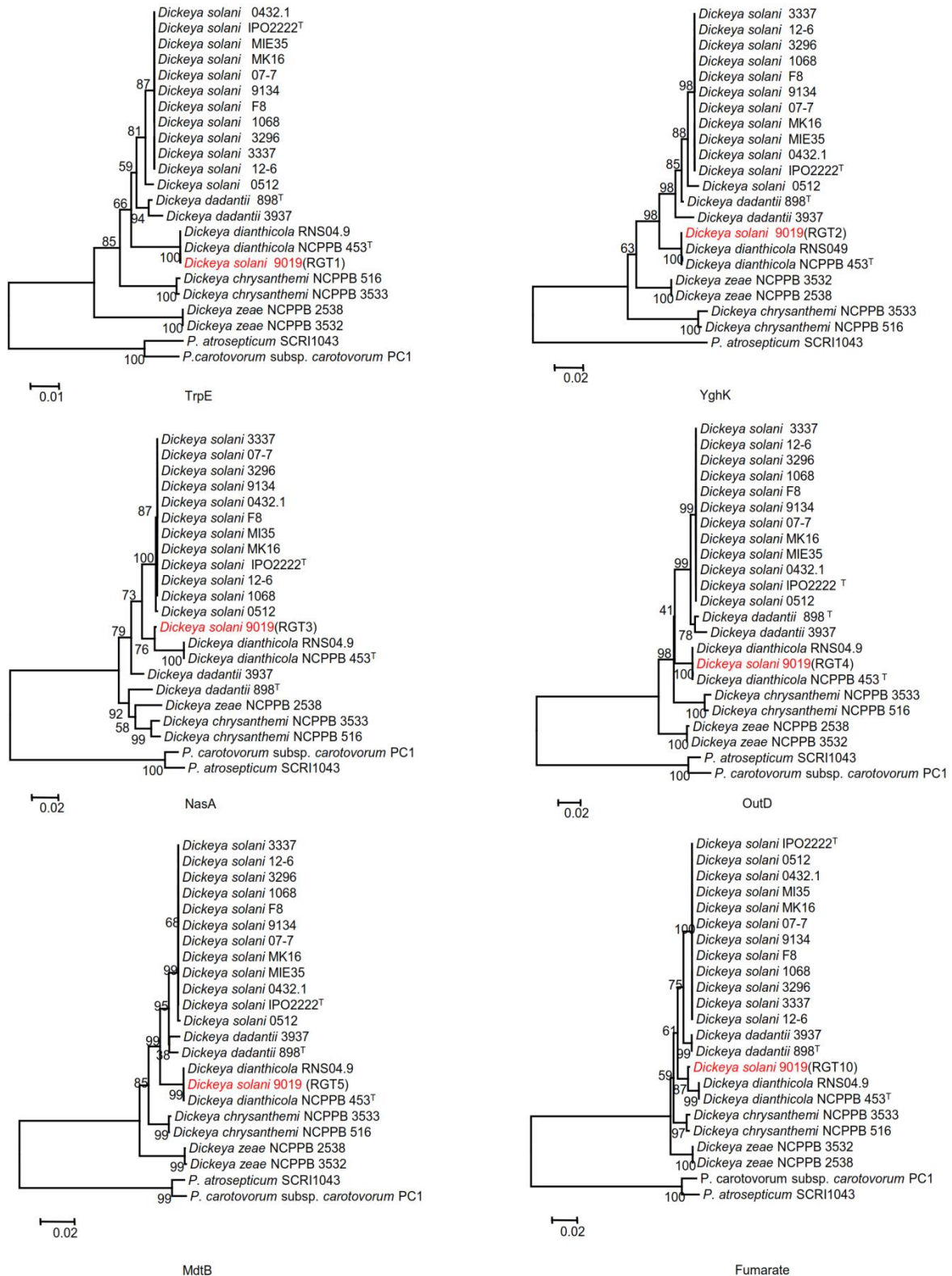


Figure III.S6 Local alignment of FliC and FliN proteins.

The variations at the positions 207 in FliC and 112 in FliN are indicated in red color, other variations are in blue color. Amino acid position is numbered according to the *D. solani* 3337 sequence of FliC and FliN. We used the draft and complete genomes of the 19 *D. solani* sequenced in this study, those of *D. solani* strains GBCC2040 and MK10, 15 *D. dianthicola* including the strains MIE32, MIE33, MIE34, CFBP1888, CFBP2015, CFBP2982, RNS04.9, RNS10.20.2A, RNS11.47.1A, DW04.9K, DS05.3.3, GBBC2039, IPO980, NBPPB3534 and NCPPB453, *D. dadantii* strains 3937, NCPPB898 and NCPPB3537, *D. chrysanthemi* strains NCPPB3533 and NCPPB516, and *D. zeae* strains Ech1591 and NCPPB2538.

```

                191
                |
FliC   D. solani 3296 TALVAAIDTALQTINTTKANIGATLNRF
           D. solani TALVAAIDTALQTINTAKANIGATLNRF
           D. dianthicola TALVSAIDTALQTINTAKANIGATLNRF
           D. dadantii TALVSAIDTALQTINTAKANIGATLNRF
           D. zeae TALVAAIDTALQTVNTAKANIGATLNRF
           D. chrysanthemi TALVQAVDAALVKINNAKANIGATLNRF

```

```

                101
                |
FliN   D. solani 3296 MINGYLIAQGE-VVVSDKYGVRITDIIT
           D. solani MINGYLIAQGEVVVSDKYGVRITDIIT
           D. dianthicola MINGYLIAQGEVVVSDKYGVRITDIIT
           D. dadantii MINGYLIAQGEVVVSDKYGVRITDIIT
           D. zeae MINGYLIAQGEVVVSDKYGVRITDIIT
           D. chrysanthemi MINGYLIAQGEVVVSDKYGVRITDIIT

```

Table III.S1 *Dickeya solani* strains used in this study

Organisms	Code	Host&Environment	Geography	Year
<i>D. solani</i> PPO 9019	Dsl 9019	Muscari	Netherlands	2006
<i>D. solani</i> PPO 9134	Dsl 9134	Hyacinth	Netherlands	2008
<i>D. solani</i> RNS 05.1.2A	Dsl 0512	Potato	France	2005
<i>D. solani</i> RNS 07.7.3B	Dsl 07-7	Potato	France	2007
<i>D. solani</i> RNS 08.23.3.1A	Dsl 3337	Potato	France	2008
<i>D. solani</i> RNS 10.68.1A	Dsl 1068	Potato	France	2010
<i>D. solani</i> RNS 12.6.1A	Dsl 12-6	Potato	Switzerland	2012
<i>D. solani</i> RNS GRC77	Dsl 3296	Potato	Israël	2008
<i>D. solani</i> RNS PPL0433	Dsl F8	Potato	Finland	2004
<i>D. solani</i> RNS 06/026	Dsl MIE35	Potato	Switzerland	2005
<i>D. solani</i> Ds0432.1	Dsl 0432.1	Potato	Finland	2004
<i>D. solani</i> MK16	Dsl MK16	River water	Scotland	-
<i>D. solani</i> 07-1E	Dsl 07E	Potato	France	2007
<i>D. solani</i> 10-06-2A	Dsl10062A	Potato	France	2010
<i>D. solani</i> 10-27-2A	Dsl10272A	Potato	Switzerland	2010
<i>D. solani</i> 10-54-2B	Dsl10542B	Potato	France	2010
<i>D. solani</i> CC2187	Dsl 2187	Potato	Ireland	-
<i>D. solani</i> CC2276	Dsl 2276	Potato	Poland	-
<i>D. solani</i> CC3239	Dsl 3239	Potato	United Kingdom	-
<i>D. solani</i> IPO2222 ^T	Dsl IPO2222	Potato	Netherlands	2007

Table III.S2 Sequencing data and mappings on the Dsl 3337 genome

Strain	Total reads	Mapped reads (%)	Unmapped reads(%)	mean coverage
Dsl0512	37 334 298	92.34	7.66	700
Dsl07-7	41 013 624	99.97	0.03	800
Dsl9019	48 631 564	98.08	1.92	900
Dsl9134	41 625 672	99.53	0.47	800
Dsl3296	40 769 110	99.45	0.55	800
DslF8	46 613 616	99.58	0.42	900
Dsl12-6	44 112 694	99.84	0.16	850
Dsl0432.1	23 156 694	99.58	0.42	450
Dsl1068	38 735 222	99.98	0.02	750
DslMK16	19 253 116	99.62	0.38	350
DslMI35	18 400 330	99.59	0.41	350
Dsl 07E	11 857 158	99.28	0.72	400
Dsl10062A	12 928 577	99.35	0.65	440
Dsl10272B	11 889 376	99.41	0.59	400
Dsl10542B	15 319 899	99.90	0.10	520
Dsl 2187	14 653 859	99.15	0.85	500
Dsl 2276	17 475 360	99.54	0.46	590
Dsl 3239	17 475 360	99.37	0.63	516
Dsl IPO2222	16 403 141	99.90	0.10	543

Table III.S3 Variants distribution on the strains vs. Dsl 3337

Strains	Variants count	Variants out CDS	Variants in CDS	Non synonymous variants
Dsl0512	37493	5350	32143	7234
Dsl07-7	1454	146	1308	292
Dsl9019	3433	331	3102	566
Dsl9134	2815	289	2526	542
Dsl3296	85	19	66	18
DslF8	65	18	47	13
Dsl12-6	63	20	43	12
Dsl0432.1	76	18	58	11
Dsl1068	70	18	52	13
DslMK16	62	19	43	12
DslMI35	75	19	56	14
Dsl 07E	49	14	35	6
Dsl10062A	48	14	34	10
Dsl10272B	43	13	30	6
Dsl10542B	45	12	33	8
Dsl 2187	51	16	35	8
Dsl 2276	44	13	31	8
Dsl 3239	52	16	36	10
Dsl IPO2222	49	15	34	8

Table III.S4 Non-synonymous variants: this table shows the unique and the shared variants within homogenous *D. solani* strains (MK16, MIE35, 0432.1, 12-6, F8, 1068, 3296, 07E, 10062A, 10272B, 10542B, 2187, 2276, 3239, IPO2222).

Position	CDS affected	Number of strains
1315848	Aconitate hydratase 2	8
1315885	Aconitate hydratase 2	15
1316262	Aconitate hydratase 2	15
1466994	Putative sensory histidine kinase YfhA	2
1471936	Phosphoribosylformylglycinamide synthase	1(DsICC2276)
1598278	Putative sugar ABC transport system	2
1704637	RNA polymerase sigma factor RpoS	1(DsI3296)
2492465	hypothetical protein	7
2492472	hypothetical protein	10
2495908	Exo-poly-alpha-D-galacturonidase precursor	14
2496505	Exo-poly-alpha-D-galacturonidase precursor	7
2754775	rRNA small subunit 7-methylguanosine (m7G)	1(DsIMK16)
2930935	VfmB protein	2
3776452	hypothetical protein	1(DsIMIE35)
3797586	putative RTX toxin	11
3802848	putative RTX toxin	7
3803181	putative RTX toxin	7
3803457	putative RTX toxin	8
4215221	Cytochrome d ubiquinol oxidase subunit I	1(DsIIPO2222)
4325693	VgrG protein	15
441030	Putative large exoprotein	1(DsICC3239)
441084	Putative large exoprotein	1(DsICC3239)
441285	Putative large exoprotein	1(DsICC3239)
4660124	Malonate decarboxylase alpha subunit	1(DsICC2187)
4889553	Fructose-specific phosphocarrier protein HPr	1(DsI10542B)
4917738	Methyl-accepting chemotaxis protein I	3
4917893	Methyl-accepting chemotaxis protein I	3
4918002	Methyl-accepting chemotaxis protein I	1(DsICC2276)
521861	Tripeptide aminopeptidase	1(DsIMK16)
952985	Flagellar biosynthesis protein FliC	1(DsI3296)
966036	Flagellar motor switch protein FliN	1(DsI3296)

Table III.S5 Other genomes used in this study.

Organisms	Accession n°.	Code
<i>Dickeya dianthicola</i> NCPPB 453 ^T	CM001841.1	Ddi 453 ^T
<i>Dickeya dianthicola</i> NCPPB 2039	CM001838.1	Ddi2039
<i>Dickeya dianthicola</i> RNS 04.9	KQ046817.1	Ddi RNS 04.9
<i>Dickeya chrysanthemi</i> NCPPB 516	CM001904.1	Dch 516
<i>Dickeya chrysanthemi</i> NCPPB3533	CM001981.1	Dch 3533
<i>Dickeya zaeae</i> NCPPB 3532	CM001858.1	Dz 3532
<i>Dickeya zaeae</i> NCPPB2538	CM001977.1	Dz 2538
<i>Dickeya dadantii</i> 3937	CP002038.1	Dda 3937
<i>Dickeya dadantii</i> NCPPB 898 ^T	CM001976.1	Dda898 ^T
<i>Pectobacterium atrosepticum</i> SCRI1043	BX950851.1	Pa SCRI1043
<i>Pectobacterium carotovorum</i> subsp. <i>carotovorum</i> PC1	CP001657.1	Pcc PC1

III.8. Acknowledgements

We thank Robert Dees (Wageningen UR/Applied Plant Research) for the gift of the ornamental strains *D. solani* PPO9019 and PPO9134, Minna Pirhonen (Department of Applied Biology, University of Helsinki) for providing PPL0433 (=F8), Leah Tsrer (Gilat Research Center, Agricultural Research Organisation) for providing GRC77 (EU3296) and Yves Dessaux (I2BC, CNRS) for his help in the manuscript editing. This work was supported by a cooperative project between France and Morocco (PRAD 14-02, Campus France n° 30229 ZK), the University Paris-Saclay (Co-tutelle funding), the excellence grant (n°H011/007) awarded by the Ministry of Higher education of Morocco, a collaborative project between CNRS(Gif sur Yvette) and FN3PT-RD3PT (Paris), the High Impact Research Grant (UM.C/625/1/HIR/MOHE/CHAN/14/01, Grant number A-000001-50001 to KGC) and the French-Malaysian exchange program awarded by French Embassy of Malaysia.

III.9. References

1. Samson R: Transfer of *Pectobacterium chrysanthemi* (Burkholder *et al.* 1953) Brenner *et al.* 1973 and *Brenneria paradisiaca* to the genus *Dickeya* gen. nov. as *Dickeya chrysanthemi* comb. nov. and *Dickeya paradisiaca* comb. nov. and delineation of four novel species, *Dickeya dadantii* sp. nov., *Dickeya dianthicola* sp. nov., *Dickeya dieffenbachiae* sp. nov. and *Dickeya zaeae* sp. nov. Int J Syst Evol Microbiol 2005, 55:1415–1427.

2. Gardan L: Elevation of three subspecies of *Pectobacterium carotovorum* to species level: *Pectobacterium atrosepticum* sp. nov., *Pectobacterium betavasculorum* sp. nov. and *Pectobacterium wasabiae* sp. nov. *Int J Syst Evol Microbiol* 2003, 53:381–391.
3. Collmer A, Keen NT: The Role of Pectic Enzymes in Plant Pathogenesis. *Annu Rev Phytopathol* 1986, 24:383–409.
4. Toth IK, van der Wolf JM, Saddler G, Lojkowska E, Hélias V, Pirhonen M, Tsrer Lahkim L, Elphinstone JG: *Dickeya* species: an emerging problem for potato production in Europe: *Dickeya* spp. on potato in Europe. *Plant Pathol* 2011, 60:385–399.
5. Van der Wolf JM, Nijhuis EH, Kowalewska MJ, Saddler GS, Parkinson N, Elphinstone JG, Pritchard L, Toth IK, Lojkowska E, Potrykus M, Waleron M, de Vos P, Cleenwerck I, Pirhonen M, Garland L, Helias V, Pothier JF, Pfluger V, Duffy B, Tsrer L, Manulis S: *Dickeya solani* sp. nov., a pectinolytic plant-pathogenic bacterium isolated from potato (*Solanum tuberosum*). *Int J Syst Evol Microbiol* 2014, 64(Pt 3):768–774.
6. Czajkowski R, de Boer WJ, van der Zouwen PS, Kastelein P, Jafra S, de Haan EG, van den Bovenkamp GW, van der Wolf JM: Virulence of “*Dickeya solani*” and *Dickeya dianthicola* biovar-1 and -7 strains on potato (*Solanum tuberosum*): Comparison of *Dickeya* spp. on potato. *Plant Pathol* 2013, 62:597–610.
7. Potrykus M, Golanowska M, Hugouvieux-Cotte-Pattat N, Lojkowska E: Regulators Involved in *Dickeya solani* Virulence, Genetic Conservation, and Functional Variability. *Mol Plant Microbe Interact* 2014, 27:700–711.
8. Czajkowski R, de Boer WJ, Velvis H, van der Wolf JM: Systemic colonization of potato plants by a soilborne, green fluorescent protein-tagged strain of *Dickeya* sp. biovar 3. *Phytopathology* 2010, 100:134–142.
9. Laurila J, Ahola V, Lehtinen A, Joutsjoki T, Hannukkala A, Rahkonen A, Pirhonen M: Characterization of *Dickeya* strains isolated from potato and river water samples in Finland. *Eur J Plant Pathol* 2008, 122:213–225.
10. Sławiak M, van Beckhoven JRCM, Speksnijder AGCL, Czajkowski R, Grabe G, van der Wolf JM: Biochemical and genetical analysis reveal a new clade of biovar 3 *Dickeya* spp. strains isolated from potato in Europe. *Eur J Plant Pathol* 2009, 125:245–261.
11. Parkinson N, Stead D, Bew J, Heeney J, Tsrer Lahkim L, Elphinstone J: *Dickeya* species relatedness and clade structure determined by comparison of *recA* sequences. *Int J Syst Evol Microbiol* 2009, 59(Pt 10):2388–2393.
12. Van Vaerenbergh J, Baeyen S, De Vos P, Maes M: Sequence Diversity in the *Dickeya fliC* Gene: Phylogeny of the *Dickeya* Genus and TaqMan® PCR for “*D. solani*”, New Biovar 3 Variant on Potato in Europe. *PLoS ONE* 2012, 7:e35738.
13. Garland L, Koskinen P, Rouhiainen L, Laine P, Paulin L, Auvinen P, Holm L, Pirhonen M: Genome Sequence of *Dickeya solani*, a New soft Rot Pathogen of Potato, Suggests its Emergence May Be Related to a Novel Combination of Non-Ribosomal Peptide/Polyketide Synthetase Clusters. *Diversity* 2013, 5:824–842.
14. Pédrón J, Mondy S, Raoul des Essarts Y, Van Gijsegem F, Faure D: Genomic and metabolic comparison with *Dickeya dadantii* 3937 reveals the emerging *Dickeya solani* potato pathogen to display distinctive metabolic activities and T5SS/T6SS-related toxin repertoire. *BMC Genomics* 2014, 15:283.
15. Smith MW, Feng D-F, Doolittle RF: Evolution by acquisition: the case for horizontal gene transfers. *Trends Biochem Sci* 1992, 17:489–493.

16. Smith JM, Smith NH, O'Rourke M, Spratt BG: How clonal are bacteria?. Proc Natl Acad Sci 1993, 90:4384–4388.
17. Ochman H, Lawrence JG, Groisman EA: Lateral gene transfer and the nature of bacterial innovation. Nature 2000, 405:299–304.
18. Koonin EV, Makarova KS, Aravind L: Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol 2001, 55:709–742.
19. Frost LS, Leplae R, Summers AO, Toussaint A: Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol 2005, 3:722–732.
20. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW: Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol Rev 2009, 33:376–393.
21. Choi SC, Rasmussen MD, Hubisz MJ, Gronau I, Stanhope MJ, Siepel A: Replacing and additive horizontal gene transfer in *Streptococcus*. Mol Biol Evol 2012, 29:3309–3320.
22. Khayi S, Mondy S, Beury-Cirou A, Moumni M, Helias V, Faure D: Genome Sequence of the Emerging Plant Pathogen *Dickeya solani* Strain RNS 08.23.3.1A. Genome Announc 2014, 2:e01270–13–e01270–13.
23. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S: MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol 2013, 30:2725–2729.
24. Darling ACE: Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. Genome Res 2004, 14:1394–1403.
25. Colston SM, Fullmer MS, Beka L, Lamy B, Gogarten JP, Graf J: Bioinformatic Genome Comparisons for Taxonomic and Phylogenetic Assignments Using *Aeromonas* as a Test Case. mBio 2014, 5:e02136–14.
26. Rogers TJ, Paton JC, Wang H, Talbot UM, Paton AW: Reduced virulence of an *fliC* mutant of Shiga-toxigenic *Escherichia coli* O113:H21. Infect Immun 2006, 74:1962–1966.
27. Jahn CE, Willis DK, Charkowski AO: The flagellar sigma factor *fliA* is required for *Dickeya dadantii* virulence. Mol Plant Microbe Interact 2008, 21:1431–1442.
28. He Y, Xu T, Fosshem LE, Zhang X-H: FliC, a flagellin protein, is essential for the growth and virulence of fish pathogen *Edwardsiella tarda*. PloS One 2012, 7:e45070.
29. Dong T, Schellhorn HE: Role of RpoS in Virulence of Pathogens. Infect Immun 2010, 78:887–897.
30. Pritchard L, Humphris S, Saddler GS, Parkinson NM, Bertrand V, Elphinstone JG, Toth IK: Detection of phytopathogens of the genus *Dickeya* using a PCR primer prediction pipeline for draft bacterial genome sequences: *Dickeya* diagnostics from draft bacterial genome sequences. Plant Pathol 2013, 62:587–596.
31. Winsor GL, Khaira B, Van Rossum T, Lo R, Whiteside MD, Brinkman FSL: The *Burkholderia* Genome Database: facilitating flexible queries and comparative analyses. Bioinforma Oxf Engl 2008, 24:2803–2804.
32. Coenye T, Mahenthiralingam E, Henry D, LiPuma JJ, Laevens S, Gillis M, Speert DP, Vandamme P: *Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including biocontrol and cystic fibrosis-related isolates. Int J Syst Evol Microbiol 2001, 51(Pt 4):1481–1490.

33. Ma B, Hibbing ME, Kim H-S, Reedy RM, Yedidia I, Breuer J, Breuer J, Glasner JD, Perna NT, Kelman A, Charkowski AO: Host Range and Molecular Phylogenies of the Soft Rot *Enterobacterial* Genera *Pectobacterium* and *Dickeya*. *Phytopathology* 2007, 97:1150–1163.
34. Glasner JD, Marquez-Villavicencio M, Kim H-S, Jahn CE, Ma B, Biehl BS, Rissman AI, Mole B, Yi X, Yang C-H, Dangl JL, Grant SR, Perna NT, Charkowski AO: Niche-specificity and the variable fraction of the *Pectobacterium* pan-genome. *Mol Plant-Microbe Interact MPMI* 2008, 21:1549–1560.
35. Jain R, Rivera MC, Lake JA: Horizontal gene transfer among genomes: The complexity hypothesis. *Proc Natl Acad Sci* 1999, 96:3801–3806.
36. Rivera MC, Jain R, Moore JE, Lake JA: Genomic evidence for two functionally distinct gene classes. *Proc Natl Acad Sci U S A* 1998, 95:6239–6244.
37. Alexandre A, Laranjo M, Young JPW, Oliveira S: *dnaJ* is a useful phylogenetic marker for *alphaproteobacteria*. *Int J Syst Evol Microbiol* 2008, 58:2839–2849.
38. Philippe H, Douady CJ: Horizontal gene transfer and phylogenetics. *Curr Opin Microbiol* 2003, 6:498–505.
39. Shapiro BJ: Signatures of Natural Selection and Ecological Differentiation in Microbial Genomes. In *Ecological Genomics*. Volume 781. Edited by Landry CR, Aubin-Horth N. Dordrecht: Springer Netherlands ; 2014:339–359.
40. Wilson K: Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol* 1987:2–4.
41. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM: DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007, 57:81–91.
42. Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S: MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* 2013, 30:2725–2729.

Chapitre IV. Genomic overview of the phytopathogen *Pectobacterium wasabiae* strain RNS 08.42.1A suggests horizontal acquisition of quorum-sensing genes

Pectobacterium wasabiae, initialement isolé à partir de raifort au Japon (Goto and Matsumoto, 1987a), est responsable de la maladie de la jambe noire et de la pourriture molle chez *Solanum tuberosum* (pomme de terre). Depuis 2007, le nombre de rapports concernant l'identification de *P. wasabiae* a augmenté. *P. wasabiae* a été récemment identifié dans plusieurs pays (Baghaee-Ravari *et al.*, 2011 ; De Boer *et al.*, 2012 ; Ma *et al.*, 2007 ; Moleleki *et al.*, 2013 ; Pitman *et al.*, 2010 ; Waleron *et al.*, 2013). Les études montrent que cette espèce était présente mais mal identifiée à cause du manque de marqueurs moléculaires fiables pour le diagnostic et l'identification. En effet, plusieurs souches de *Pectobacterium* spp on été reclassifiées en *P. wasabiae*, dont des isolats qui datent des années 1960 (Nabhan *et al.*, 2012 ; Nykyri *et al.*, 2012 ; Waleron *et al.*, 2013).

P. wasabiae RNS 08.42.1A a été isolé en 2008 à partir d'une plante de pomme de terre (Bintje CV) exprimant les symptômes de jambe noire lors d'une expérience en serre menée sur *Rhizoctonia solani* à INRA du Rheu.

Notre intérêt portait sur la caractérisation génomique de *P. wasabiae* RNS 08.42.1A par séquençage et assemblage de son génome et par utilisation d'approche de génomique comparative avec d'autre *P. wasabiae* présents dans NCBI et notamment la souche type *P. wasabiae* CFBP 3304^T.

Notre analyse de génomique comparative a révélé une forte synténie entre *P. wasabiae* RNS 08.42.1A et les espèces de *P. wasabiae* isolées à partir de pomme de terre. Les déterminants de virulence caractérisés dans *P. wasabiae* SCC3193 (Nykyri *et al.*, 2012) sont conservés dans toutes les souches *P. wasabiae* utilisées dans cette étude y compris *P. wasabiae* RNS 08.42.1A. Le trait caractéristique des *P. wasabiae* est l'absence du système de sécrétion de type III.

P. wasabiae RNS 08.42.1A présentait un complexe protéique Expl-ExpR1 qui a montré une faible similitude avec ceux identifiés dans les autres souches de *P. wasabiae*. En effet, il pourrait avoir été acquis par transfert horizontal à partir d'autres *Pectobacterium* spp.

Dans ce travail, j'ai réalisé l'assemblage génomique et les analyses de génomique comparative, la souche a été isolée et caractérisée par V. Hélias et A. Quêtu-Laurent. Y. Raoul des Essarts a réalisé l'extraction de l'ADN. Le texte a bénéficié des relectures de D. Faure, V. Hélias, M. Mourni et Y. Dessaux. Les résultats sont présentés sous la forme d'un article publié dans la revue *Genetica* (Khayl *et al.*, 2015).

Chapitre V. Discussion générale et perspectives

V.1. Discussion générale

Les phytopathogènes constituent un des risques majeurs pour la sécurité alimentaire mondiale. Ils sont capables de réduire de manière significative les rendements des cultures et d'accroître les pertes post-récolte lors des étapes de transport et de stockage. L'émergence de nouveaux phytopathogènes constitue ainsi une menace pour les cultures d'importance économique comme le riz, le blé, le maïs et la pomme de terre. Le risque d'insécurité alimentaire s'accroît dans le contexte des changements climatiques que connaît la Terre (Strange et Scott, 2005). Cette émergence peut en outre être favorisée par le réchauffement climatique qui influence largement les cultures, ainsi que la fécondité, la dispersion et la distribution des maladies et parasites des plantes (Schaad, 2008; van der Waals *et al.*, 2013). L'effet des conditions environnementales sur le développement et l'intensité des maladies illustre parfaitement, en pathologie, les relations triangulaires qui relient ; environnement, hôte et pathogène (Francl, 2001).

D'autres facteurs complexes entourent l'émergence des phytopathogènes, tels que les méthodes de pratiques culturales et les échanges internationaux des produits issus de l'agriculture (Anderson *et al.*, 2004; Schaad, 2008). De plus, le développement des agro-écosystèmes contribue également à l'émergence des pathogènes et à la modification de la population de pathogènes préexistante (Stukenbrock et McDonald, 2008). L'agro-écosystème constitue un milieu très dense propice aux échanges entre les plantes infectées et les plantes saines. Ainsi, la densité de la population pathogène augmente fortement à la faveur des conditions telles que la texture du sol, la fertilisation et l'irrigation. Ces conditions sont différentes de celles retrouvées dans l'écosystème naturel (Stukenbrock et McDonald, 2008).

L'insuffisance des ressources alimentaires dans les pays à forte croissance démographique et le coût croissant des céréales sur les marchés internationaux ont accru l'intérêt pour la culture de la pomme de terre comme denrée de première nécessité (DeFauw *et al.*, 2012). En effet, les surfaces de production sont en constante augmentation dans le monde entier avec plus de 19

millions d'hectares (DeFauw *et al.*, 2012 ; FAOSTAT, 2013). Cette culture est sujette à plus de 160 maladies d'origines diverses dont 10 maladies sont causées par des bactéries (Arora et Khurana, 2004; M. van der Wolf et H. De Boer, 2007). Les *Pectobacterium* et *Dickeya* spp. font partie des dix pathogènes de plantes les plus importants de point de vue économique et scientifique (Mansfield *et al.*, 2012), ils sont responsables de dégâts importants pour la filière de pomme de terre et d'autres cultures en réduisant la qualité ainsi que le rendement lors de la culture ou au moment du stockage des tubercules (Pérombelon, 2002).

Au cours de cette dernière décennie, plusieurs nouvelles espèces de *Pectobacterium* et *Dickeya* ont été identifiées dans différents pays associées aux maladies de la jambe noire et pourriture molle sur la pomme de terre. En 2004, Duarte et ces collègues ont identifié une nouvelle espèce émergente au Brésil *P. carotovorum* sous-espèce *brasiliensis*. Récemment une nouvelle espèce a été identifiée dans plusieurs pays européens, il s'agit de *Dickeya solani*. Celle-ci a contribué à l'augmentation de l'incidence des maladies causée par les *Dickeya* spp. (Toth *et al.*, 2011; van der Wolf *et al.*, 2014).

Comprendre comment ces phytopathogènes évoluent et s'adaptent pour devenir plus agressifs ou capables d'infecter de nouveaux hôtes constitue un enjeu majeur en agriculture. Ainsi, l'émergence de l'ère de la génomique a ouvert le champ à l'identification rapide des phytopathogènes émergents ainsi qu'à la compréhension des mécanismes qui président leur émergence (Thynne *et al.*, 2015). D'après Stukenbrock et McDonald (2008) l'émergence de nouveaux pathogènes est susceptible de se faire via 4 principaux scénarios ; la domestication des espèces végétales avec leurs pathogènes, les changements d'hôtes, le transfert horizontal de gènes (THG) et les hybridations génétiques.

Ce travail de doctorat contribue à la compréhension de la structure et de l'évolution génétique des deux pathogènes *D. solani* et *P. wasabiae* par approches de séquençage et génomique comparée. Dans le cas de *D. solani*, aucun génome complet et de bonne qualité n'était disponible

dans les bases de données publiques. C'est ainsi que le séquençage et l'assemblage du génome de *D. solani* 3337 que nous avons réalisé, constitue une étape importante dans l'étude de cette espèce émergente (**chapitre II**). La majorité des projets d'assemblage partiel de *D. solani* ont utilisé le séquençage 454 (NCBI). Dans le cas de la souche *D. solani* 3337, le séquençage a été réalisé en combinant deux technologies : PacBio et Illumina. Ainsi, nous avons pu faire un double contrôle de la séquence génomique et éliminer les erreurs de séquençage liées aux homopolymères (Loman *et al.*, 2012). Ces erreurs qui surviennent au niveau des régions d'homopolymères peuvent biaiser l'analyse du polymorphisme nucléotidique entre les souches de *D. solani* et générer de faux SNPs/InDels (**Chapitre III**). La séquence complète de *D. solani* 3337 a été assemblée sous forme d'un seul chromosome sans éléments extrachromosomiques. L'ensemble des séquences codantes et des gènes ribosomiaux ont été identifiées et caractérisées.

La séquence génomique de *D. solani* 3337 a été ensuite utilisée comme référence dans la recherche des SNPs/InDels et les analyses de génomique comparative par rapport à 19 souches de *D. solani* (**Chapitre III**). Ce travail est la première étude de génomique des populations chez *Dickeya* et *Pectobacterium* spp. Ces analyses ont débouché sur l'acquisition d'une nouvelle vision de la diversité et l'évolution génomique de *D. solani*. Le transfert horizontal de gènes (THG) par remplacement a été identifié chez plusieurs souches où des régions génomiques ont été complètement remplacées par leurs homologues provenant de *D. solani* (Intra-espèce) et de *D. dianthicola* (inter-espèce). Ces événements de THG inter et intra-espèces constituent une source de diversification de l'espèce, chose qui a été omise dans les études précédentes basées sur la séquence de quelques gènes. Les isolats qui ont été identifiés dans les différents pays européens étaient génétiquement homogènes et les isolements ont été effectués sur des échantillons de semences provenant de l'importation et aussi de pomme de terre de consommation ou de plants produit localement, ce qui renforce l'hypothèse de dissémination via les semences et le caractère

clonal du pathogène (Laurila *et al.*, 2008; Sławiak *et al.*, 2009; Van Vaerenbergh *et al.*, 2012; van der Wolf *et al.*, 2014). En revanche, notre analyse a révélé l'existence d'un sous-groupe au sein de *D. solani*. Les analyses génomiques réalisées sur les souches de *D. solani* (disponibles sur NCBI) ont montré que ce groupe était représenté jusqu'alors par la seule souche *D. solani* 0512. De plus, le THG par remplacement intra-espèce identifié chez *D. solani* 07-7 constitue une deuxième évidence qui renforce notre hypothèse sur l'existence d'un sous-groupe de *D. solani* 0512-*like*.

La souche *D. solani* MK16 a été isolée à partir des eaux d'une rivière alors qu'elle présente une forte similarité génétique comparée aux autres *D. solani* isolées sur pomme de terre. Également, elle a montré une forte agressivité sur tubercules de pomme de terre ce qui met l'accent sur les risques liées aux contaminations et dissémination des pathogènes par les eaux d'irrigation des cultures.

Les fragments d'ADN échangés entre *D. solani* et *D. dianthicola* pourraient ainsi avoir un effet sur le comportement de ces pathogènes vis à vis de leurs plantes hôtes et suggèrent une co-infection de la même plante hôte ainsi qu'une cohabitation des deux pathogènes dans la même niche écologique. Par contre, les analyses fonctionnelles n'ont pas montré de lien entre ces événements de THG et la variabilité d'agressivité de *D. solani* observée sur tubercules de pomme de terre. Généralement le THG affecte les îlots de pathogénicité (PAIs : *pathogenicity islands*) qui comprennent des clusters de gènes liés à la virulence (Hacker et Kaper, 2000; Juhas *et al.*, 2009), alors que dans notre étude, les régions RGT (*Replacing horizontal gene transfer*) comprenaient des gènes impliqués dans le métabolisme bactérien général, des gènes de ménage et des gènes qui codent pour des transporteurs membranaires qui semblaient jouer un rôle lié plus à l'adaptation écologique qu'aux processus de virulence du pathogène. Ces gènes sont en outre très conservés au sein de l'espèce. Ils sont liés à des fonctions qui pourraient être importantes pour la survie et l'adaptation des bactéries, ce qui favoriserait leur transfert horizontal au sein de

l'espèce ou entre des espèces phylogénétiquement proches telles que *D. solani* et *D. dianthicola*. Jain *et al.* (1999) ont rapporté que le THG entre bactéries affecte plus les gènes dits *operational genes* que les gènes dits *informational genes* tels que les gènes impliqués dans la transcription et la traduction.

En plus, du THG par remplacement, les analyses des lectures non mappées à la référence *D. solani* 3337, ont révélé que la souche *D. solani* 9019 a acquis un plasmide de 43 kpb par TGH additif (Choi *et al.*, 2012). Ce plasmide a été identifié chez *Burkholderia ambifaria* AMMD qui est une souche utilisée dans la lutte biologique contre les espèces pathogènes de *Pythium* (Coenye *et al.*, 2001 ; Winsor *et al.*, 2008). La stabilité de ce plasmide a été vérifiée par PCR sur colonie de *D. solani* 9019. Il comprend plusieurs gènes qui codent pour la machinerie assurant son maintien et son transfert. Ces gènes sont retrouvés chez les plasmides de la classe d'incompatibilité Inc-P (Adamczyk et Jagura-Burdzy, 2003 ; Jain et Srivastava, 2013). Les plasmides à large spectre d'hôtes (BHR : *Broad-Host-Range*) sont impliqués dans le transfert de gènes de virulence et de résistance aux antibiotiques (Jain et Srivastava, 2013) et ils sont largement retrouvés et transférés chez les bactéries à Gram négatif (Hayes, 2001).

Par ailleurs, Philippe et Douady (2003) ont souligné l'importance du THG dans la reconstruction phylogénétique des micro-organismes. Ainsi, les analyses des régions RGT ont mis en évidence des échanges entre *D. solani* et *D. dianthicola* de gènes utilisés dans la taxinomie, ce qui pourrait conduire à des assignations taxinomiques erronées de ces espèces. L'utilisation de plusieurs marqueurs moléculaires lors des processus de diagnostic et d'identification taxinomique de ces phytopathogènes s'avère donc d'une grande importance que ce soit sur le plan de lutte contre ces pathogènes ou pour limiter leur propagation via les semences.

Jusqu'à présent les mécanismes moléculaires qui sont à l'origine de l'adaptation de *D. solani* à l'hôte *S. tuberosum*, n'ont pas encore été déchiffrés. Par ailleurs, l'optimum élevé de température de croissance (35-37°C) (Janse et Ruissen, 1988) de cette espèce et le fait qu'elle soit très

virulente dans les régions tempérées telles qu'en Afrique du nord (Egypte), Australie et dans certaines régions du Pérou (Tsrer *et al.*, 2009), pourraient s'expliquer par une adaptation du pathogène aux changements climatiques (Hélias, 2008). Dans ce sens, l'émergence de plusieurs espèces a été associée au réchauffement climatique telle qu'*Acidovorax avenae* sous-espèce *avenae*, responsable des rayures bactériennes du riz qui a émergé en Italie en 2003, pendant les périodes de températures élevés (Cortesi *et al.*, 2005).

Le séquençage et l'assemblage du génome de *P. wasabiae* RNS 08.42.1A ont permis la caractérisation génomique de cette souche en décrivant les facteurs de virulence, les systèmes de sécrétion ainsi que les clusters souche-spécifiques par comparaison à d'autres *P. wasabiae* disponibles sur NCBI (**Chapitre IV**).

P. wasabiae a été isolé en premier sur le raifort (Goto et Matsumoto, 1987), puis il a été retrouvé comme pathogène sur pomme de terre en Amérique du nord (Ma *et al.*, 2007; De Boer *et al.*, 2012), en Nouvelle Zélande (Pitman *et al.*, 2010), en Europe (Finlande et Pologne) (Nykyri *et al.*, 2012) et en France (**Chapitre IV**). Dans ce travail, l'analyse du génome de *P. wasabiae* a montré une forte synténie entre toutes les souches de *P. wasabiae* isolées à partir de la pomme de terre et qui provenaient de différentes régions. Ce résultat est en concordance avec la structure de l'arbre phylogénétique (MLSA) qui regroupe les souches de pomme de terre dans un clade séparé de la souche type *P. wasabiae* CFBP 3344^T isolé du raifort. L'hypothèse émise pour expliquer la structure phylogénétique de cette espèce réside dans l'existence d'un groupe, associé à la maladie sur pomme de terre, qui a évolué au cours du temps par acquisition de plusieurs caractéristiques génétiques lui permettant de s'adapter à sa nouvelle niche écologique. L'analyse des séquences génomiques, des 5 souches de *P. wasabiae* étudiées, a montré que les isolats provenant de l'hôte pomme de terre partagent des caractéristiques génétiques que l'on ne retrouve pas chez la souche type *P. wasabiae* CFBP3344^T provenant du raifort, tels que les gènes impliqués dans le métabolisme des sucres maltose, maltodextrine et l'utilisation de l'histidine et

galactosamine. En se basant sur cette étude et celles de Nykyri *et al.* (2012) et Waleron *et al.* (2013), les analyses phylogénétiques ont mis en évidence la séparation entre les deux groupes de *P. wasabiae*. De plus les ANI (*Average nucleotide identity*) calculées entre la souche type et les autres souches indiquent des valeurs inférieures à 95% ce qui représente le seuil pour définir l'espèce bactérienne proposé par Goris *et al.* (2007). En effet, les souches associées à la maladie de la jambe noire et la pourriture molle sur pomme de terre pourraient constituer une sous-espèce de *P. wasabiae*.

Toutefois, d'après les travaux qui ont été faits sur *P. wasabiae* (Pitman *et al.*, 2010; Nabhan *et al.*, 2012; Nykyri *et al.*, 2012; Moleleki *et al.*, 2013), il n'est pas évident d'admettre que *P. wasabiae* constitue une nouvelle espèce qui a émergé sur pomme de terre en étendant son spectre d'hôte. En effet, cette espèce pourrait être présente comme pathogène de pomme de terre depuis longtemps alors que les moyens de caractérisation phénotypique, biochimique et génétique ne pouvaient pas l'identifier. Dans ce sens, plusieurs *Pectobacterium* ont été reclassifiés en *P. wasabiae* (Pitman *et al.*, 2010 ; Nykyri *et al.*, 2012)

V.2. Perspectives

À l'issue de ce travail de thèse, plusieurs perspectives se dégagent quant à la diversification des espèces de *Dickeya solani* et *Pectobacterium wasabiae*. Par approche transcriptomique, il serait intéressant d'étudier la variation de l'expression génique au niveau des régions génomiques concernées par le THG entre *D. solani* et *D. dianthicola* en comparant les souches qui ont eu du THG en condition de macération sur tubercules de pomme de terre et en prenant la souche *D. solani* 3337 ou IPO2222^T qui n'ont pas de THG comme référence. Par ailleurs, sachant que les essais d'agressivité ont été effectués sur tubercules, il serait important de réaliser des essais de virulence sur plante entière qui permettraient de relier des traits fonctionnels à la variabilité génétique au sein des souches de *D. solani*.

À notre connaissance, jusqu'à présent aucun plasmide n'a été identifié chez les espèces de *Dickeya*. La souche *D. solani* 9019 est la seule *Dickeya* possédant un plasmide d'environ 43 kpb qui a par ailleurs été identifié chez plusieurs espèces de *Burkholderia*. Il serait opportun d'étudier la permissivité des espèces de *Dickeya* à l'acquisition de ce plasmide ainsi que sa stabilité dans les bactéries receveuses. Les plasmides pourraient conférer des traits phénotypiques intéressants à leurs cellules hôtes. Dans cette optique, le rôle du plasmide pourrait être déterminé, notamment son implication dans la virulence ou la colonisation de *D. solani* 9019 sur *S. tuberosum*. Pour cela, il serait envisageable d'évaluer l'agressivité de la souche *D. solani* 9019 dépourvue de plasmide comparée au phénotype de la souche sauvage sur tubercule et sur plante entière.

Dans cette étude nous avons caractérisé une souche de *D. solani* 0512 qui est divergente des autres *D. solani*. Sur l'arbre MLSA, cette souche se situe dans une branche séparée des autres *D. solani* y compris la souche type *D. solani* IPO2222^T (fig III.1). Elle pourrait par conséquent être la représentante d'un sous-groupe au sein de l'espèce *D. solani*. Il serait intéressant de faire un échantillonnage plus large associé à des analyses taxinomiques plus profondes pour vérifier l'existence et la répartition de ce sous-groupe de *D. solani*.

Concernant *P. wasabiae*, les analyses phylogénétiques présentées dans ce travail et notamment celles des études précédentes, ont révélé une séparation entre les isolats de *S. tuberosum* et la souche type isolée du raifort. Il serait envisageable de faire un échantillonnage plus large afin d'obtenir des isolats d'autres plantes hôtes que la pomme de terre pour ensuite procéder à une caractérisation taxinomique plus poussée par une approche de génomique comparée.

Au terme de cette thèse, nous avons apporté de nouveaux éléments quant à la diversité génétique de *D. solani* et *P. wasabiae*, grâce aux outils de séquençage à haut débit et de génomique comparative, qui sont de plus en plus incontournables pour l'analyse structurelle et fonctionnelle des génomes bactériens. A travers ce travail, se dessine une méthodologie qui pourrait être utile

pour l'évaluation systématique de la diversité et de l'évolution génétique des espèces bactériennes émergentes associées aux plantes ou aux animaux. À partir d'une collection d'isolats bactériens (de régions géographiques différentes) de l'espèce étudiée, le séquençage des génomes permettrait une analyse de la diversité génétique par la recherche de SNPs/InDels et de THGs. Cette méthodologie nous permettrait d'acquérir des connaissances importantes quant à la diversité et l'évolution des pathogènes ainsi que des outils moléculaires de diagnostic.

Références bibliographiques

- Acosta Muniz, C., Jaillard, D., Lemaitre, B., and Boccard, F. (2007) *Erwinia carotovora* Evf antagonizes the elimination of bacteria in the gut of *Drosophila* larvae. *Cell. Microbiol.* 9: 106–119.
- Adamczyk, M. and Jagura-Burdzy, G. (2003) Spread and survival of promiscuous IncP-1 plasmids. *Acta Biochim. Pol.* 50: 425–453.
- Adriaenssens, E.M., Van Vaerenbergh, J., Vandenheuvel, D., Dunon, V., Ceysens, P.-J., De Proft, M., et al. (2012) T4-Related Bacteriophage LIMestone Isolates for the Control of Soft Rot on Potato Caused by “*Dickeya solani*.” *PLoS ONE* 7: e33227.
- Altermann, E. (2012) Tracing Lifestyle Adaptation in Prokaryotic Genomes. *Front. Microbiol.* 3:
- 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.
- Anderson, S. (1981) Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res.* 9: 3015–3027.
- Arora, R.K. and Khurana, S.M.P. (2004) Major Fungal and Bacterial Diseases of Potato and their Management. In, Mukerji, K.G. (ed), *Fruit and Vegetable Diseases*. Kluwer Academic Publishers, Dordrecht, pp. 189–231.
- Baghaee-Ravari, S., Rahimian, H., Shams-Bakhsh, M., Lopez-Solanilla, E., Antúñez-Lamas, M., and Rodríguez-Palenzuela, P. (2011) Characterization of *Pectobacterium* species from Iran using biochemical and molecular methods. *Eur. J. Plant Pathol.* 129: 413–425.
- Barbey, C., Crépin, A., Bergeau, D., Ouchiha, A., Mijouin, L., Taupin, L., et al. (2013) In Planta Biocontrol of *Pectobacterium atrosepticum* by *Rhodococcus erythropolis* Involves Silencing of Pathogen Communication by the Rhodococcal Gamma-Lactone Catabolic Pathway. *PLoS ONE* 8: e66642.
- Barbey, C., Crépin, A., Cirou, A., Budin-Verneuil, A., Orange, N., Feuilloley, M., et al. (2012) Catabolic pathway of gamma-caprolactone in the biocontrol agent *Rhodococcus erythropolis*. *J. Proteome Res.* 11: 206–216.
- Barnard, A.M.L. and Salmond, G.P.C. (2007) Quorum sensing in *Erwinia* species. *Anal. Bioanal. Chem.* 387: 415–423.
- Basset, A., Khush, R.S., Braun, A., Gardan, L., Boccard, F., Hoffmann, J.A., and Lemaitre, B. (2000) The phytopathogenic bacteria *Erwinia carotovora* infects *Drosophila* and activates an immune response. *Proc. Natl. Acad. Sci. U. S. A.* 97: 3376–3381.
- Basset, A., Tzou, P., Lemaitre, B., and Boccard, F. (2003) A single gene that promotes interaction of a phytopathogenic bacterium with its insect vector, *Drosophila melanogaster*. *EMBO Rep.* 4: 205–209.
- Bell, K.S., Sebaihia, M., Pritchard, L., Holden, M.T.G., Hyman, L.J., Holeva, M.C., et al. (2004) Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. *Proc. Natl. Acad. Sci.* 101: 11105–11110.
- Bentley, D.R. (2006) Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* 16: 545–552.
- Bétencourt, A. and Prunier, J.P. (1965) A propos de la pourriture sèche lenticellaire des tubercules de pommes de terre provoquée par *Erwinia Carotovora* (Jones) Holland. *Eur. Potato J.* 8: 230–242.
- De Boer, S.H. (2004) Blackleg of potato. *Plant Health Instr.* Doi:10.1094/PHI-I-2004-0712-01.

- De Boer, S.H., Li, X., and Ward, L.J. (2012) *Pectobacterium* spp. associated with bacterial stem rot syndrome of potato in Canada. *Phytopathology* 102: 937–947.
- De Boer, S.H. and Ward, L.J. (1995) PCR Detection of *Erwinia carotovora* subsp. *atroseptica* Associated with Potato Tissue. *Phytopathology* 85: 854.
- Boetzer, M. and Pirovano, W. (2012) Toward almost closed genomes with GapFiller. *Genome Biol.* 13: R56.
- Brenner, D.J., Steigerwalt, A.G., Miklos, G.V., and Fanning, G.R. (1973) Deoxyribonucleic Acid Relatedness Among *Erwiniae* and Other *Enterobacteriaceae*: The Soft-Rot Organisms (Genus *Pectobacterium* Waldee). *Int. J. Syst. Bacteriol.* 23: 205–216.
- Brochier, C., Forterre, P., and Gribaldo, S. (2004) Archaeal phylogeny based on proteins of the transcription and translation machineries: tackling the *Methanopyrus kandleri* paradox. *Genome Biol.* 5: R17.
- Brown, E.W., Davis, R.M., Gouk, C., and van der Zwet, T. (2000) Phylogenetic relationships of necrogenic *Erwinia* and *Brenneria* species as revealed by glyceraldehyde-3-phosphate dehydrogenase gene sequences. *Int. J. Syst. Evol. Microbiol.* 50 Pt 6: 2057–2068.
- Cappaert, M.R. and Powelson, M.L. (1987) The presence of *Erwinia carotovora* in surface and well water sources in central oregon. *Irrig. Crops Res. Cent. Or.* 1987 75.
- Chao, Y.C., Feng, C.T., and Ho, W.C. (2006) First Report of Aglaonema Bacterial Blight Caused by *Erwinia chrysanthemi* in Taiwan. *Plant Dis.* 90: 1358–1358.
- Chapelle, E., Alunni, B., Malfatti, P., Solier, L., Pédrón, J., Kraepiel, Y., and Van Gijsegem, F. (2015) A straightforward and reliable method for bacterial in planta transcriptomics: application to the *Dickeya dadantii* / *Arabidopsis thaliana* pathosystem. *Plant J.* 82: 352–362.
- Charkowski, A., Blanco, C., Condemine, G., Expert, D., Franza, T., Hayes, C., et al. (2012) The Role of Secretion Systems and Small Molecules in Soft-Rot *Enterobacteriaceae* Pathogenicity. *Annu. Rev. Phytopathol.* 50: 425–449.
- Charkowski, A.O. (2006) The soft rot *Erwinia*. In, *Plant-associated bacteria*. Springer, pp. 423–505.
- Charkowski, A.O., Lind, J., and Rubio-Salazar, I. (2014) Genomics of Plant-Associated Bacteria: The Soft Rot *Enterobacteriaceae*. In, Gross, D.C., Lichens-Park, A., and Kole, C. (eds), *Genomics of Plant-Associated Bacteria*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 37–58.
- Chin, C.-S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., et al. (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10: 563–569.
- Cirou, A., Mondy, S., An, S., Charrier, A., Sarrazin, A., Thoison, O., et al. (2012) Efficient Biostimulation of Native and Introduced Quorum-Quenching *Rhodococcus erythropolis* Populations Is Revealed by a Combination of Analytical Chemistry, Microbiology, and Pyrosequencing. *Appl. Environ. Microbiol.* 78: 481–492.
- Cirou, A., Raffoux, A., Diallo, S., Latour, X., Dessaux, Y., and Faure, D. (2011) Gamma-caprolactone stimulates growth of quorum-quenching *Rhodococcus* populations in a large-scale hydroponic system for culturing *Solanum tuberosum*. *Res. Microbiol.* 162: 945–950.
- Coenye, T., Mahenthiralingam, E., Henry, D., LiPuma, J.J., Laevens, S., Gillis, M., et al. (2001) *Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including biocontrol and cystic fibrosis-related isolates. *Int. J. Syst. Evol. Microbiol.* 51: 1481–1490.

- Compant, S., Duffy, B., Nowak, J., Clement, C., and Barka, E.A. (2005) Use of Plant Growth-Promoting Bacteria for Biocontrol of Plant Diseases: Principles, Mechanisms of Action, and Future Prospects. *Appl. Environ. Microbiol.* 71: 4951–4959.
- Cortesi, P., Bartoli, F., Pizzatti, C., Song, W.Y., and Schaad, N.W. (2005) First report of *Acidovorax avenae* spp. *avenae* on rice in Italy. *J. Plant Pathol.* 87: 76.
- Costechareyre, D., Chich, J.-F., Strub, J.-M., Rahbé, Y., and Condemine, G. (2013) Transcriptome of *Dickeya dadantii* Infecting *Acyrtosiphon pisum* Reveals a Strong Defense against Antimicrobial Peptides. *PLoS ONE* 8: e54118.
- Costechareyre, D., Dridi, B., Rahbé, Y., and Condemine, G. (2010) Cyt toxin expression reveals an inverse regulation of insect and plant virulence factors of *Dickeya dadantii*. *Environ. Microbiol.* 12: 3290–3301.
- Crépin, A., Barbey, C., Cirou, A., Tannières, M., Orange, N., Feuilloley, M., et al. (2012) Biological control of pathogen communication in the rhizosphere: A novel approach applied to potato soft rot due to *Pectobacterium atrosepticum*. *Plant Soil* 358: 27–37.
- Cronin, D., Moëne-Loccoz, Y., Fenton, A., Dunne, C., Dowling, D.N., and O’Gara, F. (2006) Ecological interaction of a biocontrol *Pseudomonas fluorescens* strain producing 2,4-diacetylphloroglucinol with the soft rot potato pathogen *Erwinia carotovora* subsp. *atroseptica*. *FEMS Microbiol. Ecol.* 23: 95–106.
- Czajkowski, R., de Boer, W.J., Velvis, H., and van der Wolf, J.M. (2010) Systemic colonization of potato plants by a soilborne, green fluorescent protein-tagged strain of *Dickeya* sp. biovar 3. *Phytopathology* 100: 134–142.
- Czajkowski, R., de Boer, W.J., van der Zouwen, P.S., Kastelein, P., Jafra, S., de Haan, E.G., et al. (2013a) Virulence of “*Dickeya solani*” and *Dickeya dianthicola* biovar-1 and -7 strains on potato (*Solanum tuberosum*). *Plant Pathol.* 62: 597–610.
- Czajkowski, R., Ozymko, Z., and Lojkowska, E. (2013b) Isolation and characterization of novel soilborne lytic bacteriophages infecting *Dickeya* spp. biovar 3 (“*D. solani*”). *Plant Pathol.* 63: 758–772.
- Czajkowski, R., Ozymko, Z., Zwirowski, S., and Lojkowska, E. (2014) Complete genome sequence of a broad-host-range lytic *Dickeya* spp. bacteriophage ϕ D5. *Arch. Virol.* 159: 3153–3155.
- Czajkowski, R., Pérombelon, M.C.M., Jafra, S., Lojkowska, E., Potrykus, M., van der Wolf, J.M., and Sledz, W. (2015) Detection, identification and differentiation of *Pectobacterium* and *Dickeya* species causing potato blackleg and tuber soft rot: a review: Detection of soft rot *Enterobacteriaceae*. *Ann. Appl. Biol.* 166: 18–38.
- Czajkowski, R., Pérombelon, M.C.M., van Veen, J.A., and van der Wolf, J.M. (2011) Control of blackleg and tuber soft rot of potato caused by *Pectobacterium* and *Dickeya* species: a review: Control of *Dickeya* and *Pectobacterium* species in potato. *Plant Pathol.* 60: 999–1013.
- Darrasse, A., Priou, S., Kotoujansky, A., and Bertheau, Y. (1994) PCR and restriction fragment length polymorphism of a pel gene as a tool to identify *Erwinia carotovora* in relation to potato diseases. *Appl. Environ. Microbiol.* 60: 1437–1443.
- DeFauw, S.L., He, Z., Larkin, R.P., and Mansour, S.A. (2012) Sustainable Potato Production and Global Food Security. In, He,Z., Larkin,R., and Honeycutt,W. (eds), Sustainable Potato Production: Global Case Studies. Springer Netherlands, Dordrecht, pp. 3–19.
- Diallo, S., Latour, X., Groboillot, A., Smadja, B., Copin, P., Orange, N., et al. (2009) Simultaneous and selective detection of two major soft rot pathogens of potato: *Pectobacterium atrosepticum*

(*Erwinia carotovora* subsp. *atrosepticum*) and *Dickeya* spp. (*Erwinia chrysanthemi*). Eur. J. Plant Pathol. 125: 349–354.

Dong, Y.H., Wang, L.H., Xu, J.L., Zhang, H.B., Zhang, X.F., and Zhang, L.H. (2001) Quenching quorum-sensing-dependent bacterial infection by an N-acyl homoserine lactonase. Nature 411: 813–817.

Dong, Y.H., Xu, J.L., Li, X.Z., and Zhang, L.H. (2000) AiiA, an enzyme that inactivates the acylhomoserine lactone quorum-sensing signal and attenuates the virulence of *Erwinia carotovora*. Proc. Natl. Acad. Sci. U. S. A. 97: 3526–3531.

Duarte, V., De Boer, S.H., Ward, L.J., and Oliveira, A.M.R. (2004) Characterization of atypical *Erwinia carotovora* strains causing blackleg of potato in Brazil. J. Appl. Microbiol. 96: 535–545.

During, K., Porsch, P., Fladung, M., and Lorz, H. (1993) Transgenic potato plants resistant to the phytopathogenic bacterium *Erwinia carotovora*. Plant J. 3: 587–598.

Dye, D. (1969) A taxonomic study of genus *Erwinia* .2. *carotovora* group. N. Z. J. Sci. 12: 81–8.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009) Real-Time DNA Sequencing from Single Polymerase Molecules. Science 323: 133–138.

Ekblom, R. and Wolf, J.B.W. (2014) A field guide to whole-genome sequencing, assembly and annotation. Evol. Appl. 7: 1026–1042.

Elphinstone, J.G. (1987) Soft Rot and Blackleg of Potato International Potato Center.

El Fahime, M and Ennaji, M.M. (2012) Évolution des techniques de séquençage. Technol. Lab. Vol 2 No 5 2007.

França, L.T.C., Carrilho, E., and Kist, T.B.L. (2002) A review of DNA sequencing techniques. Q. Rev. Biophys. 35.:

Francl, L.J. (2001) The Disease Triangle: A Plant Pathological Paradigm Revisited. Plant Health Instr.

Fravel, D.R. (1988) Role of Antibiosis in the Biocontrol of Plant Diseases*. Annu. Rev. Phytopathol. 26: 75–91.

Frechon, D., Exbrayat, P., Helias, V., Hyman, L.J., Jouan, B., Llop, P., et al. (1998) Evaluation of a PCR kit for the detection of *Erwinia carotovora* subsp. *atroseptica* on potato tubers. Potato Res. 41: 163–173.

Garbeva, P., Veen, J.A., and Elsas, J.D. (2004) Assessment of the diversity, and antagonism towards *Rhizoctonia solani* AG3, of *Pseudomonas* species in soil from different agricultural regimes. FEMS Microbiol. Ecol. 47: 51–64.

Gardan, L., Gouy, C., Christen, R., and Samson, R. (2003) Elevation of three subspecies of *Pectobacterium carotovorum* to species level: *Pectobacterium atrosepticum* sp. nov., *Pectobacterium betavasculorum* sp. nov. and *Pectobacterium wasabiae* sp. nov. Int. J. Syst. Evol. Microbiol. 53: 381–391.

Garlant, L., Koskinen, P., Rouhiainen, L., Laine, P., Paulin, L., Auvinen, P., et al. (2013) Genome Sequence of *Dickeya solani*, a New soft Rot Pathogen of Potato, Suggests its Emergence May Be Related to a Novel Combination of Non-Ribosomal Peptide/Polyketide Synthetase Clusters. Diversity 5: 824–842.

Gill, E.D., Schaerer, S., and Dupuis, B. (2014) Factors impacting blackleg development caused by *Dickeya* spp. in the field. Eur. J. Plant Pathol. 140: 317–327.

- Glasner, J.D., Marquez-Villavicencio, M., Kim, H.-S., Jahn, C.E., Ma, B., Biehl, B.S., et al. (2008) Niche-specificity and the variable fraction of the *Pectobacterium* pan-genome. *Mol. Plant-Microbe Interact.* MPMI 21: 1549–1560.
- Glasner, J.D., Yang, C.-H., Reverchon, S., Hugouvieux-Cotte-Pattat, N., Condemine, G., Bohin, J.-P., et al. (2011) Genome Sequence of the Plant-Pathogenic Bacterium *Dickeya dadantii* 3937. *J. Bacteriol.* 193: 2076–2077.
- Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57: 81–91.
- Goto, M. and Matsumoto, K. (1987) *Erwinia carotovora* subsp. *wasabiae* subsp. nov. Isolated from Diseased Rhizomes and Fibrous Roots of Japanese Horseradish (*Eutrema wasabi* Maxim.). *Int. J. Syst. Bacteriol.* 37: 130–135.
- de Haan, E.G., Dekker-Nooren, T.C.E.M., van den Bovenkamp, G.W., Speksnijder, A.G.C.L., van der Zouwen, P.S., and van der Wolf, J.M. (2008) *Pectobacterium carotovorum* subsp. *carotovorum* can cause potato blackleg in temperate climates. *Eur. J. Plant Pathol.* 122: 561–569.
- Hacker, J. and Kaper, J.B. (2000) Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* 54: 641–679.
- Harrison, M.D., Franc, G.D., Maddox, D.A., Michaud, J.E., and McCarter-Zorner, N.J. (1987) Presence of *Erwinia carotovora* in surface water in North America. *J. Appl. Bacteriol.* 62: 565–570.
- Harris, T.D., Buzby, P.R., Babcock, H., Beer, E., Bowers, J., Braslavsky, I., et al. (2008) Single-Molecule DNA Sequencing of a Viral Genome. *Science* 320: 106–109.
- Hauben, L., Moore, E.R., Vauterin, L., Steenackers, M., Mergaert, J., Verdonck, L., and Swings, J. (1998) Phylogenetic position of phytopathogens within the *Enterobacteriaceae*. *Syst. Appl. Microbiol.* 21: 384–397.
- Hawkes, J.G. and Francisco-Ortega, J. (1993) The early history of the potato in Europe. *Euphytica* 70: 1–7.
- Hayes, F. (2001) The Horizontal Gene Pool — Bacterial Plasmids and Gene Spread. *Heredity* 86: 251–252.
- Hélias, V. (2008) *Pectobacterium* spp. et *Dickeya* spp. de la pomme de terre nouvelle nomenclatures pour *Erwinia* spp. pdf. *Cah. Agric.* 17: 349–354.
- Hélias, V., Andrivon, D., and Jouan, B. (2000) Development of symptoms caused by *Erwinia carotovora* spp. *atroseptica* under field conditions and their effects on the yield of individual potato plants. *Plant Pathol.* 49: 23–32.
- Hommais, F., Oger-Desfeux, C., Van Gijsegem, F., Castang, S., Ligor, S., Expert, D., et al. (2008) PecS is a global regulator of the symptomatic phase in the phytopathogenic bacterium *Erwinia chrysanthemi* 3937. *J. Bacteriol.* 190: 7508–7522.
- Howarth, F.G. (1991) Environmental Impacts of Classical Biological Control. *Annu. Rev. Entomol.* 36: 485–509.
- Hugouvieux-Cotte-Pattat, N., Condemine, G., Nasser, W., and Reverchon, S. (1996) Regulation of pectinolysis in *Erwinia chrysanthemi*. *Annu. Rev. Microbiol.* 50: 213–257.
- Hyytiäinen, H. (2005) Regulatory networks controlling virulence in the plant pathogen *Erwinia carotovora* spp. *carotovora*. Thèse de doctorat, Université of Helsinki, 57p.

- Jahn, C.E., Selimi, D.A., Barak, J.D., and Charkowski, A.O. (2011) The *Dickeya dadantii* biofilm matrix consists of cellulose nanofibres, and is an emergent property dependent upon the type III secretion system and the cellulose synthesis operon. *Microbiology* 157: 2733–2744.
- Jain, A. and Srivastava, P. (2013) Broad host range plasmids. *FEMS Microbiol. Lett.* 348: 87–96.
- Jain, R., Rivera, M.C., and Lake, J.A. (1999) Horizontal gene transfer among genomes: The complexity hypothesis. *Proc. Natl. Acad. Sci.* 96: 3801–3806.
- Janse, J.D. and Ruissen, T. (1988) Characterization and Classification of *Erwinia chrysanthemi* Strains from Several Hosts in The Netherlands. *Phytopathology* 78: 800.
- Jiang, X., Sobetzko, P., Nasser, W., Reverchon, S., and Muskhelishvili, G. (2015) Chromosomal “Stress-Response” Domains Govern the Spatiotemporal Expression of the Bacterial Virulence Program. *mBio* 6: e00353–15.
- Juhas, M., van der Meer, J.R., Gaillard, M., Harding, R.M., Hood, D.W., and Crook, D.W. (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.* 33: 376–393.
- Kamysz, W., Krolicka, A., Bogucka, K., Ossowski, T., Lukasiak, J., and Lojkowska, E. (2005) Antibacterial Activity of Synthetic Peptides Against Plant Pathogenic *Pectobacterium* Species. *J. Phytopathol.* 153: 313–317.
- Kang, H.W., Kwon, S.W., and Go, S.J. (2003) PCR-based specific and sensitive detection of *Pectobacterium carotovorum* spp. *carotovorum* by primers generated from a URP-PCR fingerprinting-derived polymorphic band. *Plant Pathol.* 52: 127–133.
- Kastelein, P., Schepel, E.G., Mulder, A., Turkensteen, L.J., and Van Vuurde, J.W.L. (1999) Preliminary selection of antagonists of *Erwinia carotovora* subsp. *atroseptica* (Van Hall) Dye for application during green crop lifting of seed potato tubers. *Potato Res.* 42: 161–171.
- Keller, L. and Surette, M.G. (2006) Communication in bacteria: an ecological and evolutionary perspective. *Nat. Rev. Microbiol.* 4: 249–258.
- Kettani Halabi, M. (2012) Etude de la diversité de *Pectobacterium* spp et des effets induits par les lipopolysaccharides chez les plantes. Thèse de doctorat, Université Paris sud, France, 218p.
- Kettani-Halabi, M., Terta, M., Amdan, M., El Fahime, E.M., Bouteau, F., and Ennaji, M.M. (2013) An easy, simple inexpensive test for the specific detection of *Pectobacterium carotovorum* subsp. *carotovorum* based on sequence analysis of the *pmrA* gene. *BMC Microbiol.* 13: 176.
- Khay, S., Mondy, S., Beury-Cirou, A., Moumni, M., Helias, V., and Faure, D. (2014) Genome Sequence of the Emerging Plant Pathogen *Dickeya solani* Strain RNS 08.23.3.1A. *Genome Announc.* 2: e01270–13–e01270–13.
- Khay, S., Raoul des Essarts, Y., Quêtu-Laurent, A., Moumni, M., Hélias, V., and Faure, D. (2015) Genomic overview of the phytopathogen *Pectobacterium wasabiae* strain RNS 08.42.1A suggests horizontal acquisition of quorum-sensing genes. *Genetica* 143: 241–252.
- Kim, H.-S., Ma, B., Perna, N.T., and Charkowski, A.O. (2009) Phylogeny and virulence of naturally occurring type III secretion system-deficient *Pectobacterium* strains. *Appl. Environ. Microbiol.* 75: 4539–4549.
- Kim, M.H., Cho, M.S., Kim, B.K., Choi, H.J., Hahn, J.H., Kim, C., et al. (2012) Quantitative Real-Time Polymerase Chain Reaction Assay for Detection of *Pectobacterium wasabiae* Using YD Repeat Protein Gene-Based Primers. *Plant Dis.* 96: 253–257.

- Kloepper, J.W. (1983) Effect of seed piece inoculation with plant growth-promoting rhizobacteria on populations of *Erwinia carotovora* on potato roots and in daughter tubers. *Phytopathology* 73: 217–219.
- Kloepper, J.W., Leong, J., Teintze, M., and Schroth, M.N. (1980) *Pseudomonas* siderophores: A mechanism explaining disease-suppressive soils. *Curr. Microbiol.* 4: 317–320.
- Kõiv, V., Andresen, L., Broberg, M., Frolova, J., Somervuo, P., Auvinen, P., et al. (2013) Lack of RsmA-Mediated Control Results in Constant Hypervirulence, Cell Elongation, and Hyperflagellation in *Pectobacterium wasabiae*. *PLoS ONE* 8: e54248.
- Konstantinidis, K.T. and Tiedje, J.M. (2005) Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 102: 2567–2572.
- Koonin, E.V., Makarova, K.S., and Aravind, L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* 55: 709–742.
- Kwasiborski, A., Mondy, S., Beury-Cirou, A., and Faure, D. (2013) Genome Sequence of the *Pectobacterium atrosepticum* Strain CFBP6276, Causing Blackleg and Soft Rot Diseases on Potato Plants and Tubers. *Genome Announc.* 1.:
- Kwon, S.W., Go, S.J., Kang, H.W., Ryu, J.C., and Jo, J.K. (1997) Phylogenetic analysis of *Erwinia* species based on 16S rRNA gene sequences. *Int. J. Syst. Bacteriol.* 47: 1061–1067.
- Latour, X., Faure, D., Diallo, S., Cirou, A., Smadja, B., Dessaux, Y., and Orange, N. (2008) Lutte contre les maladies bactériennes de la pomme de terre dues aux *Pectobacterium* spp. (*Erwinia carotovora*). *Cah. Agric.* 17: 355–360.
- Laurila, J., Ahola, V., Lehtinen, A., Joutsjoki, T., Hannukkala, A., Rahkonen, A., and Pirhonen, M. (2008) Characterization of *Dickeya* strains isolated from potato and river water samples in Finland. *Eur. J. Plant Pathol.* 122: 213–225.
- Laurila, J., Hannukkala, A., Nykyri, J., Pasanen, M., Hélias, V., Garland, L., and Pirhonen, M. (2010) Symptoms and yield reduction caused by *Dickeya* spp. strains isolated from potato and river water in Finland. *Eur. J. Plant Pathol.* 126: 249–262.
- Lawrence, J.G. and Hendrickson, H. (2005) Genome evolution in bacteria: order beneath chaos. *Curr. Opin. Microbiol.* 8: 572–578.
- Lee, D.H., Kim, J.-B., Lim, J.-A., Han, S.-W., and Heu, S. (2014) Genetic Diversity of *Pectobacterium carotovorum* subsp. *brasiliensis* Isolated in Korea. *Plant Pathol. J.* 30: 117–124.
- Leff, B., Ramankutty, N., and Foley, J.A. (2004) Geographic distribution of major crops across the world: GLOBAL CROP DISTRIBUTION. *Glob. Biogeochem. Cycles* 18: n/a–n/a.
- Liu, H., Coulthurst, S.J., Pritchard, L., Hedley, P.E., Ravensdale, M., Humphris, S., et al. (2008) Quorum Sensing Coordinates Brute Force and Stealth Modes of Infection in the Plant Pathogen *Pectobacterium atrosepticum*. *PLoS Pathog.* 4: e1000093.
- Li, Z., Bai, X., Ruparel, H., Kim, S., Turro, N.J., and Ju, J. (2003) A photocleavable fluorescent nucleotide for DNA sequencing and analysis. *Proc. Natl. Acad. Sci. U. S. A.* 100: 414–419.
- Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J., and Pallen, M.J. (2012) Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.* 30: 434–439.
- Lund, B.M. and Kelman, A. (1977) Determination of the potential for development of bacterial soft rot of potatoes. *Am. Potato J.* 54: 211–225.

- Lyon, G.D. (1989) The biochemical basis of resistance of potatoes to soft rot *Erwinia* spp.: a review. *Plant Pathol.* 38: 313–339.
- Ma, B., Hibbing, M.E., Kim, H.-S., Reedy, R.M., Yedidia, I., Breuer, J., et al. (2007) Host Range and Molecular Phylogenies of the Soft Rot Enterobacterial Genera *Pectobacterium* and *Dickeya*. *Phytopathology* 97: 1150–1163.
- Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., et al. (2012) Top 10 plant pathogenic bacteria in molecular plant pathology: Top 10 plant pathogenic bacteria. *Mol. Plant Pathol.* 13: 614–629.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bembien, L.A., et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*.
- Medini, D., Donati, C., Tettelin, H., Masignani, V., and Rappuoli, R. (2005) The microbial pan-genome. *Curr. Opin. Genet. Dev.* 15: 589–594.
- Mendonca, M. de (1979) Endemic and Soilborne Nature of *Erwinia carotovora* var. *atroseptica*, a Pathogen of Mature Sugar Beets. *Phytopathology* 69: 1096.
- Van der Merwe, J.J., Coutinho, T.A., Korsten, L., and van der Waals, J.E. (2010) *Pectobacterium carotovorum* subsp. *brasilensis* causing blackleg on potatoes in South Africa. *Eur. J. Plant Pathol.* 126: 175–185.
- Miller, J.R., Koren, S., and Sutton, G. (2010) Assembly algorithms for next-generation sequencing data. *Genomics* 95: 315–327.
- Mills, A.A.S., (Bud) Platt, H.W., and Hurta, R.A.R. (2006) Sensitivity of *Erwinia* spp. to salt compounds in vitro and their effect on the development of soft rot in potato tubers in storage. *Postharvest Biol. Technol.* 41: 208–214.
- Moh, A.A. (2012) Etude des facteurs écologiques influençant la croissance et le développement des *Pectobacterium* spp. infectant les tubercules de pomme de terre. Thèse de doctorat, Université de Liège, Gembloux Agro-Bio Tech, Belgique. 113 p., 12 Tabl., 14 Fig.
- Moleleki, L.N., Onkendi, E.M., Mongae, A., and Kubheka, G.C. (2013) Characterisation of *Pectobacterium wasabiae* causing blackleg and soft rot diseases in South Africa. *Eur. J. Plant Pathol.* 135: 279–288.
- Molina, J.J. and Harrison, M.D. (1980) The role of *Erwinia carotovora* in the epidemiology of potato blackleg. II. The effect of soil temperature on disease severity. *Am. Potato J.* 57: 351–363.
- Monson, R., Burr, T., Carlton, T., Liu, H., Hedley, P., Toth, I., and Salmond, G.P.C. (2013) Identification of genes in the VirR regulon of *Pectobacterium atrosepticum* and characterization of their roles in quorum sensing-dependent virulence. *Environ. Microbiol.* 15: 687–701.
- M. van der Wolf, J. and H. De Boer, S. (2007) Bacterial pathogens of potato. In, *Potato Biology and Biotechnology*. Elsevier, pp. 595–617.
- Nabhan, S., Wydra, K., Linde, M., and Debener, T. (2012) The use of two complementary DNA assays, AFLP and MLSA, for epidemic and phylogenetic studies of pectolytic enterobacterial strains with focus on the heterogeneous species *Pectobacterium carotovorum*: Phylogenetic focus on soft-rot plant pathogens. *Plant Pathol.* 61: 498–508.
- Nassar, A., Darrasse, A., Lemattre, M., Kotoujansky, A., Dervin, C., Vedel, R., and Bertheau, Y. (1996) Characterization of *Erwinia chrysanthemi* by pectinolytic isozyme polymorphism and restriction fragment length polymorphism analysis of PCR-amplified fragments of pel genes. *Appl. Environ. Microbiol.* 62: 2228–2235.

- Naushad, H.S., Lee, B., and Gupta, R.S. (2014) Conserved signature indels and signature proteins as novel tools for understanding microbial phylogeny and systematics: identification of molecular signatures that are specific for the phytopathogenic genera *Dickeya*, *Pectobacterium* and *Brenneria*. *Int. J. Syst. Evol. Microbiol.* 64: 366–383.
- Neilands, J.B. (1995) Siderophores: structure and function of microbial iron transport compounds. *J. Biol. Chem.* 270: 26723–26726.
- Nomura, N., Yamashita, M., and Murooka, Y. (1996) Genetic organization of a DNA-processing region required for mobilization of a non-self-transmissible plasmid, pEC3, isolated from *Erwinia carotovora* subsp. *carotovora*. *Gene* 170: 57–62.
- Nykyri, J., Fang, X., Dorati, F., Bakr, R., Pasanen, M., Niemi, O., et al. (2014) Evidence that nematodes may vector the soft rot-causing enterobacterial phytopathogens. *Plant Pathol.* 63: 747–757.
- Nykyri, J., Niemi, O., Koskinen, P., Nokso-Koivisto, J., Pasanen, M., Broberg, M., et al. (2012) Revised Phylogeny and Novel Horizontally Acquired Virulence Determinants of the Model Soft Rot Phytopathogen *Pectobacterium wasabiae* SCC3193. *PLoS Pathog.* 8: e1003013.
- Ordax, M., Marco-Noales, E., López, M.M., and Biosca, E.G. (2006) Survival strategy of *Erwinia amylovora* against copper: induction of the viable-but-nonculturable state. *Appl. Environ. Microbiol.* 72: 3482–3488.
- Park, D.S., Shim, J.K., Kim, J.S., Kim, B.Y., Kang, M.J., Seol, Y.J., et al. (2006) PCR-based sensitive and specific detection of *Pectobacterium atrosepticum* using primers based on Rhs family gene sequences. *Plant Pathol.* 55: 625–629.
- Parkinson, J.S. and Kofoid, E.C. (1992) Communication Modules in Bacterial Signaling Proteins. *Annu. Rev. Genet.* 26: 71–112.
- Pasco C. (2005) Comportement variétal de pomme de terre vis-à-vis des pourritures molles : test su tubercules en conditions contrôlées. *Cah TechINRA* 2005 : 183-7.
- Pédron, J., Mondy, S., Raoul des Essarts, Y., Van Gijsegem, F., and Faure, D. (2014) Genomic and metabolic comparison with *Dickeya dadantii* 3937 reveals the emerging *Dickeya solani* potato pathogen to display distinctive metabolic activities and T5SS/T6SS-related toxin repertoire. *BMC Genomics* 15: 283.
- Pérez-Mendoza, D., Coulthurst, S.J., Humphris, S., Campbell, E., Welch, M., Toth, I.K., and Salmond, G.P.C. (2011) A multi-repeat adhesin of the phytopathogen, *Pectobacterium atrosepticum*, is secreted by a Type I pathway and is subject to complex regulation involving a non-canonical diguanylate cyclase. *Mol. Microbiol.* 82: 719–733.
- Pérombelon, M.C.M. (1992) Potato blackleg: epidemiology, host-pathogen interaction and control. *Neth. J. Plant Pathol.* 98: 135–146.
- Pérombelon, M.C.M. (2002) Potato diseases caused by soft rot *erwinias*: an overview of pathogenesis. *Plant Pathol.* 51: 1–12.
- Pérombelon, M.C.M. and Hyman, L.J. (1989) Sources and Pathways of Contamination of Potatoes by Soft Rot *Erwinias* in Scotland. In: Tjamos, E.C. and Beckman, C.H. (eds), *Vascular Wilt Diseases of Plants*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 143–151.
- Pérombelon, M.C.M. and Kelman, A. (1980) Ecology of the Soft Rot *Erwinias*. *Annu. Rev. Phytopathol.* 18: 361–387.
- Pérombelon, M.C.M. and Lowe, R. (1975) Studies on the initiation of bacterial soft rot in potato tubers. *Potato Res.* 18: 64–82.

- Pérombelon, M.C.M., Lumb, V.M., Zutra, D., Hyman, L.J., and Burnett, E.M. (1989) Factors Affecting Potato Blackleg Development. In, Tjamos, E.C. and Beckman, C.H. (eds), *Vascular Wilt Diseases of Plants*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 421–431.
- Peters, J., Sledz, W., Bergervoet, J.H.W., and van der Wolf, J.M. (2007) An enrichment microsphere immunoassay for the detection of *Pectobacterium atrosepticum* and *Dickeya dianthicola* in potato tuber extracts. *Eur. J. Plant Pathol.* 117: 97–107.
- Philippe, H. and Douady, C.J. (2003) Horizontal gene transfer and phylogenetics. *Curr. Opin. Microbiol.* 6: 498–505.
- Pirhonen, M., Heino, P., Helander, I., Harju, P., and Palva, E.T. (1988) Bacteriophage T4 resistant mutants of the plant pathogen *Erwinia carotovora*. *Microb. Pathog.* 4: 359–367.
- Pitman, A.R., Harrow, S.A., and Visnovsky, S.B. (2010) Genetic characterisation of *Pectobacterium wasabiae* causing soft rot disease of potato in New Zealand. *Eur. J. Plant Pathol.* 126: 423–435.
- Potrykus, M., Golanowska, M., Hugouvieux-Cotte-Pattat, N., and Lojkowska, E. (2014) Regulators Involved in *Dickeya solani* Virulence, Genetic Conservation, and Functional Variability. *Mol. Plant. Microbe Interact.* 27: 700–711.
- Powelson, M.L. (1985) Potato early dying disease in the Pacific Northwest caused by *Erwinia carotovora* pv. *carotovora* and *E. carotovora* pv. *atroseptica*. *Am. Potato J.* 62: 173–176.
- Priou, S. and Jouan, B. (1996) Priou S, Jouan B. Les maladies provoquées par les bactéries pathogènes du genre *Erwinia*. In : Rousselle P, Robert Y, Crosnier JC, eds. *Mieux comprendre la pomme de terre*. Paris (France : Inra éditions, 1996. Rousselle P Robert Crosnier JC Eds Mieux Compr. Pomme Terre Paris Fr. Inra Éditions 1996.
- Pritchard, L., Humphris, S., Saddler, G.S., Parkinson, N.M., Bertrand, V., Elphinstone, J.G., and Toth, I.K. (2013a) Detection of phytopathogens of the genus *Dickeya* using a PCR primer prediction pipeline for draft bacterial genome sequences: *Dickeya* diagnostics from draft bacterial genome sequences. *Plant Pathol.* 62: 587–596.
- Pritchard, L., Humphris, S., Saddler, G.S., Parkinson, N.M., Bertrand, V., Elphinstone, J.G., and Toth, I.K. (2013b) Detection of phytopathogens of the genus *Dickeya* using a PCR primer prediction pipeline for draft bacterial genome sequences: *Dickeya* diagnostics from draft bacterial genome sequences. *Plant Pathol.* 62: 587–596.
- Quevillon-Cheruel, S., Leulliot, N., Muniz, C.A., Vincent, M., Gallay, J., Argentini, M., et al. (2009) Evf, a virulence factor produced by the *Drosophila* pathogen *Erwinia carotovora*, is an S-palmitoylated protein with a new fold that binds to lipid vesicles. *J. Biol. Chem.* 284: 3552–3562.
- Raoul des Essarts, Y. (2015) Pathogénie de *Dickeya dianthicola* et *Dickeya solani* chez *Solanum tuberosum*, développement et évaluation de stratégie de lutte biologique. Thèse de doctorat, Université Paris sud, France, 153p.
- Raoul des Essarts, Y., Mondy, S., Hélias, V., and Faure, D. (2015) Genome Sequence of the Potato Plant Pathogen *Dickeya dianthicola* Strain RNS04.9. *Genome Announc.* 3: e00581–15.
- Rhodes, D.J. and Logan, C. (1987) A method for selecting fluorescent *Pseudomonads* inhibitory to seed tuber decay. *Potato Res.* 30: 603–611.
- Rodionov, D.A., Gelfand, M.S., and Hugouvieux-Cotte-Pattat, N. (2004) Comparative genomics of the KdgR regulon in *Erwinia chrysanthemi* 3937 and other gamma-proteobacteria. *Microbiol. Read. Engl.* 150: 3571–3590.
- Rombouts, F.M. (1972) Occurrence and properties of bacterial pectate lyases.

Rousselle, P., Robert, Y., and Crosnier, J.-C. (1996) La pomme de terre production, amélioration, ennemis et maladies, utilisations Institut national de la recherche agronomique : Institut technique des céréales et des fourrages, Paris.

Samson, R., Legendre, J.B., Christen, R., Fischer-Le Saux, M., Achouak, W., and Gardan, L. (2005) Transfer of *Pectobacterium chrysanthemi* (Burkholder et al. 1953) Brenner et al. 1973 and *Brenneria paradisiaca* to the genus *Dickeya* gen. nov. as *Dickeya chrysanthemi* comb. nov. and *Dickeya paradisiaca* comb. nov. and delineation of four novel species, *Dickeya dadantii* sp. nov., *Dickeya dianthicola* sp. nov., *Dickeya dieffenbachiae* sp. nov. and *Dickeya zeae* sp. nov. Int. J. Syst. Evol. Microbiol. 55: 1415–1427.

Sanger, F. and Coulson, A.R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. 94: 441–448.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. 74: 5463–5467.

Schaad, N.W. (2008) Emerging plant pathogenic bacteria and global warming. In, *Pseudomonas syringae* Pathovars and Related Pathogens—Identification, Epidemiology and Genomics. Springer, pp. 369–379.

Sengenès, J. (2012) Développement de méthodes de séquençage de seconde génération pour l'analyse des profils de méthylation de l'ADN. Thèse de doctorat, Université Pierre et Marie Curie, 216p.

Sharga, B.M. and Lyon, G.D. (1998) *Bacillus subtilis* BS 107 as an antagonist of potato blackleg and soft rot bacteria. Can. J. Microbiol. 44: 777–783.

Sjöblom, S. (2009) Quorum sensing in the plant pathogen *Erwinia carotovora* subsp. *carotovora*.

Sławiak, M., van Beckhoven, J.R.C.M., Speksnijder, A.G.C.L., Czajkowski, R., Grabe, G., and van der Wolf, J.M. (2009) Biochemical and genetical analysis reveal a new clade of biovar 3 *Dickeya* spp. strains isolated from potato in Europe. Eur. J. Plant Pathol. 125: 245–261.

Smadja, B., Latour, X., Trigui, S., Burini, J.F., Chevalier, S., and Orange, N. (2004) Thermodependence of growth and enzymatic activities implicated in pathogenicity of two *Erwinia carotovora* subspecies (*Pectobacterium* spp.). Can. J. Microbiol. 50: 19–27.

Smid, E.J., Jansen, A.H.J., and Gorris, L.G.M. (1995) Detection of *Erwinia carotovora* subsp. *atroseptica* and *Erwinia chrysanthemi* in potato tubers using polymerase chain reaction. Plant Pathol. 44: 1058–1069.

Sticklen, M.B. (2008) Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nat. Rev. Genet. 9: 433–443.

Strange, R.N. and Scott, P.R. (2005) Plant Disease: A Threat to Global Food Security. Annu. Rev. Phytopathol. 43: 83–116.

Stukenbrock, E.H. and McDonald, B.A. (2008) The Origins of Plant Pathogens in Agro-Ecosystems. Annu. Rev. Phytopathol. 46: 75–100.

Thomas, G. and Sansonetti, G. (2009) New light on a hidden treasure: International Year of the Potato 2008, an end-of-year review Food & Agriculture Org.

Thynne, E., McDonald, M.C., and Solomon, P.S. (2015) Phytopathogen emergence in the genomics era. Trends Plant Sci. 20: 246–255.

- Toth, I.K., Bertheau, Y., Hyman, L.J., Laplaze, L., López, M.M., McNicol, J., et al. (1999) Evaluation of phenotypic and molecular typing techniques for determining diversity in *Erwinia carotovora* subsp. *atroseptica*. J. Appl. Microbiol. 87: 770–781.
- Toth, I.K., Pritchard, L., and Birch, P.R. (2006) Comparative genomics reveals what makes an enterobacterial plant pathogen. Annu Rev Phytopathol 44: 305–336.
- Toth, I.K., van der Wolf, J.M., Saddler, G., Lojkowska, E., Hélias, V., Pirhonen, M., et al. (2011) *Dickeya* species: an emerging problem for potato production in Europe: *Dickeya* spp. on potato in Europe. Plant Pathol. 60: 385–399.
- Tsrur, L., Erlich, O., Lebiush, S., Hazanovsky, M., Zig, U., Slawiak, M., et al. (2009) Assessment of recent outbreaks of *Dickeya* sp. (syn. *Erwinia chrysanthemi*) slow wilt in potato crops in Israel. Eur. J. Plant Pathol. 123: 311–320.
- Uroz, S., Chhabra, S.R., Cámara, M., Williams, P., Oger, P., and Dessaux, Y. (2005) N-Acylhomoserine lactone quorum-sensing molecules are modified and degraded by *Rhodococcus erythropolis* W2 by both amidolytic and novel oxidoreductase activities. Microbiol. Read. Engl. 151: 3313–3322.
- Uroz, S., Oger, P.M., Chapelle, E., Adeline, M.-T., Faure, D., and Dessaux, Y. (2008) A *Rhodococcus* qsdA-Encoded Enzyme Defines a Novel Class of Large-Spectrum Quorum-Quenching Lactonases. Appl. Environ. Microbiol. 74: 1357–1366.
- Van Vaerenbergh, J., Baeyen, S., De Vos, P., and Maes, M. (2012) Sequence Diversity in the *Dickeya fliC* Gene: Phylogeny of the *Dickeya* Genus and TaqMan® PCR for “*D. solani*”, New Biovar 3 Variant on Potato in Europe. PLoS ONE 7: e35738.
- Venkatesh, B., Babujee, L., Liu, H., Hedley, P., Fujikawa, T., Birch, P., et al. (2006) The *Erwinia chrysanthemi* 3937 PhoQ sensor kinase regulates several virulence determinants. J. Bacteriol. 188: 3088–3098.
- van der Waals, J.E., Krüger, K., Franke, A.C., Haverkort, A.J., and Steyn, J.M. (2013) Climate Change and Potato Production in Contrasting South African Agro-Ecosystems 3. Effects on Relative Development Rates of Selected Pathogens and Pests. Potato Res. 56: 67–84.
- Waldee, E.L. (1942) Comparative Studies of Some Peritrichous Phytopathogenic Bacteria . Iowa State College.
- Waleron, M., Waleron, K., and Lojkowska, E. (2013) Occurrence of *Pectobacterium wasabiae* in potato field samples. Eur. J. Plant Pathol. 137: 149–158.
- Webb, L.E. and Wood, R.K.S. (1974) Infection of potato tubers with soft rot bacteria. Ann. Appl. Biol. 76: 91–98.
- Wegener, C.B. (2001) Transgenic potatoes expressing an *Erwinia* pectate lyase gene — results of a 4-year field experiment. Potato Res. 44: 401–410.
- Willats, W.G.T., McCartney, L., Mackie, W., and Knox, J.P. (2001) Pectin: cell biology and prospects for functional analysis. In, Carpita, N.C., Campbell, M., and Tierney, M. (eds), Plant Cell Walls. Springer Netherlands, Dordrecht, pp. 9–27.
- Williamson, N.R., Commander, P.M.B., and Salmond, G.P.C. (2010) Quorum sensing-controlled Evr regulates a conserved cryptic pigment biosynthetic cluster and a novel phenomycin-like locus in the plant pathogen, *Pectobacterium carotovorum*. Environ. Microbiol. 12: 1811–1827.
- Winslow, C.-E., Broadhurst, J., Buchanan, R.E., Krumwiede Jr, C., Rogers, L.A., and Smith, G.H. (1920) The families and genera of the bacteria: final report of the committee of the Society of

- American Bacteriologists on characterization and classification of bacterial types. *J. Bacteriol.* 5: 191.
- Winsor, G.L., Khaira, B., Van Rossum, T., Lo, R., Whiteside, M.D., and Brinkman, F.S.L. (2008) The *Burkholderia* Genome Database: facilitating flexible queries and comparative analyses. *Bioinforma. Oxf. Engl.* 24: 2803–2804.
- van der Wolf, J.M., Nijhuis, E.H., Kowalewska, M.J., Saddler, G.S., Parkinson, N., Elphinstone, J.G., et al. (2014) *Dickeya solani* sp. nov., a pectinolytic plant-pathogenic bacterium isolated from potato (*Solanum tuberosum*). *Int. J. Syst. Evol. Microbiol.* 64: 768–774.
- Xu, G.-W. and Gross, D.C. (1986) Selection of Fluorescent *Pseudomonads* Antagonistic to *Erwinia carotovora* and Suppressive of Potato Seed Piece Decay. *Phytopathology* 76: 414.
- Yaganza, E.-S., Rioux, D., Simard, M., Arul, J., and Tweddell, R.J. (2004) Ultrastructural alterations of *Erwinia carotovora* subsp. *atroseptica* caused by treatment with aluminum chloride and sodium metabisulfite. *Appl. Environ. Microbiol.* 70: 6800–6808.
- Yap, M.-N., Yang, C.-H., and Charkowski, A.O. (2008) The Response regulator HrpY of *Dickeya dadantii* 3937 regulates virulence genes not linked to the hrp cluster. *Mol. Plant-Microbe Interact. MPMI* 21: 304–314.
- Van Der Zaag, D.E. and Horton, D. (1983) Potato production and utilization in world perspective with special reference to the tropics and sub-tropics. *Potato Res.* 26: 323–362.

Annexes

Genome Sequence of the Emerging Plant Pathogen *Dickeya solani* Strain RNS 08.23.3.1A

Slimane Khayi,^{a,b} Samuel Mondy,^a Amélie Beury-Cirou,^c Mohieddine Mounni,^b Valérie Hélias,^d Denis Faure^a

Institut des Sciences du Végétal, CNRS, Gif-sur-Yvette, France^a; Université Moulay Ismail, Faculté des Sciences, Département de Biologie, Meknès, Morocco^b; Comité Nord Plant de Pomme de Terre (CNPPT), Semences, Innovation, Protection Recherche et Environnement (SIPRE), Achicourt, France^c; Fédération Nationale des Producteurs de Plantes de Pomme de Terre (FN3PT-RD3PT), UMT Innoplant (FN3PT-INRA), Le Rheu, France^d

S. K. and S. M. contributed equally to this work.

Here we present the genome sequence of *Dickeya solani* strain RNS 08.23.3.1A (PRI3337), isolated from *Solanum tuberosum*. *Dickeya solani*, recently described on potato cultures in Europe, is a proposed new taxon closely related to the *Dickeya dianthicola* and *Dickeya dadantii* species.

Received 30 December 2013 Accepted 6 January 2014 Published 30 January 2014

Citation Khayi S, Mondy S, Beury-Cirou A, Mounni M, Hélias V, Faure D. 2014. Genome sequence of the emerging plant pathogen *Dickeya solani* strain RNS 08.23.3.1A. *Genome Announc.* 2(1):e01270-13. doi:10.1128/genomeA.01270-13.

Copyright © 2014 Khayi et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Denis Faure, faure@isv.cnrs-gif.fr.

Bacteria belonging to the *Dickeya* genus cause soft rot disease on a wide range of plants (1, 2), particularly on ornamentals (*Dianthus*, *Dahlia*, *Kalanchoe*, *Pelargonium*, *Chrysanthemum*, *Parthenium*, *Dieffenbachiae*, and *St. ntpaulia*) and crops (chicory, banana, sunflower, rice, artichoke, pineapple, tomato, banana, maize, and potato). Infection is characterized by the maceration of plant tissues due to degradation of pectin, the major component of primary cell walls in plants (3). The resulting symptoms on potato stems are named blackleg and soft rot on tubers. On this plant host, disease symptoms caused by *Dickeya* spp. are similar to those caused by *Pectobacterium atrosepticum*. Three main *Dickeya* species, *D. dadantii*, *D. zaeae* and *D. dianthicola*, were previously described on potato. A new taxon, tentatively named *D. solani*, has been recently identified largely in Europe and beyond (4). The *D. solani* strain RNS 08.23.3.1A (PRI3337) was isolated from potatoes in France in 2008 (5).

Here we report the *de novo* genome assembly of *D. solani* strain RNS 08.23.3.1A using Illumina HiSeq 2000 v3 technology. Two libraries were constructed using the TruSeq SBS v3 sequencing kit, a shotgun (SG) paired-end library with a fragment size of 150 to 500 bp and a long-jumping-distance (LJD) mate-pair library with an insert size of 6,000 bp. The two libraries were sequenced using the 2×100-bp paired-end read module by Eurofins Genomics (France). Assembly of the sequences was carried out by use of the software CLC Genomics Workbench (v5.1) from CLC bio. Sequence reads with low-quality (limit 0.05), ambiguous nucleotides ($n \leq 2$) and sequence lengths of <30 nucleotides were dropped for assembly. In total, 5,682,625 mate-pair and 52,496,166 paired-end reads were obtained, corresponding to 419,377,725 and 4,903,141,904 bases, respectively. The average length was 73.8 bp for the mate-pair reads and 93.4 bp for the paired-end reads. The *de novo* assembly (length fraction, 0.5; similarity, 0.8) generated 42 contigs (>2,000 bp) with an average coverage of 1,075-fold. The average length of the contigs was 121,605 bp, and the largest contig was 483,425 bp, with an N_{50}

contig size of 299,659 bp. Scaffolding of the contigs was processed using SSPACE basic v2.0 (6). Two scaffolds with gaps from 367 bp to 4,771 bp in size were obtained. For finishing, the gaps were closed by the mapping of mate pairs using as reference the 5 kbp from each of the contig ends (read length, 0.9; identity, 0.95). Then, using homemade script and fastqselect.tcl from the MIRA3 package, the mapped reads for both orientations (R1 and R2) were retrieved and *de novo* assembled (read length, 0.5; identity, 0.8). The remaining gaps were resolved by Sanger sequencing of PCR amplicons.

The assembled genome of *D. solani* strain RNS 08.23.3.1A is composed of one circular chromosome, containing 4,923,734 bp. The G+C content was 56.11%. Gene prediction using the RAST (Rapid Annotation using Subsystem Technology) v4.0 automated pipeline (7) revealed the presence of 4,337 open reading frames.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [AMYI000000000](http://www.ncbi.nlm.nih.gov/nuccore/AMYI000000000). The version described in this paper is version AMYI01000000.

ACKNOWLEDGMENTS

This work was supported by a cooperative project between France and Morocco (PRAD 14-02, Campus France no. 30229 ZK) and a collaborative project between CNRS, FN3PT, and CNPPT-SIPRE.

REFERENCES

- Burkholder WD, McFadden LA, Dimock AW. 1953. A bacterial blight of chrysanthemums. *Phytopathology* 43:522–526.
- Samson R, Legendre JB, Christen R, Fischer-Le Saux M, Achouak W, Gardan L. 2005. Transfer of *Pectobacterium chrysanthemi* (Burkholder et al. 1953) Brenner et al. 1973 and *Brenneria paradisiaca* to the genus *Dickeya* a gen. nov. as *Dickeya chrysanthemi* comb. nov. and *Dickeya paradisiaca* comb. nov. and delineation of four novel species, *Dickeya dadantii* sp. nov., *Dickeya dianthicola* sp. nov., *Dickeya dieffenbachiae* sp. nov. and *Dickeya zaeae* sp. nov. *Int. J. Syst. Evol. Microbiol.* 55(Pt 4):1415–1427. <http://dx.doi.org/10.1099/ijs.0.02791-0>.

3. Hugouvieux-Cotte-Pattat N, Condemine G, Nasser W, Reverchon S. 1996. Regulation of pectinolysis in *Erwinia chrysanthemi*. *Annu. Rev. Microbiol.* 50:213–257. <http://dx.doi.org/10.1146/annurev.micro.50.1.213>.
4. Toth IK, van der Wolf JM, Saddler G, Lojkowska E, Pirhonen M, Tsror L, Elphinstone JG, Hélias V. 2011. *Dickey a* species: an emerging problem for potato production in Europe. *Rev. Plants Pathol.* 60:385–399. <http://dx.doi.org/10.1111/j.1365-3059.2011.02427.x>.
5. Hélias V. 2012. Jambé noire: evolution des souches et risques associés. *Pomme Terre Fr.* 580:48–49.
6. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27(4): 578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
7. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.

RESEARCH ARTICLE

Open Access



Population genomics reveals additive and replacing horizontal gene transfers in the emerging pathogen *Dickeya solani*

Slimane Khayi^{1,2}, Pauline Blin¹, Jacques Pédrón³, Teik-Min Chong⁴, Kok-Gan Chan⁴, Mohieddine Moumni², Valérie Hélias^{5,6}, Frédérique Van Gijsegem⁷ and Denis Faure^{1*}

Abstract

Background: *Dickeya solani* is an emerging pathogen that causes soft rot and blackleg diseases in several crops including *Solanum tuberosum*, but little is known about its genomic diversity and evolution.

Results: We combined Illumina and PacBio technologies to complete the genome sequence of *D. solani* strain 3337 that was used as a reference to compare with 19 other genomes (including that of the type strain IPO2222^T) which were generated by Illumina technology. This population genomic analysis highlighted an unexpected variability among *D. solani* isolates since it led to the characterization of two distinct sub-groups within the *D. solani* species. This approach also revealed different types of variations such as scattered SNP/InDel variations as well as replacing and additive horizontal gene transfers (HGT). Infra-species (between the two *D. solani* sub-groups) and inter-species (between *D. solani* and *D. dianthicola*) replacing HGTs were observed. Finally, this work pointed that genetic and functional variation in the motility trait could contribute to aggressiveness variability in *D. solani*.

Conclusions: This work revealed that *D. solani* genomic variability may be caused by SNPs/InDels as well as replacing and additive HGT events, including plasmid acquisition; hence the *D. solani* genomes are more dynamic than that were previously proposed. This work alerts on precautions in molecular diagnosis of this emerging pathogen.

Keywords: *Dickeya*, Soft rot, Potato, Population genomics, Horizontal gene transfer

Background

Pectinolytic bacteria belonging to the *Dickeya* and *Pectobacterium* genera are pathogens that cause soft rot and blackleg diseases in a wide range of plants and crops including *Solanum tuberosum* [1, 2]. These phytopathogens produce plant cell-wall degrading enzymes that are able to macerate the tuber and stem tissues, thus provoking the disease symptoms [3]. Since 2000s, the emerging *D. solani* species has been proposed as a contributor to the increased incidence of blackleg and soft rot diseases on potato crop in Europe and the Mediterranean basin [4]. The *D. solani* species has been officially described recently [5].

Little is known about the ecological and genetic traits that may support the relative success of *D. solani* in

invading potato fields [6, 7]. *D. solani* can initiate disease from a low inoculum level in warm climates and was described in some studies to spread more easily through vascular tissues than other *Dickeya* species [4, 8]. Besides classical intergenic spacers 16S-23S rDNA [9], several molecular studies have proposed different marker genes for the identification of *D. solani* strains collected from potato and ornamental plants, such as *dnaX* [10], *recA* [11] and *fliC* [12]. At whole genome level, genomic and metabolic comparisons of two *D. solani* strains Ds0432-1 (isolated in Finland) and 3337 (isolated in France) vs. *D. dadantii* 3937 indicated a conserved synteny between the two species, but also the presence of distinctive traits [13, 14]. *D. solani* and *D. dadantii* diverged in their battery of non-ribosomal peptide/polyketide synthase clusters, T5SS/T6SS-related toxin-antitoxin systems and several metabolic abilities. Some of these traits would contribute to the successful invasion of this pathogen [13, 14].

* Correspondence: denis.faure@i2bc.paris-saclay.fr

¹Institute for Integrative Biology of the Cell (I2BC), CNRS CEA Univ. Paris-Sud, Université Paris-Saclay, Saclay Plant Sciences, Avenue de la Terrasse, 91198 Gif-sur-Yvette cedex, France

Full list of author information is available at the end of the article

More recently, a reverse genetic approach revealed that the virulence master regulators are quite the same in *D. solani* and *D. dadantii* [7].

The analysis of population genome structure and dynamics, including additive or replacing horizontal gene transfer (HGT) may bring valuable clues on the mechanisms of emergence of *D. solani*. While additive HGT allows the acquisition of novel genes by a population [15–20], replacing HGT provokes the replacement of an allele by another from close relatives [21]. HGT events inform about the genome diversification and adaptation processes, but also on the companion populations that the pathogens met during the emergence and dissemination steps. Replacing HGT is also of a major stake in pathogen diagnostic, as it may provoke false identification when the alleles exchanged by replacing HGT are used as molecular taxonomic markers.

Here, we analyzed the whole genome polymorphism of 20 *D. solani* isolates, including the type strain IPO2222^T, collected from different geographic locations, dates of isolation and plant hosts. We combined Illumina and PacBio technologies to complete the 3337 *D. solani* strain genome that we used as a reference in the comparative genomics. While most strains belonged to a core-population that exhibited less than one hundred variant positions between two given genomes, some other genomes revealed massive replacing HGT from the companion pathogen *D. dianthicola* and a plasmid acquisition from *Burkholderia ambifaria*. Moreover, we were able to correlate SNPs in virulence genes with a decrease in aggressiveness, highlighting the power of genomics as a tool to reveal functional variability in *D. solani* population. To our knowledge this is the first study that reports whole genome analysis of a *D. solani* population and describes its diversity.

Results

Complete genome of the *D. solani* 3337

The *D. solani* 3337 genome was previously sequenced by Illumina technology using two libraries (mate-pair and paired-end) and *de novo* assembled in a high quality draft genome deposited at NCBI [22]. In this work, the 3337 *D. solani* genome was re-sequenced using PacBio technology. The PacBio sequencing generated six contigs (2 473 62 pb, 1 512 701 bp, 894 591 bp, 49 337 bp, 10 627 bp, and 4 290 bp) with an average 150 fold coverage. The published Illumina-scaffolding was confirmed and the remaining gaps were filled using the PacBio contigs. Hence, combining the Illumina and PacBio sets of sequences, we obtained a complete sequence of the unique circular chromosome (4 922 460 bp). The RAST annotation generated 4 536 CDS and 97 RNAs. The *D. solani* 3337 complete genome was used as a reference for comparative genomics.

Positioning the sequenced *D. solani* strains within the *Dickeya* genus

In addition to *D. solani* 3337, 19 *D. solani* strains including the type strain IPO2222^T were collected at different years and geographical locations (Additional file 1: Table S1) and their genomes sequenced by Illumina technology. All these draft and complete genomes were used in multi-locus sequence analysis (MLSA) and average nucleotide identity (ANI) calculation. For MLSA, the eleven concatenated *rpoD*, *gyrB*, *recA*, *rpoS*, *dnaX*, *dnaA*, *gapA*, *fusA*, *rplB*, *purA*, *gyrA* housekeeping genes (17 298 bp) were aligned to construct a relation-tree using Neighbor-Joining method, the evolutionary distances were computed using the Maximum Composite Likelihood method [23]. All the *D. solani* (Dsl) isolates were grouped in a same cluster that was separated from the other pectinolytic enterobacteria (Fig. 1). Noticeably, the strain Dsl 0512 was the unique strain that was consistently distant from the other *D. solani* strains. As previously reported [5], within the genus *Dickeya* the most related species to *D. solani* were *D. dadantii* and *D. dianthicola*. ANI values which were calculated using the strain Dsl 3337 as a reference were in accordance with the MLSA clustering. All the *D. solani* strains exhibited an ANI value

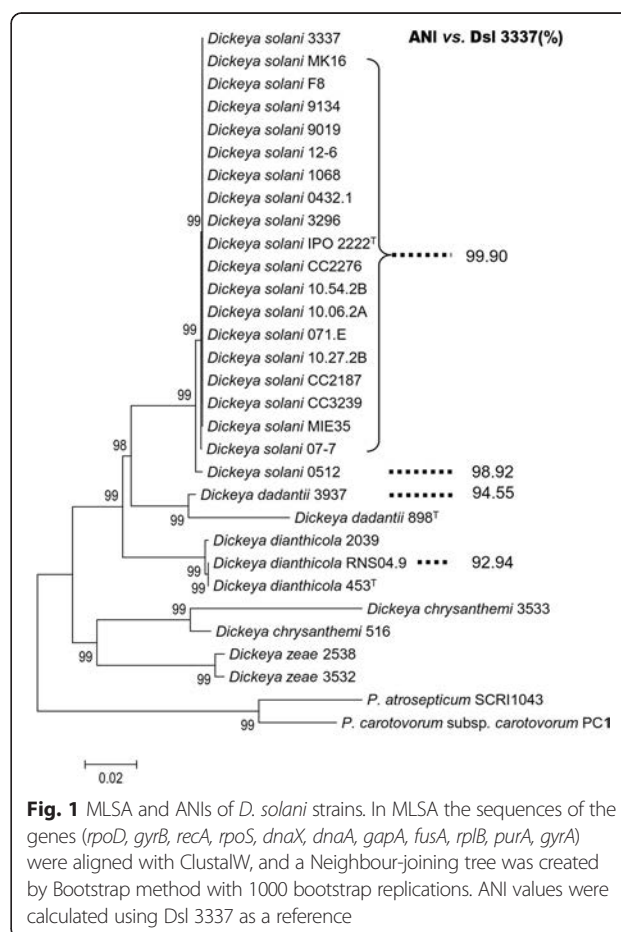


Fig. 1 MLSA and ANIs of *D. solani* strains. In MLSA the sequences of the genes (*rpoD*, *gyrB*, *recA*, *rpoS*, *dnaX*, *dnaA*, *gapA*, *fusA*, *rplB*, *purA*, *gyrA*) were aligned with ClustalW, and a Neighbour-joining tree was created by Bootstrap method with 1000 bootstrap replications. ANI values were calculated using Dsl 3337 as a reference

equal to or above 99.9 %, but that of Dsl 0512 was below 99 %. Among strains of the closest species, *D. dadantii* and *D. dianthicola* ANI values dropped to 94 % and 92 %, respectively.

Overview of the SNP and InDel variations in *D. solani* genomes

Illumina reads of the *D. solani* strains were mapped on the complete genome sequence of Dsl 3337. The percentage of mapped reads was above 99 % for all strains with the exception of Dsl 9019 (98.08 %) and 0512 (92.34 %) (Additional file 1: Table S2). The mapping vs. Dsl 3337, which reached a high mean coverage value (between 400 and 900), allowed us to identify variations (SNPs and InDels) in each of the genomes (Additional file 1: Table S3). According to the number of variations, the *D. solani* strains could be clustered into three groups. The first group, which we thereafter term as the core-population, encompassed most of the strains (including IPO2222^T and the reference Dsl 3337) with a variation number ranging from 43 to 85. In the second group were the strains Dsl 07-7, 9019 and 9134 with a variation number between 1454 and 3433. The third group consisted in the only strain Dsl 0512 with a very high variation number that reached 37493. RAST annotation of the strain Dsl 3337 was used to position the variations in or out coding DNA sequences (CDSs), as well as to identify non-synonymous variations in CDSs (Additional file 1: Table S3). Non-synonymous variations ranged between 14 and 21 % of the total number of variations, hence only 6 to 18 non-synonymous variations were identified in strains of the *D. solani* core-population (Additional file 1: Table S3).

Heterogeneous distribution of the SNP and InDel variations in *D. solani* genes

We refined our analysis by calculating the number of genes (CDSs) that were affected by SNPs and InDels as well as non-synonymous variations (Fig. 2a-b). In the core-population, 9 to 17 genes exhibited variations and about one half of them (4 to 10) harbored non-synonymous variations. In Dsl 07-7, 9019 and 9134, 56 to 144 genes were affected; and among them, 46 to 81 contained non-synonymous variations. In Dsl 0512, 2760 genes, hence half of the genome showed variations. To compare variation abundance in genes, a mean value of the number of all variations (synonymous and non-synonymous) per affected gene was calculated (Fig. 2c). In the core-population, this value ranged from 2 to 5. In Dsl 0512 and 07-7, it was similar (11 and 9, respectively) while the highest value was observed in Dsl 9134 and 9019 (45 and 42, respectively). Overall, these analyses revealed that Dsl 0512, 07-7, 9134 and 9019 harbored genes with different numbers of variations as compared to those in the

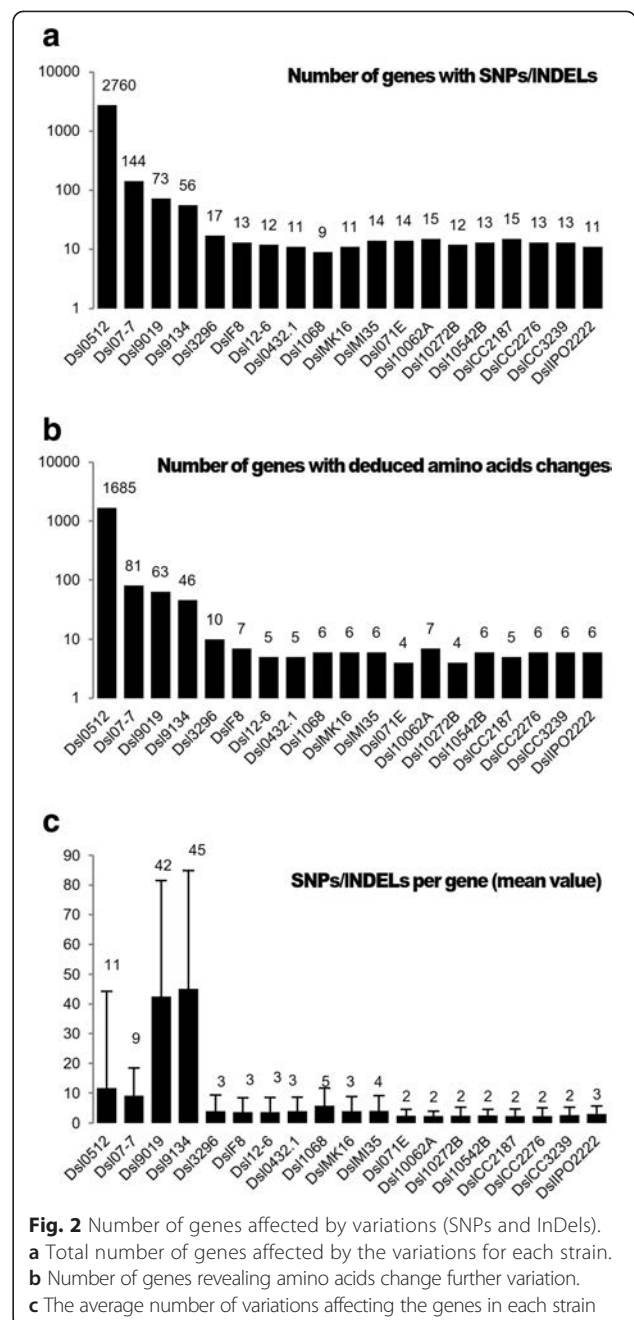
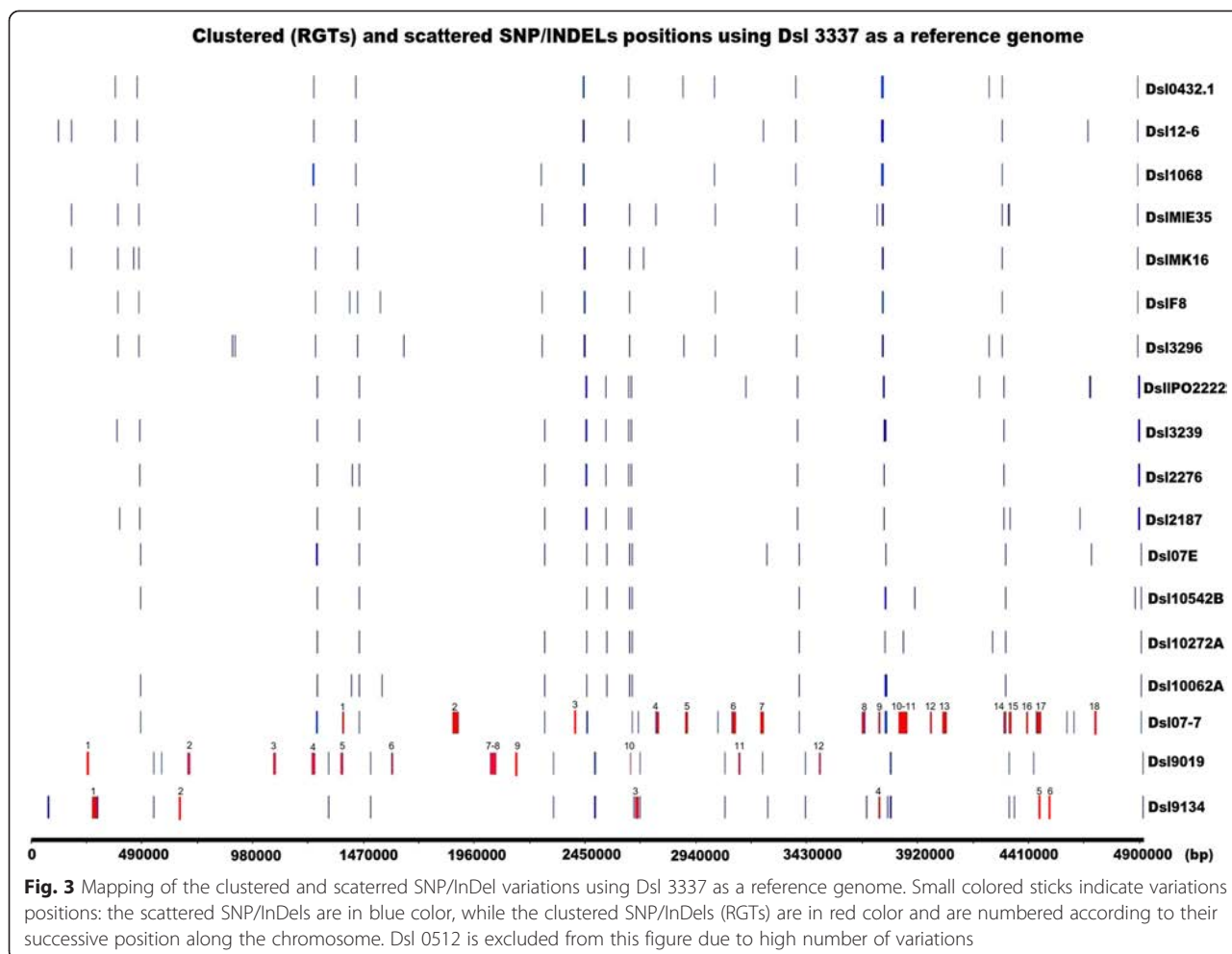


Fig. 2 Number of genes affected by variations (SNPs and InDels). **a** Total number of genes affected by the variations for each strain. **b** Number of genes revealing amino acids change further variation. **c** The average number of variations affecting the genes in each strain

core-population, suggesting putative HGTs from distinct sources.

Finally, all these different variations were positioned along the Dsl chromosome (Fig. 3). In the core-population, the rare variations appeared to be scattered with a mean distribution of 0.015 variations per kbp (when all SNPs and InDels were counted) and 0.012 variations per gene (when only SNPs and InDels in CDSs were counted). In Dsl 9134, 9019 and 07-7, most of the variations affected several tens of genes that are



clustered in distinct regions, while only a few variations remain scattered. In Dsl 0512, variations exhibited a genome wide distribution. In the next part of the work, the three types of SNP/InDel distribution (scattered, clustered and wide genome distribution) have been analyzed in details.

The mosaic genome of *D. solani* 0512 might define a novel *D. solani* sub-group

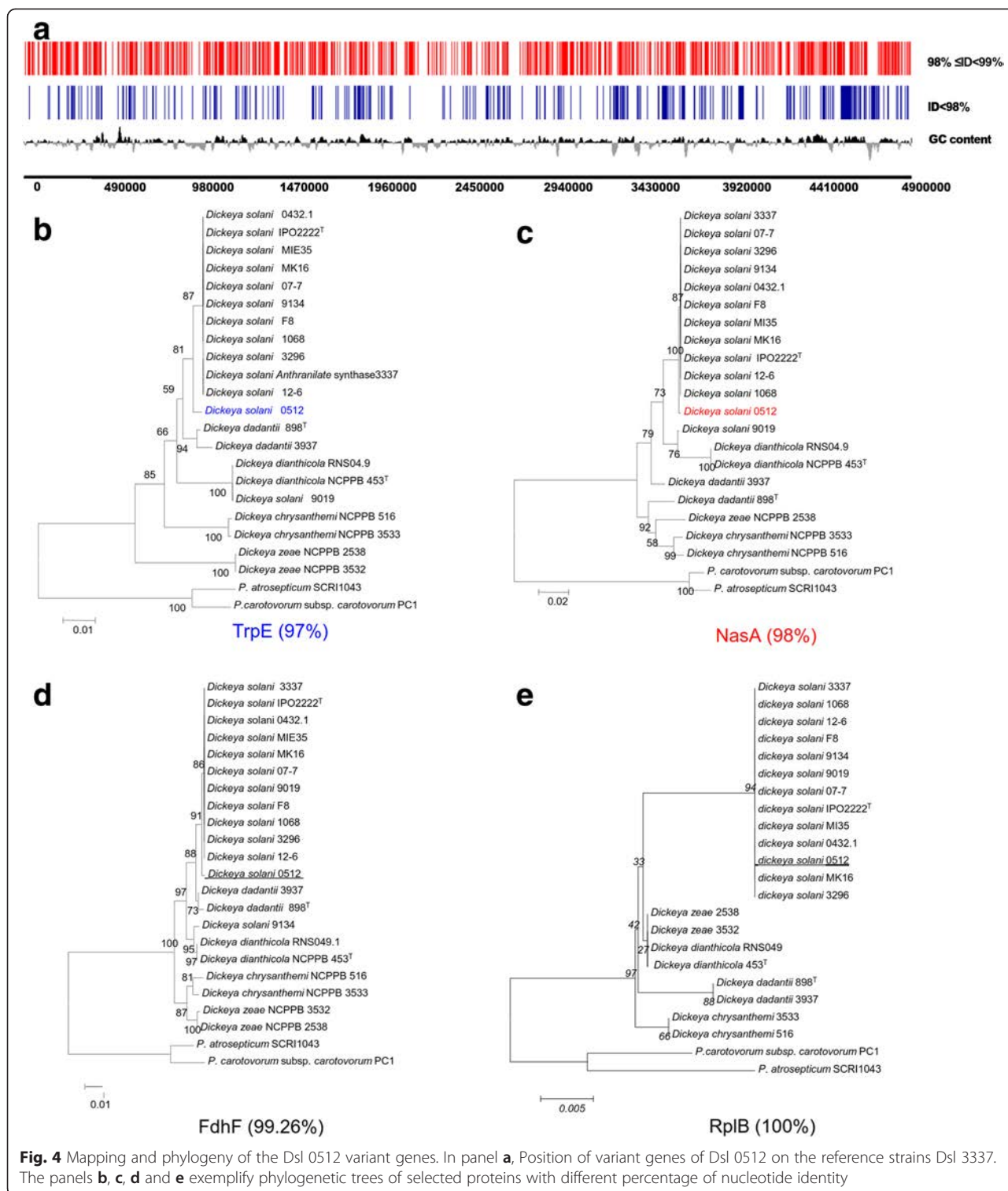
Dsl 0512 differed from the other *D. solani* by the high number and wide distribution of variations (Additional file 1: Table S3, Figs. 2 and 3), a unique phylogenetic position in MLSA (Fig. 1), and a high percentage (7.66 %) of unmapped reads against Dsl 3337 genome (Additional file 1: Table S2). Unmapped reads were used for a *de novo* assembly which generated six contigs with a size ranging from 13 248 bp to 36 630 bp. All these six contigs were absent from the other *D. solani* strains. Using MAUVE [24], these six sequences were positioned on the draft genome of Dsl 0512 that was constructed using the strain Dsl 3337 as a reference (Additional file 2: Figure S1). RAST annotation indicated that most of the genes belonging to these 6 contigs coded for phage elements and hypothetical

or unknown proteins, with the exception of some genes coding for two putative ABC transporters, two putative virulence factors and one methyl-accepting chemotaxis protein, all being carried by the contig4. The similarity scores were too weak to assign a more precise function and phylogenetic origin to these putative genes/proteins.

Another characteristic of Dsl 0512 was a high number of genes (half of the genome) that exhibited variations. These genes were distributed along the genome without any clustering in specific regions. Constructed phylogenetic trees revealed that the analyzed genes exhibiting a nucleotide identity below 98 % (compared to Dsl 3337 genes) did not belong to the core population gene cluster (Fig. 4, Additional file 3: Figure S2, Additional file 4: Figure S3, Additional file 5: Figure S4, Additional file 6: Figure S5). All these features supported the existence of a novel *D. solani* sub-group. The strain Dsl 0512 could be proposed as the eponym of the *D. solani* 0512 sub-group.

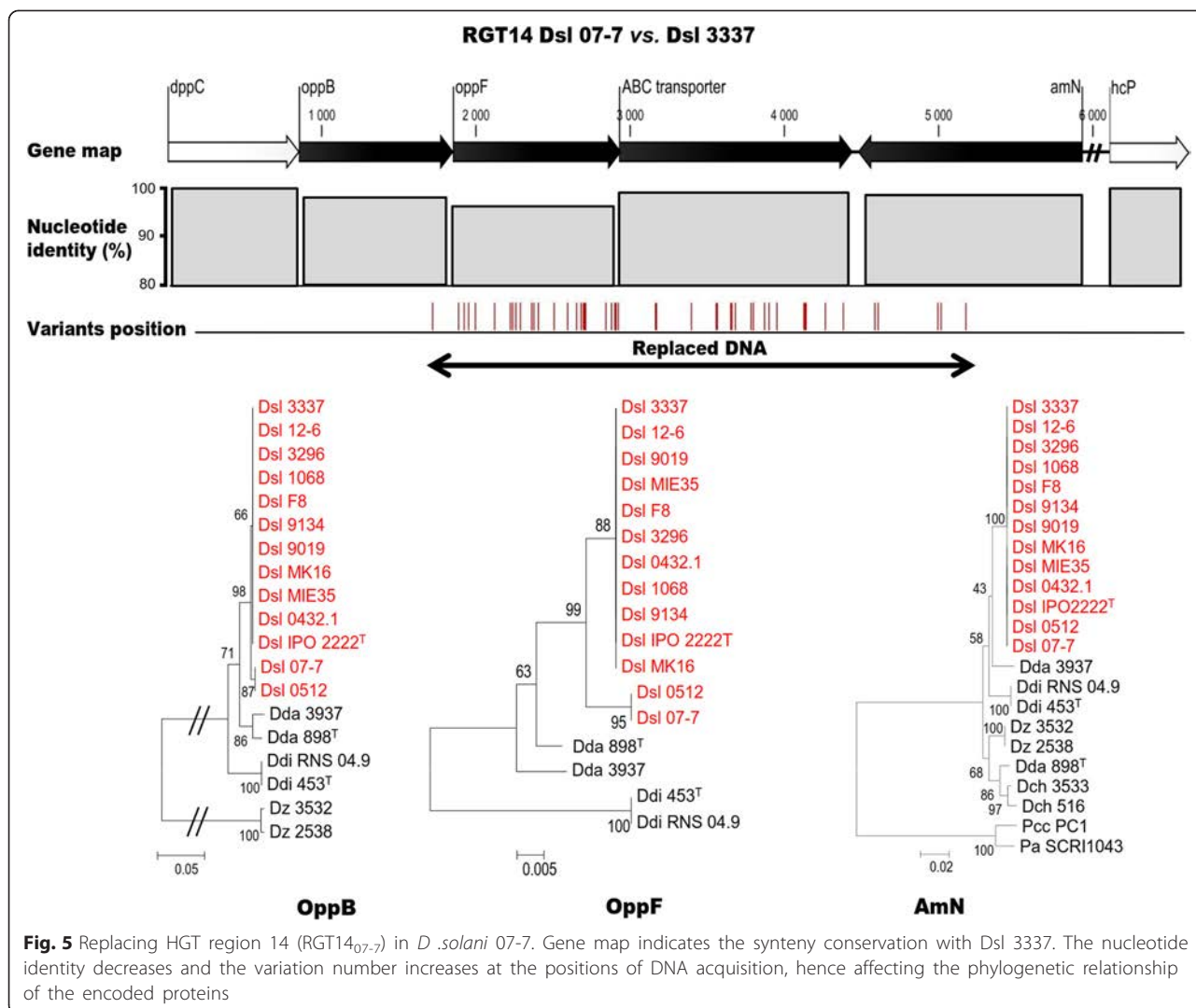
Infra-species replacing HGT in Dsl 07-7

The 144 variant genes of Dsl 07-7 showed a non-uniform distribution along the chromosome, since most of them



were clustered in 18 separate regions. One of these regions, which is presented in Fig. 5, contained four genes: *oppB*, *oppF*, an ABC transporter gene and *amiN*. These genes contained in total 47 variations leading to

a decrease of their nucleotide identity as compared to the corresponding genes in the Dsl 3337 genome. Moreover, the phylogenetic analysis of the protein sequence coded by the *oppB* and *oppF* genes, which were

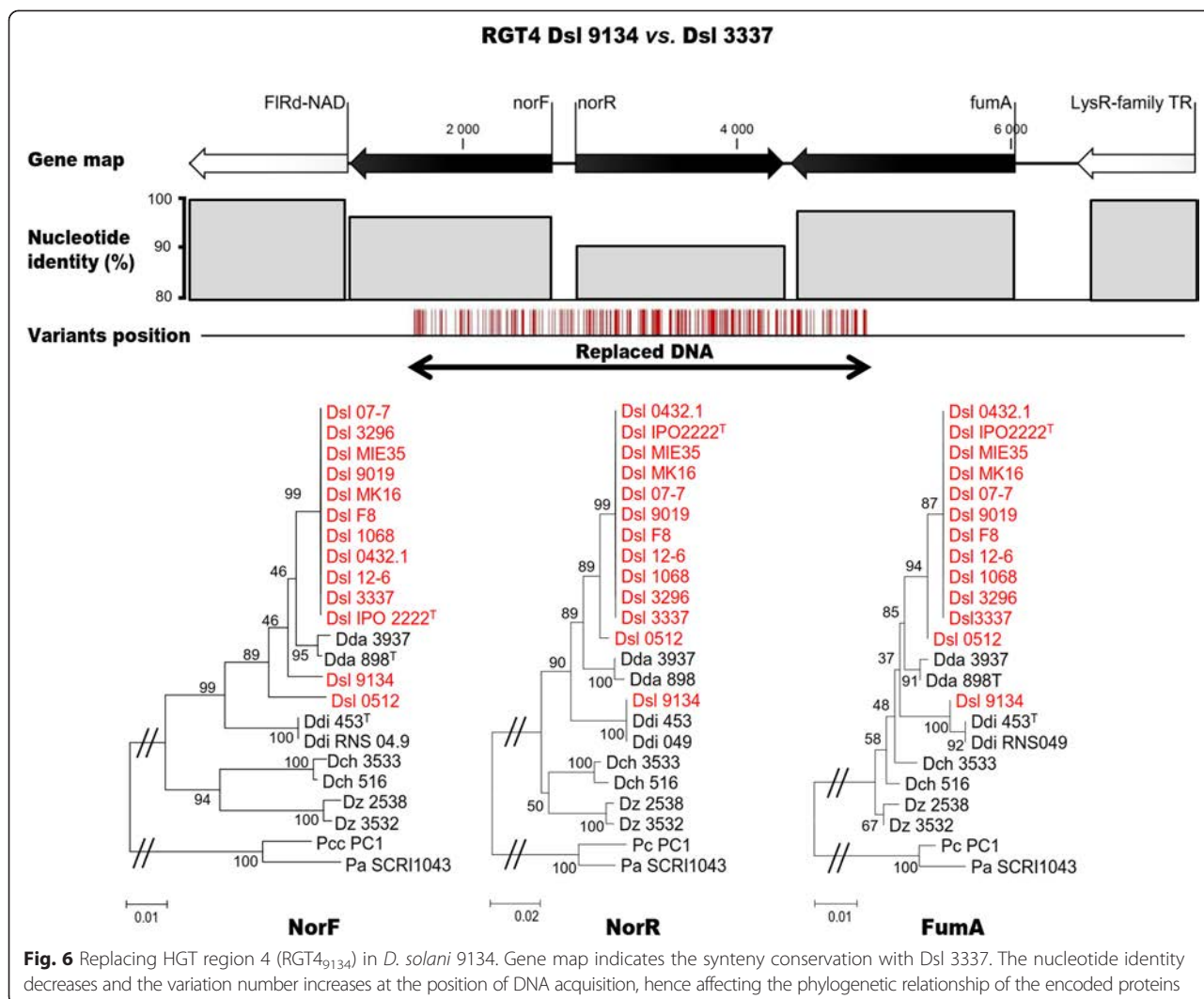


the most affected by variations, revealed a replacing HGT from a strain belonging to the *D. solani* 0512 subgroup. The 17 other regions exhibited a similar gene organization and a phylogenetic clustering with Dsl 0512 genes. Hence, all these 18 regions were called as RGT (replacing HGT) regions. They were numbered according to their successive position along the chromosome with the strain name in subscript position: RGT1₀₇₋₇, RGT2₀₇₋₇, RGT3₀₇₋₇ ... (Fig. 3). This analysis suggested that Dsl 07-7 acquired a dozen of gene fragments during massive replacing HGT(s) from strain(s) belonging to the Dsl 0512 sub-group. Hence, Dsl 07-7 exemplified the occurrence of an infra-specific gene exchange among the *D. solani* population, and also supported the possible co-existence of strains of the *D. solani* 0512 sub-group with those of the core-population.

Inter-species replacing HGT in *D. solani* strains 9134 and 9019

In Dsl 9134, 39 among the 56 genes with variations were clustered in 6 RGT regions, the other genes with variations being scattered along the chromosome. In Dsl 9019, 63 among the 73 genes with variations were clustered in 12 RGT regions. In both strains, the RGT regions were named according to the same nomenclature as in Dsl 07-7 (Fig. 3).

The RGT4₉₁₃₄ illustrated the typical organization of these RGTs in Dsl 9134 (Fig. 6). RGT4₉₁₃₄ (4860 bp) exhibited 229 positions of variations that were distributed in three genes: *norF*, *norR*, and *fumA*. These genes were related to the nitric oxide metabolism. Because of the high number of variations, the gene identity with *D. solani* strain 3337 decreased in RGT4₉₁₃₄, especially in the *norR* gene that was located in the central part of

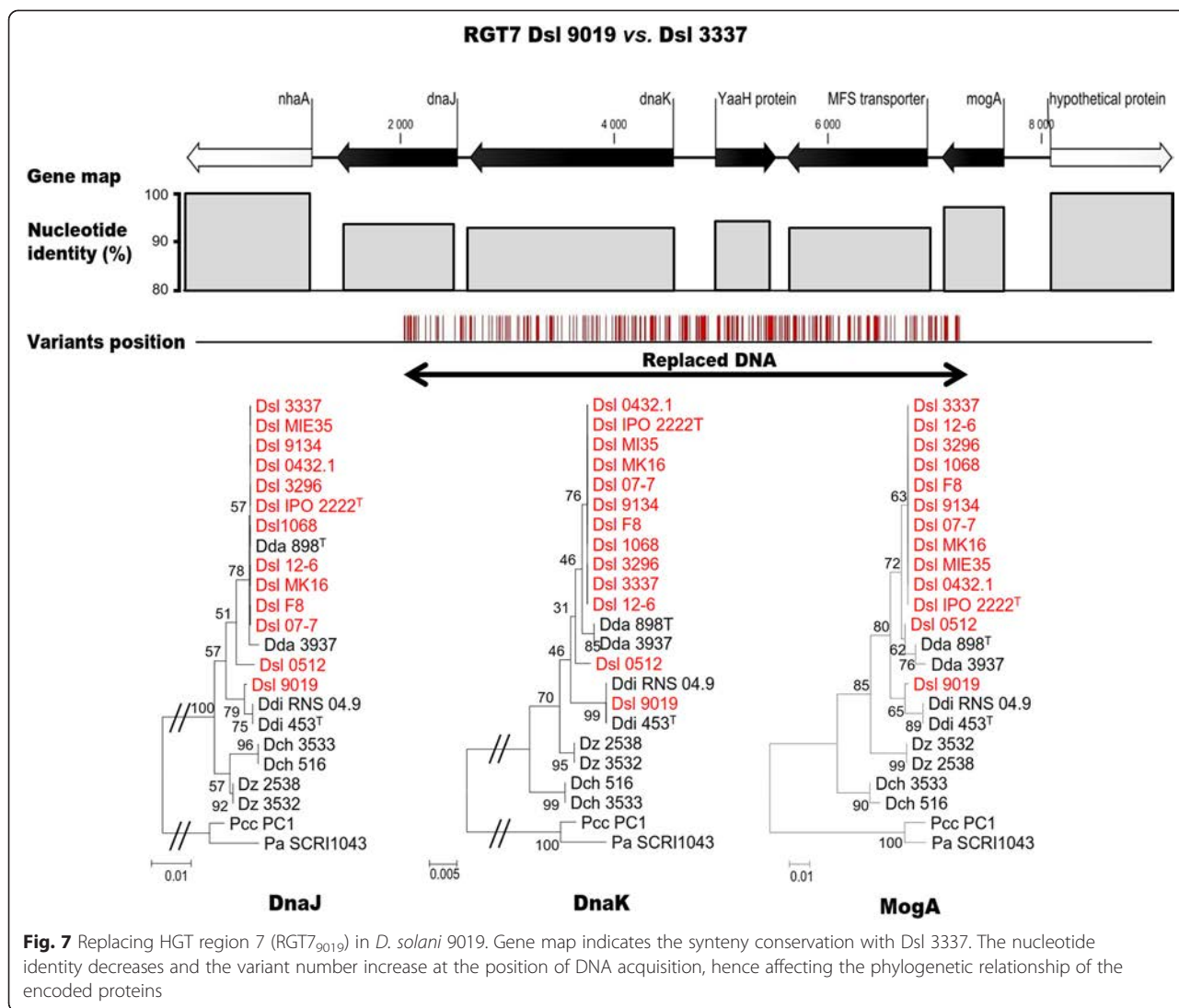


the RGT region. Protein phylogeny revealed that the three proteins encoded by the RGT4₉₁₃₄ genes did not branch with their *D. solani* counterparts but were most closely related to those of *D. dianthicola*. The variation positions suggested that replacing HGT occurred in the middle of the genes *norF* and *fumA*, and hence generated proteins with an intermediate position between the *D. solani* and *D. dianthicola* proteins in the phylogenetic trees. A second example of inter-species replacing HGT is given with the RGT7₉₀₁₉ (6248 bp) of Dsl 9019, which contained five genes *dnaJ*, *dnaK*, *yaaH*, a MFS transporter gene and *mogA* (Fig. 7). This example highlighted that replacing HGT might also affect genes such as *dnaK* and *dnaJ* which are used for MLSA and taxonomic identification [25]. In RGT7₉₀₁₉, 269 variants were detected. Discrepancies within DnaJ, DnaK and MogA phylogenies suggested the occurrence of a replacing HGT from *D. dianthicola*. In all the other RGTs of Dsl 9134 and 9019, a phylogeny approach (Additional file 5:

Figure S4, Additional file 6: Figure S5) also supported the occurrence of a replacing HGT using *D. dianthicola* population as the unique source.

Plasmid acquisition in *D. solani* strain 9019 from Burkholderia

In addition to replacing HGT, an additive HGT event that consisted in a plasmid acquisition occurred in Dsl 9019. The Dsl 9019 unmapped reads, which represented 1.9 % of the total read number (Additional file 1: Table S2), allowed the generation of a single contig (43 564 bp) by *de novo* assembly. This plasmid exhibited a complete identity (100 %) with a plasmid of *Burkholderia ambifaria* AMMD (CP000443.1). The stable replication of this plasmid in Dsl 9019 was verified in sub-cultures using plasmid-specific primers (pF1: cagcgaagagcaagac aa, pR1: tcatggaagcgatctcgg and pF2: ttaccgagccga gctgtggcgt, pR2 :caggaagtgtcgttatcgcgagt).



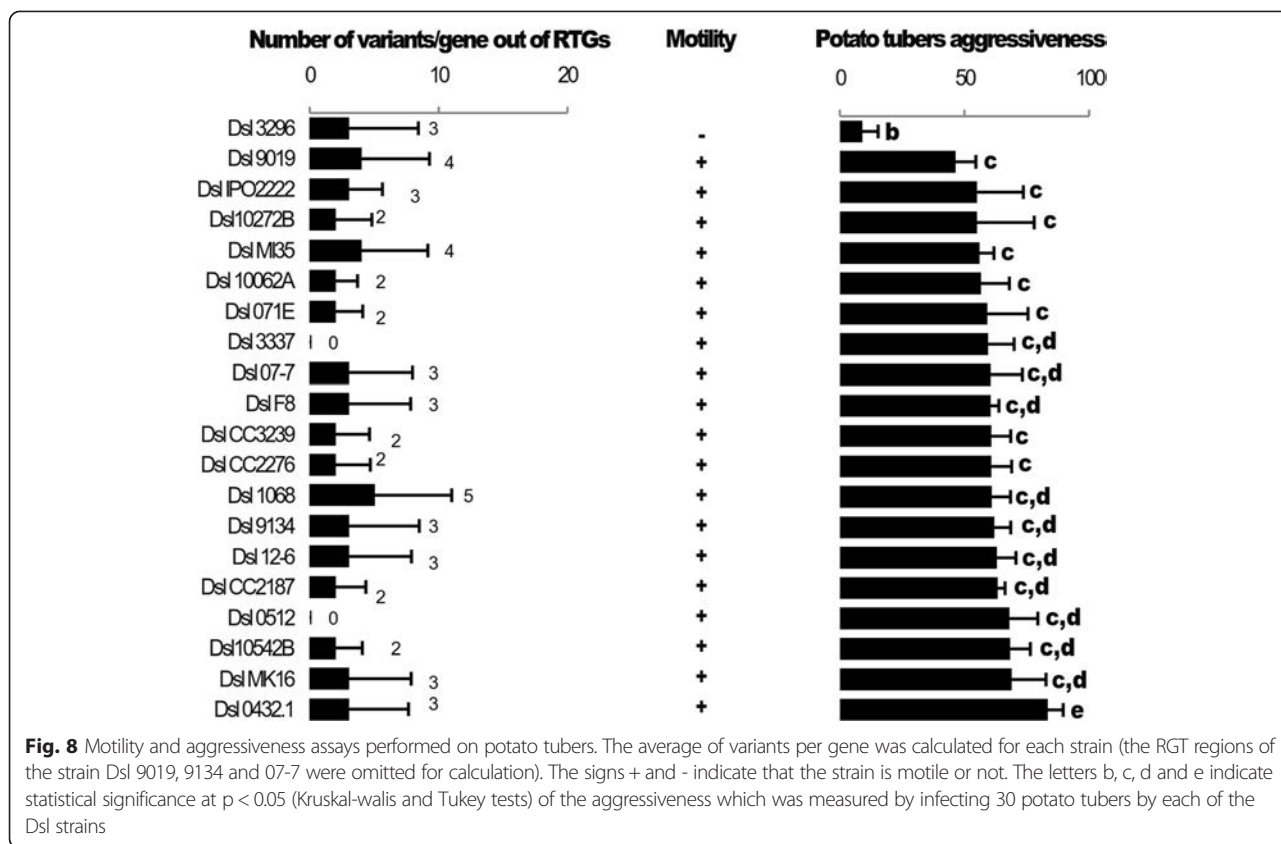
In *D. solani* 3296, variations in flagellar genes correlated motility and virulence decrease

All the non-synonymous variations of the core-population were listed in Additional file 1: Table S4. Remarkably, two unique non-synonymous variations that affected the *fliC* and *fliN* flagellar genes were present in Dsl 3296. The substitution C to T at the position 952985 lead to conversion of Ala207 to Thr in FliC, while deletion of the GTC codon starting at the position 966 038 provoke the loss of the Val112 in FliN. The nucleotide variations were verified by Sanger sequencing. These two variations were unique among the sequenced *D. solani* strains, as well as the known *Dickeya* genomes (Additional file 7: Figure S6). These genes retained our attention as *fli* genes are required for aggressiveness in *Dickeya* and in other *Enterobacteriaceae* [26–29]. We hypothesized that Dsl 3296 could be impaired in motility, hence also exhibited a reduced aggressiveness on potato host plants. We compared motility

and virulence of all the 20 Dsl analyzed in this study (Fig. 8). All strains except Dsl 3296 were motile. Moreover, a weak aggressiveness of the strain 3296 was observed in virulence assay on potato tuber, hence correlating genomic variants in *fli* genes with motility and virulence deficiency. As a consequence, even if SNP/InDel variations are scarce, some of them may affect virulence functions in Dsl strains.

Discussion

This work provided new insights into the analysis of the emerging plant-pathogen *D. solani*. We combined Illumina and PacBio technologies to determine a high quality genome sequence of *D. solani* 3337 that we used as a reference to compare 19 other genome sequences generated by Illumina technology. While previous studies reported pairwise comparison between a single *D. solani* genome with that of other *Dickeya* and *Pectobacterium* species [13, 14], this work was also based on a population genomic approach.



This approach revealed the unexpected diversity of the *D. solani* genomes that resulted from a combination of scattered SNP/InDel variations as well as replacing and additive HGT events.

The majority of analyzed *D. solani* strains (16 among 20) that we called the core-population contained only 43 to 85 variants (SNPs and InDels). This result is in accordance with the high ANI values (>99.9 %) that were calculated between each strain against the reference 3337. Other studies have pointed the high homogeneity within genetic equipment of *D. solani* population [9, 12, 14, 30]. All these molecular analyses support the clonal hypothesis of the *D. solani* population. In spite of this high homogeneity, Dsl strains may exhibit some variability in aggressiveness on potato tubers [7]. A previous pairwise comparative study of two Dsl strains did not succeed in the identification of the genes and functions that could explain the different aggressiveness trait [7]. However, using a population comparative approach, we pointed out that genetic and functional variations in the motility trait could contribute to an aggressiveness decrease. This observation exemplifies the powerfulness of genomic diversity analyses on field isolates for the identification of genes that modulate aggressiveness.

Another important result was the characterization of a sub-group within the *D. solani* species, highlighting that

D. solani population structure was more complex than described previously. The prototypic strain of this sub-group was Dsl 0512 (RNS 05.1.2A) that has been isolated from potato plant showing blackleg and soft rot symptoms in France (in 2005). The existence of the 0512 sub-group was supported by ANI value, MLSA, genomic architecture (presence of specific regions) and SNP/InDel abundance and distribution. The Dsl 0512 genome appeared as a mosaic of genes with a phylogenetic position inferred to either the *D. solani* core population or the Dsl 0512 sub-group. Remarkably, genes that belong to the Dsl 0512 phylogenetic sub-cluster have also been discovered in the 18 RGTs (143 genes) of the strain Dsl 07-7 that was also isolated in France. The involvement of the 0512 sub-group as a gene resource in replacing HGT reinforced its importance in the generation of variability in *D. solani* isolates. The strains Dsl 0512 and 07-7 showed aggressiveness level similar to that of most of the studied *D. solani* strains, suggesting that the 0512 sub-group is not associated to any particular aggressiveness behavior, at least on potato tubers.

This study also highlighted that additive and replacing HGT occurred in inter-species exchanges. Additive HGT was observed in the strain 9019 which acquired a plasmid from *B. ambifaria* AMMD. *B. ambifaria* AMMD was isolated from the rhizosphere of healthy pea plants in

Wisconsin (USA) in 1985 [31] and it has been reported as very effective in controlling phytopathogenic *Pythium* species [32]. Moreover we discovered replacing HGT events that recruited *D. dianthicola* genes in the two strains Dsl 9019 (63 genes distributed in 12 RGT regions) and 9134 (39 genes in 6 RGT regions) isolated from ornamental plants (respectively Muscari and Hyacinth). These exchange events between *D. solani* and *D. dianthicola* suggested that these two pathogens could coexist in the same ecological niche. In the case of *Pectobacterium spp.*, multiple species isolations from the same infected plants have been reported [33, 34]. Importantly, the replacing HGT events did not correlate with an aggressiveness increase in Dsl 9019 and 9134 at least in potato tubers. However, replacing HGT generates major impact on phylogenetic inference by generating incongruities that could impair pathogen molecular diagnostics which are based on housekeeping genes. Importantly, it has been reported that succeeded HGT events between distantly related bacteria mostly implicate housekeeping genes that are also the most conserved between different species [35, 36]. Our work revealed that in the Dsl 9019 strain, the *dnaJ* and *dnaK* genes, which are usually used in phylogenetic classifications [25, 37], have been recruited from *D. dianthicola*. Since the *Dickeya* pathogens are genetically very close (ANI \geq 93 %), replacing HGT could be predicted to interfere recurrently with taxonomical diagnostics, hence provoking assignation errors. The impact of HGT on taxonomy has been discussed in different *Enterobacteriaceae* [38, 39]. An immediate applied recommendation from our work is that even though *D. solani* is mainly described as a homogeneous population, the existence of HGT events should encourage the use of multiple taxonomical markers.

Conclusions

As a conclusion, this work revealed that *D. solani* genomic variability may be caused by SNPs/InDels as well as replacing and additive HGT events, including plasmid acquisition. From this work, the question arises about the dynamics of the *D. solani* diversity in the course of its emergence and spreading in crop cultures. This might be further investigated by a larger scale sampling and genomic analysis.

Methods

Bacterial strains and growth conditions

D. solani strains were collected from different geographical locations and dates of isolation and also from different hosts or environments (Additional file 1: Table S1). All the strains were routinely cultured in TY medium (tryptone 5 g/L, yeast extract 3 g/L and agar 1.5 %) at 28 °C.

DNA extraction and sequencing

Genomic DNA from each strain was extracted from overnight culture using a phenol-chloroform purification method followed by an ethanol precipitation as described by [40] Wilson. Quantity and quality control of the DNA was completed using a NanoDrop (ND 1000) device and agarose gel electrophoresis at 1.0 %.

Paired-end libraries with an insert size of 270 to 390 bp were constructed for each strain, and DNA sequencing was performed by Illumina HiSeq 2000 v3 technology. Sequencing of the library was carried out using 2×100 or 2×150 bp paired-end read module. Illumina sequencing was performed at the CNRS IMAGIF platform (Gif-sur-Yvette).

The genomic DNA of Dsl 3337 was subjected to PacBio RSII sequencing technology (Pacific Biosciences, CA, USA) using library targeted at 10kbp in insert size. Prior to assembly, short reads that are less than 500 bp were filtered off and minimum polymerase read quality used for mapping of subreads from a single zero-mode waveguides (ZMW) was set at 0.75. The 112 228 filtered reads (N50 value was 13 159 bp and total bp number was 814 445 948) were assembled using RS_HGAP_Assembly (version 3.0), which is an analysis pipeline module from Pacific Biosciences SMRT portal incorporating Celera Assembler, BLASR mapper and Quiver consensus caller algorithm. The cut-off length of seeding reads was set at 3 606 bp in order to serve as a reference for the recruitment of shorter reads for preassembly. The resulted consensus accuracy based on multiple sequence alignment of the subreads was at 99.99 %.

Assembly, variants calling and genome sequence analysis

Assembly of the sequences was performed using the CLC Genomics Workbench v7.0.0 software (CLC Inc, Aarhus, Denmark). After quality (quality score threshold 0.05) and length (above 40 nucleotides) trimming of the sequences, contigs were generated by *de novo* assembly (CLC parameters: automatic determination of the word and bubble sizes with no scaffolding) for each strain.

Paired end reads for each strain were mapped against the reference sequence of the strain *D. solani* 3337 at mild stringency threshold (0.8 of identity on 0.5 of read length) using CLC Genomics Workbench version 7.0.0 software. The unmapped reads for each strain were collected. The mappings were used for detection of variations (SNPs and InDels) using basic variant calling tool from CLC genomic workbench version 7.0.0. Draft genome sequences composed of the contigs of each strain were used to search and analyze the variations detected. Variations with an occurrence below 99 % in the mapping step were discarded from the study.

The nucleotide identity (ANI) values were calculated as previously proposed [41] using the ANI calculator

from the Kostas lab with default settings (<http://enve-omics.ce.gatech.edu/ani/>). Phylogenetic and molecular evolutionary analyses were conducted using MEGA, version 6 [23]. An MLSA (Multi-locus sequence analysis) was performed using eleven housekeeping genes (*rpoD*, *gyrB*, *recA*, *rpoS*, *dnaX*, *dnaA*, *gapA*, *fusA*, *rplB*, *purA*, *gyrA*) retrieved from the twenty *D. solani* strains in order to confirm their phylogenetic position within known pectinolytic *Dickeya* and *Pectobacterium* strains.

Nucleotide sequence accession number

Draft genome sequences of *Dickeya solani* strains 9109, 0512, 9134, 07-7 have been deposited at DDBJ/EMBL/GenBank under the following accession numbers: (JWLS00000000) *D. solani* 9019, (JWMJ00000000) *D. solani* 0512, (JWLT00000000) *D. solani* 9134, (JWLR00000000) *D. solani* 07-7. The versions described in this paper are versions (JWLS01000000) *D. solani* 9019, (JWMJ01000000) *D. solani* 0512, (JWLT01000000) *D. solani* 9134, (JWLR01000000) *D. solani* 07-7. Genomes of other *Dickeya* and *Pectobacterium* species were collected from public database (Table S4).

Aggressiveness and motility assays

Motility assays were conducted on semi-solid SM medium (beef extract at 3 g/L, peptone at 5 g/L, and 25 ml/L of 20 % glucose) with 0.5 % of agar. Two μ L of an overnight bacterial suspension of each strain were used to inoculate agar plates which were incubated 16 h at 28 °C. The experiment was performed twice with 2 replicates each time.

Assessment of the aggressiveness of the strains was performed on potato tubers (cv. Binjete). To this end, 10^6 CFU were used to infect 10 potato tubers for each strain. After 24 h of incubation at 25 °C, five aggressiveness categories were considered and attributed to tuber samples to assess the virulence of the strains. The experiments were performed three times, hence 600 tubers were infected and analyzed. The results were represented as normalized values.

Virulence assays were statistically analyzed to infer the aggressiveness variability within strains on potato tubers. Heterogeneity of strains was assessed using a Kruskal-Wallis test with $p < 0.05$. Statistical significance of the pairwise comparisons between strains was calculated using a post hoc Tukey test with $p < 0.05$.

Availability of supporting data

The alignments and phylogenetical tree for MLSA are available through the Dryad data repository doi: 10.5061/dryad.h26hs.

Additional files

Additional file 1: Tables S1-S5. Table S1. *Dickeya solani* strains in this study. **Table S2.** Sequencing data and mappings on the Dsl 3337 genome. **Table S3** Variants distribution on the strains vs. Dsl 3337. **Table S4.** Non-synonymous variants: this table shows the unique and the shared variants within homogeneous *D. solani* strains (MK16, MIE35, 0432.1, 12-6, F8, 1068, 3296, 07E, 10062A, 10272B, 10542B, 2187, 2276, 3239, IPO2222^T). **Table S5.** Other genomes used in this study (DOC 130 kb)

Additional file 2: Figure S1. Synteny between the strain *D. solani* 3337 and the draft genome *D. solani* 0512. The alignment was performed using MAUVE software, underlining a high conservation of the synteny. The numbers indicate the positions of the strain-specific genomic regions generated by *de novo* assembly of the unmapped reads. (TIFF 956 kb)

Additional file 3: Figure S2. Protein-based phylogenetic trees revealing Dsl 0512 as a member of a distinct sub-cluster within the *D. solani* species. (TIFF 2843 kb)

Additional file 4: Figure S3. Protein-based phylogenetic trees of different RGTs in Dsl 07-7. The genes were retrieved from RGT1, RGT2, RGT3, RGT6, RGT10 and RGT12 of Dsl 07-7. The phylogenetic positions indicate replacing HGT events from the *D. solani* 0512 sub-group. (TIFF 2881 kb)

Additional file 5: Figure S4. Protein-based phylogenetic trees of different RGTs in Dsl 9134. The genes were retrieved from RGT1, RGT3, RGT5 and RGT6 of Dsl 9134. The phylogenetic positions highlight replacing HGT events from the *D. dadantii* species. (TIFF 2117 kb)

Additional file 6: Figure S5. Protein-based phylogenetic trees of different RGTs in Dsl 9019. The genes were retrieved from RGT1, RGT2, RGT3, RGT4, RGT5 and RGT10 of Dsl9019. The phylogenetic positions highlight replacing HGT events from the *D. dadantii* species. (TIFF 2989 kb)

Additional file 7: Figure S6. Local alignment of FlIC and FlIN proteins. The variations at the positions 207 in FlIC and 112 in FlIN are indicated in red color, other variations are in blue color. Amino acid position is numbered according to the *D. solani* 3337 sequence of FlIC and FlIN. We used the draft and complete genomes of the 19 *D. solani* sequenced in this study, those of *D. solani* strains GBCC2040 and MK10, 15 *D. dianthicola* including the strains MIE32, MIE33, MIE34, CFBP1888, CFBP2015, CFBP2982, RNS04.9, RNS10.20.2A, RNS11.47.1A, DW04.9 K, DS05.3.3, GBBC2039, IPO980, NBPPB3534 and NCPPB453, *D. dadantii* strains 3937, NCPPB898 and NCPPB3537, *D. chrysanthemi* strains NCPPB3533 and NCPPB516, and *D. zeae* strains Ech1591 and NCPPB2538. (TIFF 1283 kb)

Abbreviations

ANI: Average nucleotide identity; CDS: Coding DNA sequence; Dsl: *Dickeya solani*; HGT: Horizontal gene transfer; InDel: Insertion deletion; MLSA: Multi-locus sequence analysis; NCBI: National center for biotechnology information; RGT: Replacing HGT; SNP: Single nucleotide polymorphism.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SK performed comparative genomics and phylogenetic analyses, PB performed virulence assays and DNA extractions, JP and FVG analyzed genomic data, TMC and KGC carried out PacBio sequencing, VH and FVG provided strains, DF coordinated the project, and all the authors, SK, PB, JP, FVG, TMC, KGC, VH, DF and MM contribute in manuscript writing. All the authors read and approved the final manuscript.

Acknowledgements

We thank Robert Dees (Wageningen UR/Applied Plant Research) for the gift of the ornamental strains *D. solani* PPO9019 and PPO9134, Minna Pirhonen (Department of Applied Biology, University of Helsinki) for providing PPL0433 (=F8), Leah Tsror (Gilat Research Center, Agricultural Research Organisation) for providing GRC77 (EU3296) and Yves Dessaux (I2BC, CNRS) for his help in the manuscript editing. This work was supported by a cooperative project between France and Morocco (PRAD 14-02, Campus France n° 30229 ZK), the University Paris-Saclay (Co-tutelle funding), the excellence grant (n°H011/007) awarded by the Ministry of Higher education of Morocco, a collaborative project between

CNRS (Gif sur Yvette) and FN3PT-RD3PT (Paris), the High Impact Research Grant (U.M.C/625/1/HIR/MOHE/CHAN/14/01, Grant number A-000001-50001 to KGC) and the French-Malaysian exchange program awarded by French Embassy of Malaysia.

Author details

¹Institute for Integrative Biology of the Cell (I2BC), CNRS CEA Univ. Paris-Sud, Université Paris-Saclay, Saclay Plant Sciences, Avenue de la Terrasse, 91198 Gif-sur-Yvette cedex, France. ²Université Moulay Ismaïl, Faculté des Sciences, Département de Biologie, Meknès, Morocco. ³UPMC Univ Paris 06, UMR 7618, IEEES Paris (Institute of Ecology and Environmental Sciences), 7 Quai Saint Bernard, 75005 Paris, France. ⁴Division of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia. ⁵Fédération Nationale des Producteurs de Plants de Pomme de Terre-Recherche développement Promotion du Plant de Pomme de Terre (FN3PT-RD3PT), 75008 Paris, France. ⁶UMR 1349 IGEPP INRA - Agrocampus Ouest Rennes, 35653 LeRheu, France. ⁷INRA, UMR 1392, IEEES Paris (Institute of Ecology and Environmental Sciences), 7 Quai Saint Bernard, 75005 Paris, France.

Received: 19 June 2015 Accepted: 3 October 2015

Published online: 14 October 2015

References

- Samson R. Transfer of *Pectobacterium chrysanthemi* (Burkholder et al. 1953) Brenner et al. 1973 and *Brenneria paradisiaca* to the genus *Dickeya* gen. nov. as *Dickeya chrysanthemi* comb. nov. and *Dickeya paradisiaca* comb. nov. and delineation of four novel species, *Dickeya dadantii* sp. nov., *Dickeya dianthicola* sp. nov., *Dickeya dieffenbachiae* sp. nov. and *Dickeya zeae* sp. nov. *Int J Syst Evol Microbiol.* 2005;55:1415–27.
- Gardan L. Elevation of three subspecies of *Pectobacterium carotovorum* to species level: *Pectobacterium atrosepticum* sp. nov., *Pectobacterium betavasculorum* sp. nov. and *Pectobacterium wasabiae* sp. nov. *Int J Syst Evol Microbiol.* 2003;53:381–91.
- Collmer A, Keen NT. The role of pectic enzymes in plant pathogenesis. *Annu Rev Phytopathol.* 1986;24:383–409.
- Toth IK, van der Wolf JM, Saddler G, Lojkowska E, Hélias V, Pirhonen M, et al. *Dickeya* species: an emerging problem for potato production in Europe: *Dickeya* spp. on potato in Europe. *Plant Pathol.* 2011;60:385–99.
- Van der Wolf JM, Nijhuis EH, Kowalewska MJ, Saddler GS, Parkinson N, Elphinstone JG, et al. *Dickeya solani* sp. nov., a pectinolytic plant-pathogenic bacterium isolated from potato (*Solanum tuberosum*). *Int J Syst Evol Microbiol.* 2014;64:768–74.
- Czajkowski R, de Boer WJ, van der Zouwen PS, Kastelein P, Jafra S, de Haan EG, et al. Virulence of “*Dickeya solani*” and *Dickeya dianthicola* biovar-1 and -7 strains on potato (*Solanum tuberosum*): Comparison of *Dickeya* spp. on potato. *Plant Pathol.* 2013;62:597–610.
- Potrykus M, Golanowska M, Hugouvieux-Cotte-Pattat N, Lojkowska E. Regulators Involved in *Dickeya solani* Virulence, Genetic Conservation, and Functional Variability. *Mol Plant Microbe Interact.* 2014;27:700–11.
- Czajkowski R, de Boer WJ, Velvis H, van der Wolf JM. Systemic colonization of potato plants by a soilborne, green fluorescent protein-tagged strain of *Dickeya* sp. biovar 3. *Phytopathology.* 2010;100:134–42.
- Laurila J, Ahola V, Lehtinen A, Joutsjoki T, Hannukkala A, Rahkonen A, et al. Characterization of *Dickeya* strains isolated from potato and river water samples in Finland. *Eur J Plant Pathol.* 2008;122:213–25.
- Ślawiak M, van Beckhoven JRCM, Speksnijder AGCL, Czajkowski R, Grabe G, van der Wolf JM. Biochemical and genetical analysis reveal a new clade of biovar 3 *Dickeya* spp. strains isolated from potato in Europe. *Eur J Plant Pathol.* 2009;125:245–61.
- Parkinson N, Stead D, Bew J, Heeney J, Tsrer Lahkim L, Elphinstone J. *Dickeya* species relatedness and clade structure determined by comparison of recA sequences. *Int J Syst Evol Microbiol.* 2009;59(Pt 10):2388–93.
- Van Vaerenbergh J, Baeyen S, De Vos P, Maes M. Sequence Diversity in the *Dickeya* *flhC* Gene: Phylogeny of the *Dickeya* Genus and TagMan® PCR for “*D. solani*”, New Biovar 3 Variant on Potato in Europe. *PLoS ONE.* 2012;7:e35738.
- Garlant L, Koskinen P, Rouhiainen L, Laine P, Paulin L, Auvinen P, et al. Genome sequence of *Dickeya solani*, a new soft rot pathogen of potato, suggests its emergence may be related to a novel combination of non-ribosomal peptide/polyketide synthetase clusters. *Diversity.* 2013;5:824–42.
- Pédron J, Mondy S, Raoul des Essarts Y, Van Gijsegem F, Faure D. Genomic and metabolic comparison with *Dickeya dadantii* 3937 reveals the emerging *Dickeya solani* potato pathogen to display distinctive metabolic activities and T5SS/T6SS-related toxin repertoire. *BMC Genomics.* 2014;15:283.
- Smith MW, Feng D-F, Doolittle RF. Evolution by acquisition: the case for horizontal gene transfers. *Trends Biochem Sci.* 1992;17:489–93.
- Smith JM, Smith NH, O'Rourke M, Spratt BG. How clonal are bacteria? *Proc Natl Acad Sci U S A.* 1993;90:4384–8.
- Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature.* 2000;405:299–304.
- Koonin EV, Makarova KS, Aravind L. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol.* 2001;55:709–42.
- Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol.* 2005;3:722–32.
- Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev.* 2009;33:376–93.
- Choi SC, Rasmussen MD, Hubisz MJ, Gronau I, Stanhope MJ, Siepel A. Replacing and additive horizontal gene transfer in *Streptococcus*. *Mol Biol Evol.* 2012;29:3309–20.
- Khayi S, Mondy S, Beury-Cirou A, Mounmi M, Hélias V, Faure D. Genome sequence of the emerging plant pathogen *Dickeya solani* strain RNS 08.23.3.1A. *Genome Announc.* 2014;2:e01270–13.
- Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol.* 2013;30:2725–9.
- Darling ACE. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 2004;14:1394–403.
- Colston SM, Fullmer MS, Beka L, Lamy B, Gogarten JP, Graf J. Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. *mBio.* 2014;5:e02136–14.
- Rogers TJ, Paton JC, Wang H, Talbot UM, Paton AW. Reduced virulence of an *flhC* mutant of Shiga-toxicogenic *Escherichia coli* O113:H21. *Infect Immun.* 2006;74:1962–6.
- Jahn CE, Willis DK, Charkowski AO. The flagellar sigma factor FlhA is required for *Dickeya dadantii* virulence. *Mol Plant Microbe Interact.* 2008;21:1431–42.
- He Y, Xu T, Fosshem LE, Zhang X-H. FlhC, a flagellin protein, is essential for the growth and virulence of fish pathogen *Edwardsiella tarda*. *PLoS ONE.* 2012;7:e45070.
- Dong T, Schellhorn HE. Role of RpoS in virulence of pathogens. *Infect Immun.* 2010;78:887–97.
- Pritchard L, Humphris S, Saddler GS, Parkinson NM, Bertrand V, Elphinstone JG, et al. Detection of phytopathogens of the genus *Dickeya* using a PCR primer prediction pipeline for draft bacterial genome sequences: *Dickeya* diagnostics from draft bacterial genome sequences. *Plant Pathol.* 2013;62:587–96.
- Winsor GL, Khaira B, Van Rossum T, Lo R, Whiteside MD, Brinkman FSL. The *Burkholderia* Genome Database: facilitating flexible queries and comparative analyses. *Bioinforma Oxf Engl.* 2008;24:2803–4.
- Coenye T, Mahenthalingam E, Henry D, LiPuma JJ, Laevens S, Gillis M, et al. *Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including biocontrol and cystic fibrosis-related isolates. *Int J Syst Evol Microbiol.* 2001;51:1481–90.
- Ma B, Hibbing ME, Kim H-S, Reedy RM, Yedidia I, Breuer J, et al. Host range and molecular phylogenies of the soft rot enterobacterial genera *Pectobacterium* and *Dickeya*. *Phytopathology.* 2007;97:1150–63.
- Glasner JD, Marquez-Villavicencio M, Kim H-S, Jahn CE, Ma B, Biehl BS, et al. Niche-specificity and the variable fraction of the *Pectobacterium* pan-genome. *Mol Plant-Microbe Interact.* 2008;21:1549–60.
- Jain R, Rivera MC, Lake JA. Horizontal gene transfer among genomes: The complexity hypothesis. *Proc Natl Acad Sci U S A.* 1999;96:3801–6.
- Rivera MC, Jain R, Moore JE, Lake JA. Genomic evidence for two functionally distinct gene classes. *Proc Natl Acad Sci U S A.* 1998;95:6239–44.
- Alexandre A, Laranjo M, Young JPW, Oliveira S. dnaJ is a useful phylogenetic marker for alphaproteobacteria. *Int J Syst Evol Microbiol.* 2008;58:2839–49.
- Philippe H, Douady CJ. Horizontal gene transfer and phylogenetics. *Curr Opin Microbiol.* 2003;6:498–505.

39. Shapiro BJ. Signatures of natural selection and ecological differentiation in microbial genomes. In: Landry CR, Aubin-Horth N, editors. *Ecological genomics*. Volume 781. Dordrecht: Springer Netherlands; 2014. p. 339–59.
40. Wilson K. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol*. 1987;Chapter 2:2–4.
41. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol*. 2007;57:81–91.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Draft Genome Sequences of the Three *Pectobacterium*-Antagonistic Bacteria *Pseudomonas brassicacearum* PP1-210F and PA1G7 and *Bacillus simplex* BA2H3

Slimane Khayi,^{a,b} Yannick Raoul des Essarts,^{a,c} Samuel Mondy,^a Mohieddine Moumni,^b Valérie Hélias,^{c,e} Amélie Beury-Cirou,^d Denis Faure^a

CNRS, Institut des Sciences du Végétal, UPR2355, Saclay Plant Sciences, Gif-sur-Yvette, France^a; Faculté des Sciences, Département de Biologie, Université Moulay Ismail, Meknès, Morocco^b; Fédération Nationale des Producteurs de Plants de Pomme de Terre-Recherche Développement Promotion du Plant de Pomme de Terre (FN3PT-RD3PT), Paris, France^c; Comité Nord Plant de Pomme de Terre (CNPPT), Semences, Innovation, Protection Recherche et Environnement (SIPRE), Achicourt, France^d; UMT Innoplant (FN3PT-INRAIGEPP1349), Le Rheu, France^e

***Pectobacterium* spp. are bacterial pathogens causing soft rot diseases on a wide range of plants and crops. We present in this paper the draft genome sequences of three bacterial strains, *Pseudomonas brassicacearum* PP1-210F and PA1G7 and *Bacillus simplex* BA2H3, which exhibit antagonistic activities against the *Pectobacterium* plant pathogens.**

Received 11 December 2014 Accepted 18 December 2014 Published 29 January 2015

Citation Khayi S, Raoul des Essarts Y, Mondy S, Moumni M, Hélias V, Beury-Cirou A, Faure D. 2015. Draft genome sequences of the three *Pectobacterium*-antagonistic bacteria *Pseudomonas brassicacearum* PP1-210F and PA1G7 and *Bacillus simplex* BA2H3. *Genome Announc.* 3(1):e01497-14. doi:10.1128/genomeA.01497-14.

Copyright © 2015 Khayi et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Denis Faure, faure@isv.cnrs-gif.fr.

Pectobacterium atrosepticum, *Pectobacterium wasabiae*, and *Pectobacterium carotovorum*, including *P. carotovorum* subsp. *brasilienis* and *carotovorum*, are worldwide pathogens responsible for blackleg and soft rot diseases on potato plants and tubers (1–3). Two biocontrol strategies against *Pectobacterium* phytopathogens have been developed, those of antibiosis and antivirulence (4). The biocontrol strain *Rhodococcus erythropolis* R138 targets the expression of the virulence functions in *Pectobacterium* spp. because of its capacity to disrupt the quorum-sensing signals *N*-acylhomoserine lactones (5). The genome sequence of the *R. erythropolis* antivirulence agent R138 was published recently (6). This antivirulence agent does not inhibit the growth of *Pectobacterium* spp. In contrast, we isolated three *Pectobacterium*-antagonistic bacteria, *Pseudomonas brassicacearum* PP1-210F and PA1G7 and *Bacillus simplex* BA2H3, which exhibit an ability to inhibit the growth of *Pectobacterium* strains *in vitro*. An assessment of their antagonistic abilities in the greenhouse and field settings is under way. Some other strains belonging to the *Pseudomonas fluorescens*, *P. brassicacearum*, and *B. simplex* species were previously described for their biocontrol activities against different microbial pathogens (7–10).

The genomic DNA of each bacterium was subjected to the next-generation Illumina HiSeq 2000 version 3 technology. A shotgun long jumping distance mate-pair library was constructed,

with an insert size of 8,000 bp. The sequencing of the library was carried out using a 2 by 100-bp paired-end read module by Eurofins Genomics (France). Assembly was performed by CLC Genomics 5.5 (CLC bio). The sequence reads were trimmed on quality (threshold 0.05), and minimal size (>60 nucleotides). Contigs were generated by *de novo* assembly (CLC parameters, automatic determination of the word and bubble sizes with no scaffolding). Scaffolding of the contigs was performed using SSPACE basic version 2.0 (11). For the finishing, automatic gap closure was processed using GapFiller version 1.11 (12). The remaining gaps were resolved by the mapping of mate pairs, using as a reference the 8 kb from each of the contig ends (read length, 0.9; identity, 0.95). Next, using homemade script and fastq select.tcl from the MIRA3 package, the mapped reads for both orientations (R1 and R2) were retrieved and *de novo* assembled (using the CLC parameters). The sequences were annotated using the Rapid Annotations using Subsystems Technology (RAST) pipeline (13). The detailed statistics for the three draft genome sequences are summarized in Table 1.

Nucleotide sequence accession numbers. The whole-genome shotgun projects for these bacteria have been deposited at DDBJ/EMBL/GenBank under the accession numbers [AYJR000000000](https://www.ncbi.nlm.nih.gov/nuccore/AYJR000000000) (*P. brassicacearum* PP1-210F), [AXBR000000000](https://www.ncbi.nlm.nih.gov/nuccore/AXBR000000000) (*B. simplex* BA2H3), and [JBON000000000](https://www.ncbi.nlm.nih.gov/nuccore/JBON000000000) (*P. brassicacearum* PA1G7). The versions de-

TABLE 1 Statistics for the 3 draft genome sequences

Organism	Accession no.	Genome size (bp)	N_{50} (bp)	No. of contigs	No. of scaffolds	G+C content (%)	No. of CDSs ^a	No. of tRNAs	No. of rRNAs
<i>B. simplex</i> BA2H3	AXBR000000000	5,542,531	339,104	34	11	40.2	5,856	75	31
<i>P. brassicacearum</i> PP1-210F	AYJR000000000	6,772,045	210,148	51	5	60.4	6,045	67	15
<i>P. brassicacearum</i> PA1G7	JBON000000000	6,789,417	301,959	53	8	60.5	6,052	57	13

^a CDSs, coding DNA sequences.

scribed in this paper are versions AYJR01000000 (*P. brassicacearum* PP1-210F), AXBR01000000 (*B. simplex* BA2H3), and JBON01000000 (*P. brassicacearum* PA1G7).

ACKNOWLEDGMENTS

S.K. received a Ph.D. grant from Paris-Sud University (Paris-Saclay University) and the Ministry of Higher Education of Morocco (no. H011/007); Y.R.D.E. received a Ph.D. grant from FN3PT-RD3PT and the Association Nationale de la Recherche et de la Technologie (ANRT-CIFRE no. 1282/2011).

This work was supported by cooperative projects between France and Morocco (PRAD 14-02, Campus France no. 30229 ZK), and between CNRS, FN3PT-RD3PT, and CNPPT-SIPRE. This project received a French State grant from LABEX Saclay Plant Sciences (reference ANR-10-LABX-0040-SPS) managed by the French National Research Agency under the Investments for the Future program (reference no. ANR-11-IDEX-0003-02).

REFERENCES

1. Pérombelon MCM. 2002. Potato diseases caused by soft rot erwinias: an overview of pathogenesis. *Plant Pathol* 51:1–12.
2. Duarte V, De Boer SH, Ward LJ, de Oliveira AMR. 2004. Characterization of atypical *Erwinia carotovora* strains causing blackleg of potato in Brazil. *J Appl Microbiol* 96:535–545. <http://dx.doi.org/10.1111/j.1365-2672.2004.02173.x>.
3. Pitman AR, Harrow SA, Visnovsky SB. 2010. Genetic characterisation of *Pectobacterium wasabiae* causing soft rot disease of potato in New Zealand. *Eur J Plant Pathol* 126:423–435. <http://dx.doi.org/10.1007/s10658-009-9551-y>.
4. Faure D, Dessaux Y. 2007. Quorum sensing as a target for developing control strategies for the plant pathogen *Pectobacterium*. *Eur J Plant Pathol* 119:353–365. <http://dx.doi.org/10.1007/s10658-007-9149-1>.
5. Cirou A, Mondy S, An S, Charrier A, Sarrazin A, Thoison O, DuBow M, Faure D. 2012. Efficient biostimulation of native and introduced quorum-quenching *Rhodococcus erythropolis* populations is revealed by a combination of analytical chemistry, microbiology, and pyrosequencing. *Appl Environ Microbiol* 78:481–492. <http://dx.doi.org/10.1128/AEM.06159-11>.
6. Kwasiborski A, Mondy S, Beury-Cirou A, Faure D. 2014. Genome sequence of the quorum-quenching *Rhodococcus erythropolis* strain R138. *Genome Announc* 2(2):e00224-14. <http://dx.doi.org/10.1128/genomeA.00224-14>.
7. Levenfors JP, Eberhard TH, Levenfors JJ, Gerhardson B, Hökeberg M. 2008. Biological control of snow mould (*Microdochium nivale*) in winter cereals by *Pseudomonas brassicacearum*, MA250. *BioControl* 53:651–665. <http://dx.doi.org/10.1007/s10526-007-9102-4>.
8. Ortet P, Barakat M, Lalaoua D, Fochesato S, Barbe V, Vacherie B, Santaella C, Heulin T, Achouak W. 2011. Complete genome sequence of a beneficial plant root-associated bacterium, *Pseudomonas brassicacearum*. *J Bacteriol* 193:3146. <http://dx.doi.org/10.1128/JB.00411-11>.
9. Zhou T, Chen D, Li C, Sun Q, Li L, Liu F, Shen Q, Shen B. 2012. Isolation and characterization of *Pseudomonas brassicacearum* J12 as an antagonist against *Ralstonia solanacearum* and identification of its antimicrobial components. *Microbiol Res* 167:388–394. <http://dx.doi.org/10.1016/j.micres.2012.01.003>.
10. Schwartz A, Ortiz I, Maymon M, Herbold C, Fujishige N, Vijanderaan J, Vilella W, Hanamoto K, Diener A, Sanders E, DeMason D, Hirsch A. 2013. *Bacillus simplex*—a little known PGPB with anti-fungal activity—alters pea legume root architecture and nodule morphology when coinoculated with *Rhizobium leguminosarum* bv. *viciae*. *Agronomy* 3:595–620. <http://dx.doi.org/10.3390/agronomy3040595>.
11. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
12. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol* 13:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.
13. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.

Titre : Génomique comparative des bactéries *Dickeya solani* et *Pectobacterium wasabiae*, pathogènes émergents chez *Solanum tuberosum*

Mots clés : Jambe noire, *Pectobacterium*, *Dickeya*, Génomique, Transfert horizontal, Phylogénie,

Résumé : Des bactéries pectinolytiques appartenant aux genres *Pectobacterium* et *Dickeya* sont des agents pathogènes chez *Solanum tuberosum*. Ces bactéries sont responsables de la maladie de la jambe noire et de la pourriture molle lors de la culture et du stockage des tubercules. Ce travail de thèse est divisé en deux axes : 1) Etude de la diversité d'une population du pathogène *D. solani* par approche de génomique comparée afin de mieux comprendre la structure génomique de cette espèce émergente, 2) L'assemblage du génome et la caractérisation génomique des facteurs de virulence chez *Pectobacterium wasabiae* RNS 08421A.

L'analyse des génomes de 20 isolats de *D. solani* issus d'environnements différents, par une approche de génomique comparative associée à des analyses fonctionnelles, a révélé une forte homogénéité génétique au sein de la majorité des souches (16/20). De plus, cette analyse a permis de caractériser un nouveau sous-groupe au sein de l'espèce *D. solani*, représenté par la souche 0512 (1/20). En revanche, d'autres isolats (3/20) montrent des variations de quelques centaines à quelques milliers de SNPs/InDels qui sont regroupés dans des îlots génomiques. Leur analyse

phylogénétique révèle qu'ils proviennent d'autres pathogènes par transferts horizontaux. Par ailleurs, l'analyse des fonctions affectées par les SNPs/InDels a permis de prédire, puis de vérifier sur pomme de terre, qu'un des isolats était faiblement virulent.

La deuxième partie de mon travail porte sur l'assemblage, la caractérisation et l'analyse du génome de la souche RNS 08.42.1A de *P. wasabiae*, qui a été isolée en France. La génomique comparative avec 3 autres souches de *P. wasabiae* d'origines géographiques différentes, a révélé à la fois une forte similitude au niveau de la séquence génomique (ANI > 99%) et une synténie conservée des gènes de virulence. En outre, notre analyse a mis en évidence une nette distinction entre ces quatre souches de *P. wasabiae* (isolées de *S. tuberosum*) et la souche type japonaise *P. wasabiae* CFBP 3304^T (isolée du raifort). Dans *P. wasabiae* RNS 08.42.1A, les gènes de synthèse et de perception du quorum sensing, *expI/expR*, présentent une plus forte homologie avec leurs orthologues chez *P. atrosepticum* et *P. carotovorum* (90%) qu'avec leurs homologues chez *P. wasabiae* (70%). Ceci suggère une acquisition de ces gènes par transfert horizontal au sein d'une population de pathogènes infectant la même plante hôte.

Title : Comparative genomics of the bacteria *Dickeya solani* and *Pectobacterium wasabiae*, emerging pathogens of *Solanum tuberosum*

Keywords : Soft rot, *Dickeya*, *Pectobacterium*, Genomics, Horizontal transfer, Phylogeny

Abstract: Some pectolytic bacteria *Pectobacterium* and *Dickeya* species cause important diseases on *Solanum tuberosum* and other arable and horticultural crops. These bacteria are responsible for blackleg in the field and tuber soft rots in storage and in transit as well as in the field worldwide. The main objectives of this thesis are: 1) To study the diversity of a *D. solani* population using comparative genomics approaches with the aim of understanding the genomic structure and evolution of this emerging species, 2) Characterization and genomic analysis of virulence factors in *Pectobacterium wasabiae* RNS 08421A.

Using comparative genomics approaches combined with functional assays, the analysis of the genomes of 20 isolates of *D. solani* from different environments, revealed a strong genetic homogeneity within the majority of the strains (16/20). Moreover, this analysis allowed to characterize a new sub-group within *D. solani* species, represented by the strain 0512 (1/20). In contrast, some strains (3/20) showed variations from hundreds to a few thousand of SNPs/Indels which are grouped in what we called "Genomic islands". Phylogenetic analysis of these regions showed that they

were acquired from other pathogens by HGT. Furthermore, the analysis of the functions affected by SNPs/Indels allowed predicting, then checking on potatoes, that one of these strain was less virulent.

The second part of my work involves assembly, characterization and analysis of the genome of the strain *P. wasabiae* RNS 08.42.1A, which was isolated in France. Comparative genomics with three other *P. wasabiae* strains from different geographical origins, revealed a strong similarity in the genome sequence (ANI > 99%) and conserved synteny of virulence genes. In addition, our analysis showed a clear distinction between the strains of *P. wasabiae* isolated from *S. tuberosum* and the type strain *P. wasabiae* CFBP 3304^T (isolated from horseradish). In *P. wasabiae* RNS 08.42.1A, the complex system for synthesis and perception of quorum sensing signal *expI/expR*, exhibit higher homology with their orthologs in *P. atrosepticum* and *P. carotovorum* (90%) than their homologs in *P. wasabiae* (70%). This suggests acquisition of these genes by HGT within a population of pathogens infecting the same host plant.

