

Practical use of metagenomics: the case study of the *Ralstonia solanacearum* species complex

Plant Health at the Age of Metagenomics
Paris, 26th September 2019

Genomics & the *R. solanacearum* species complex

Definitions

Genomics: analysis of the structure and function of genomes

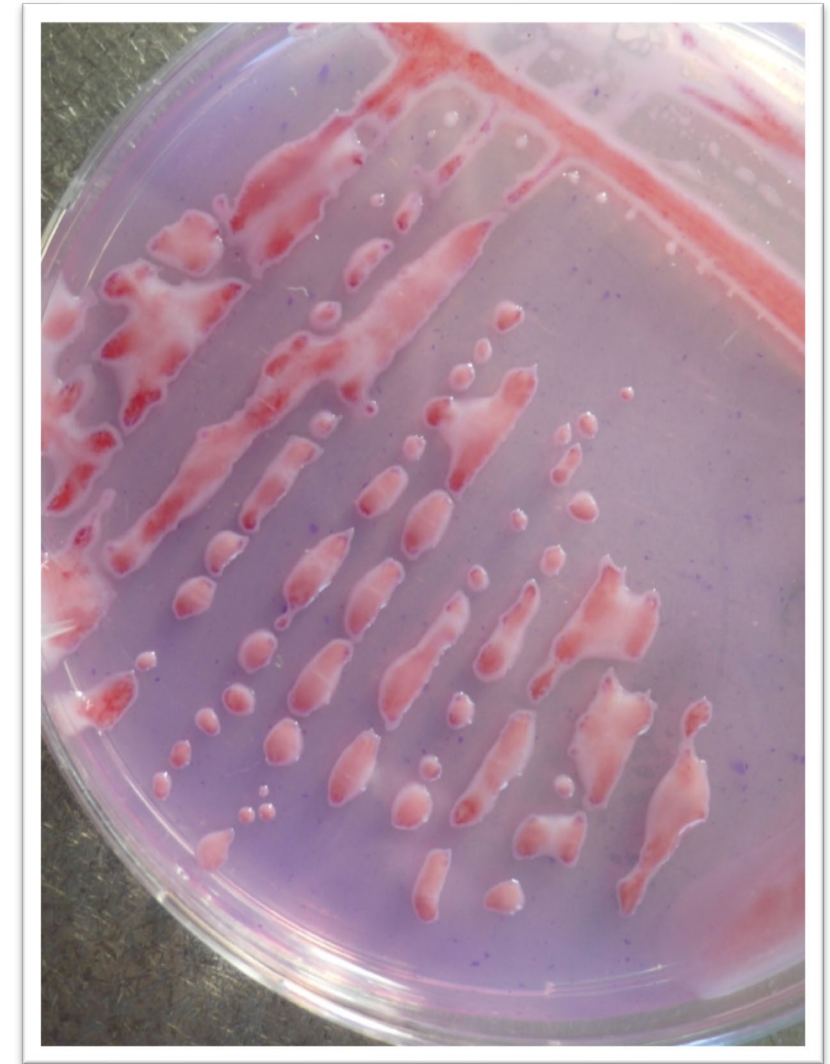
Comparative genomics : comparison of different sequenced genomes to better understand their biological functions through their commonalities and differences, and to study evolutionary history of pests

The *Ralstonia solanacearum* species complex (RSSC): a fitted model for studying plant-pathogen interactions, including basic biology of pathogenesis and non-host resistance in the context of an unusually broad host range and latent infections

Genomics & the *R. solanacearum* species complex

A complex plant pathogen

- **Broad genetic & phenotypic diversity**
- **Considered as a quarantine pest in many countries**
- Rhizospheric / soil born phytopathogenic bacteria
- Worldwide distribution with high host range diversity and high economic and social impacts
- 3 species : *R. solanacearum*; *R. pseudosolanacearum*; *R. syzygii*
- 4 Phylotypes / ~60 sequevars



Knowledge evolved with technologies

Past

90's

00's

10's

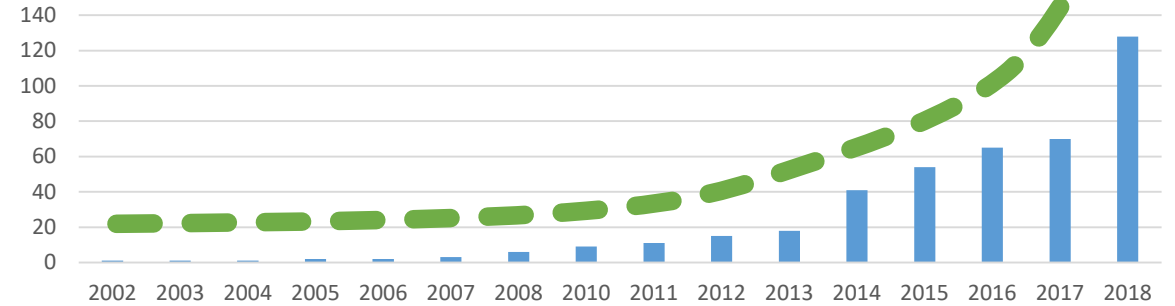
Now

Race & Biovar

Beginning of the PCR Era

Phylotype & Sequevar

Number of Sequenced Genomes for the RSSC



Source: GOLD/Berkeley University USA

1st Genome (2002)

Beginning of the Genomic Era

1st pangenome (2010)

1st genomic μ array (2010)

1st MLSA scheme (2012)

1st MLVA scheme (2013)



What have genomics allowed for the *R. solanacearum* species complex?

An essential step toward scientific breakthroughs

- i) Taxonomically reviewing species, mainly through high throughput sequencing and proteomic analysis
- ii) Unravelling the genetic background of well-known characterized phylogenetic lineages and identify lineage-specific features potentially involved in host range variation
- iii) Redefining plant-host relationship for specific lineages through genome sequencing and RNA sequencing (expression profiling)
- iv) Producing massive amounts of data on gene content to study epidemiology and for diagnostic applications

How genomics can help species delineation?

How genomics can help species delineation?

Two Rs studies confirmed the classification into 3 species

Safni et al. 2014	<ul style="list-style-type: none">• Phenotypic characterization• Whole-cell fatty acid composition• <u>DNA–DNA hybridization (DDH)</u>	Phenotype microarrays identified major variation in the core metabolisms, but without clear distinction between the three proposed species
Prior et al. 2015	<ul style="list-style-type: none">• Whole-genome sequencing• Proteomic analysis• Metabolic characterization	Full sequenced genomes allowed to dissect functional as well as genotypic differences in the denitrification metabolic pathway, which is associated with several quantifiable and biologically relevant phenotypic traits that play major roles in virulence (Dalsing et al. 2015)

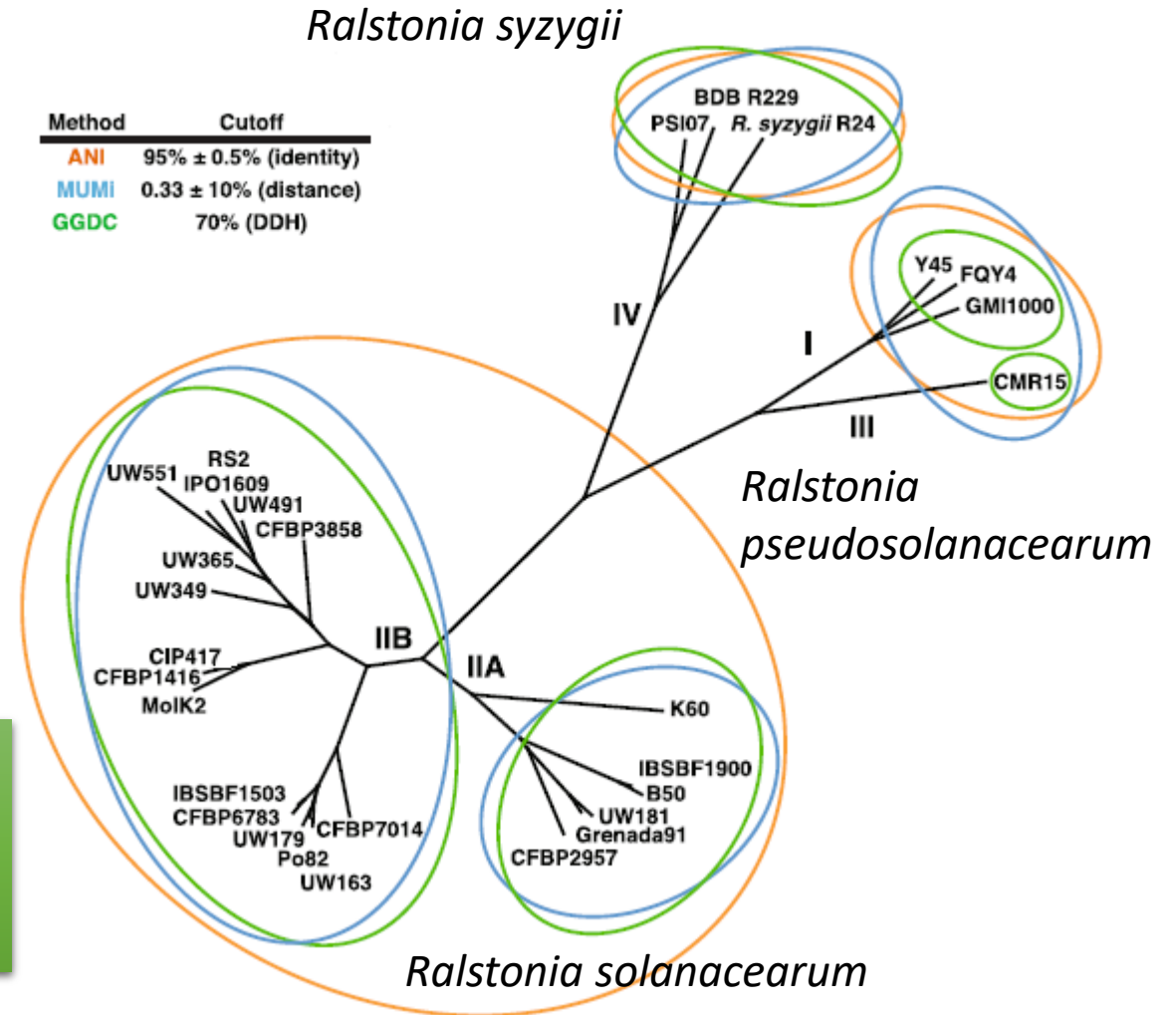
Additionally, DDH protocol showed significant drawbacks: it is technically difficult, is performed only in a few specialized laboratories, and is prone to experimental errors (Auch et al. 2010)

How genomics can help species delineation?

Two Rs studies confirmed the classification into 3 species

- The increased availability of genome sequences has advanced the development of genomic distance methods to describe bacterial diversity
- Results of these fast-evolving methods are highly correlated with those of the historically standard DNA-DNA hybridization technique
- However, these genomic-based methods can be done more rapidly and less expensively and are less prone to technical and human error. They are thus a technically accessible replacement for species delineation

Genomics applied to species delineation benefits many different applications, including breeding plant resistance to bacterial wilt, the identification of new pathological variants, management of quarantine containment...




How genomics can help understand host adaptation?

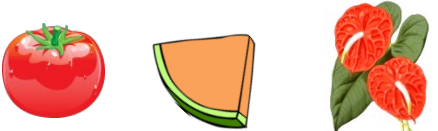
How genomics can help understand host adaptation?

Host-adapted strains from Phylotype II offer a model to study host adaptation


Moko
IIA-6/24 & IIB-3/4



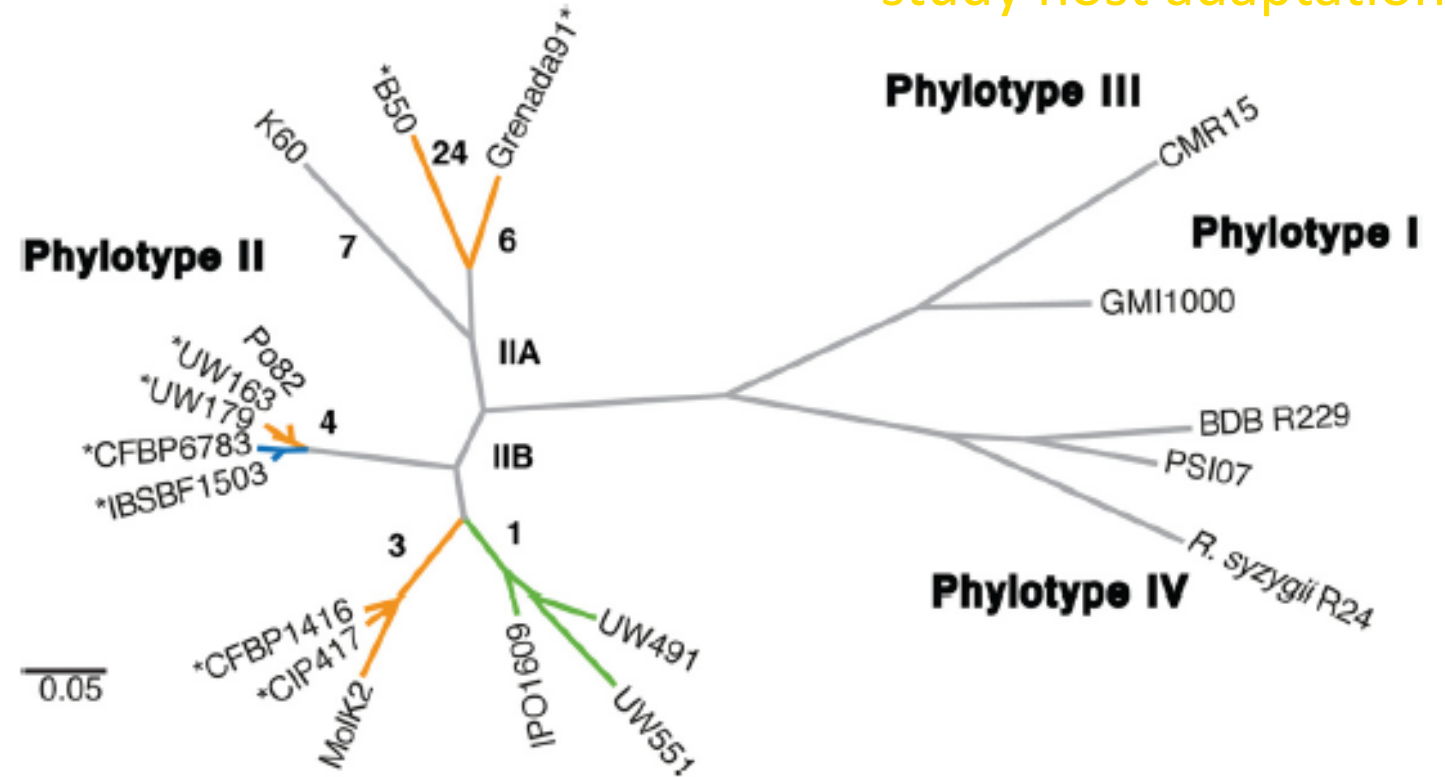
NPB
IIB-4



Brown rot
IIB-1



<24 °C

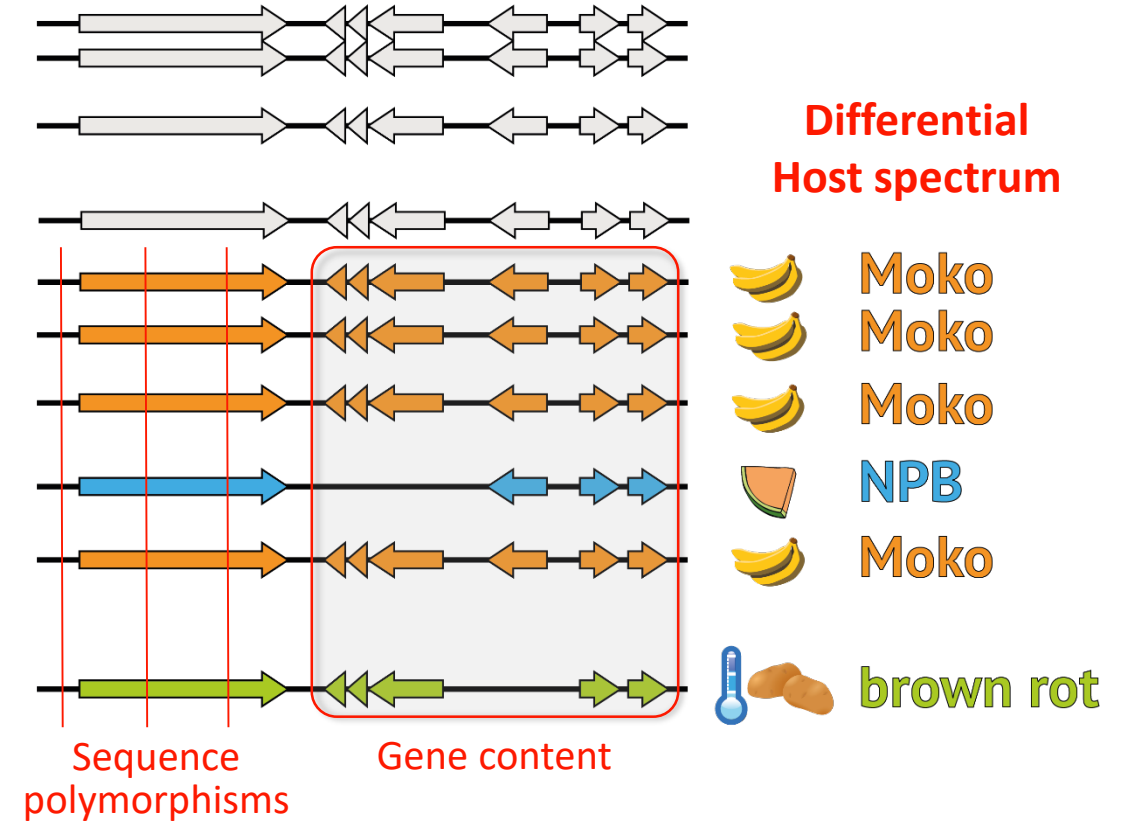
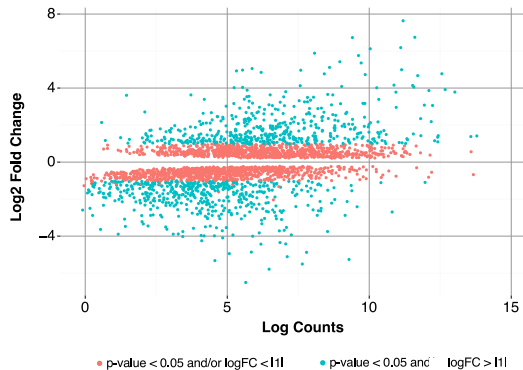


How to explain significant phenotypic trait differences in spite of close phylogenetic relationship?

How genomics can help understand host adaptation?

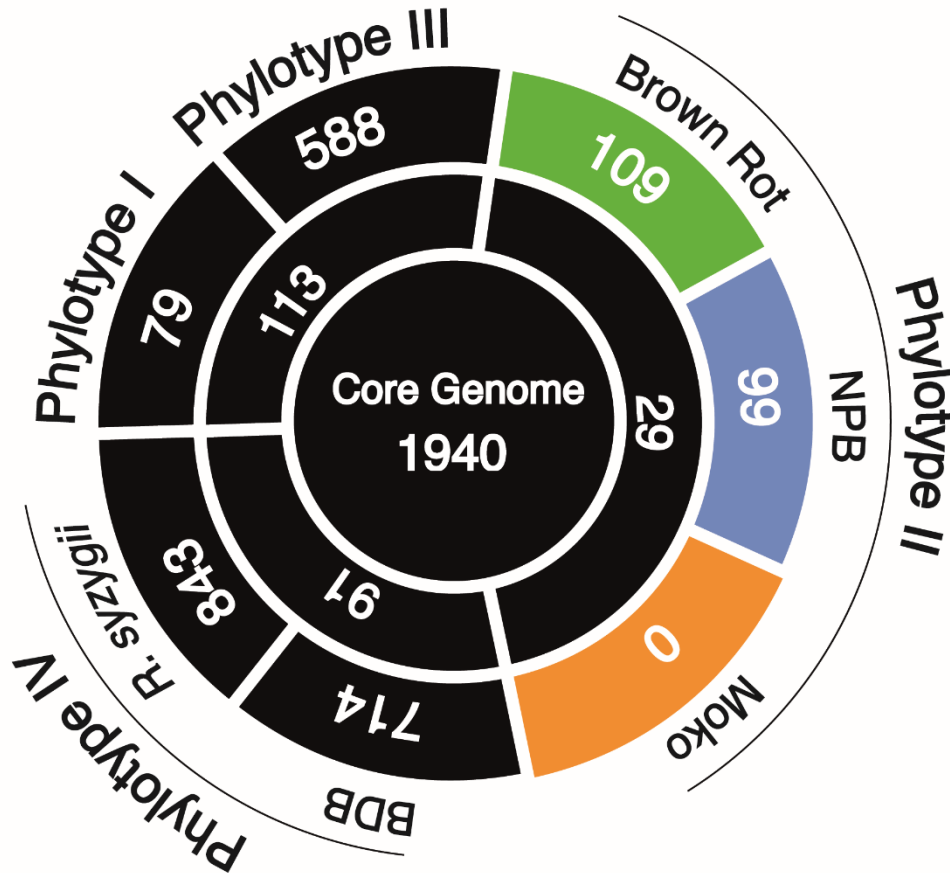
The search for Host-Adapted Polymorphisms features

Horizontal transfers



How genomics can help understand host adaptation?

Gene content appears to not be the lead to explain host specificity



19 genomes

Core-genome: 1,940 genes

Pan-genome: 16,757 genes

227 virulence factors

Few gene content associated with host specificity

How genomics can help understand host adaptation?

Type III effectors show high plasticity among the RSSC phylogeny

Pan-effectome
113 T3E

Core-effectome
14 T3E



...let's skip to the conclusion

How genomics can help understand host adaptation?

Deep genome analysis is the key to understand this complex plant pathogen

Few gene content associated with host specificity

Emergence of **NPB**, **Brown Rot**, and **Moko** associated to the evolution of the repertoire of type III effectors

Loss of effectors

Mutations

Horizontal transfers

Host range could evolve through changes in regulation

Genomics helps understanding pathogen evolution and can predict biological feature of a given ecotype

How genomics can improve the diagnostic field?

- *MOKO/NPB diagnostic PCR*
- *RSSC Research and diagnostic Microarrays*

How genomics can improve the diagnostic field?

- *MOKO/NPB diagnostic PCR*
- *RSSC Research and diagnostic Microarrays*

How genomics can improve the diagnostic field?

Find MOKO/NPB specific gene repertoires to design PCR primers



- **Blast** (*Basic Local Alignment Search Tool*)
- **GGB** (*Generic Genome Browser*) : graphical interface for various databases (sequence, annotation, synthenies...)
- **MaGe** (*Magnifying Genomes Microbial*) : annotation system of genomes of microorganisms

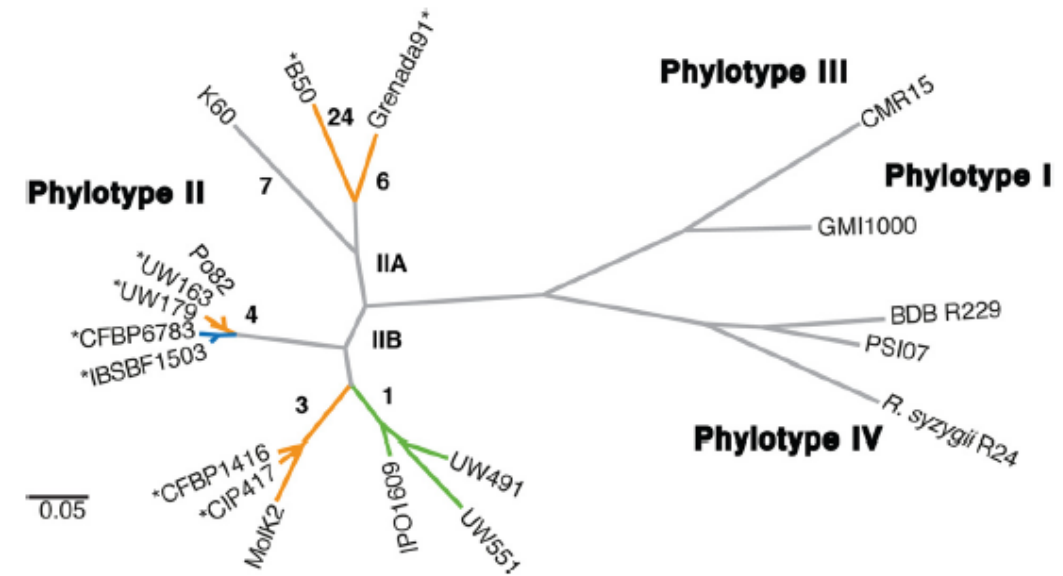
Available sequenced genomes under MaGe: 26

- Moko: 9
- NPB: 3

Results: Few specific genes in the MOKO strain ecotype, but selection of 2 candidates for design:

RALMO_2431: Integrase / recombinase like protein

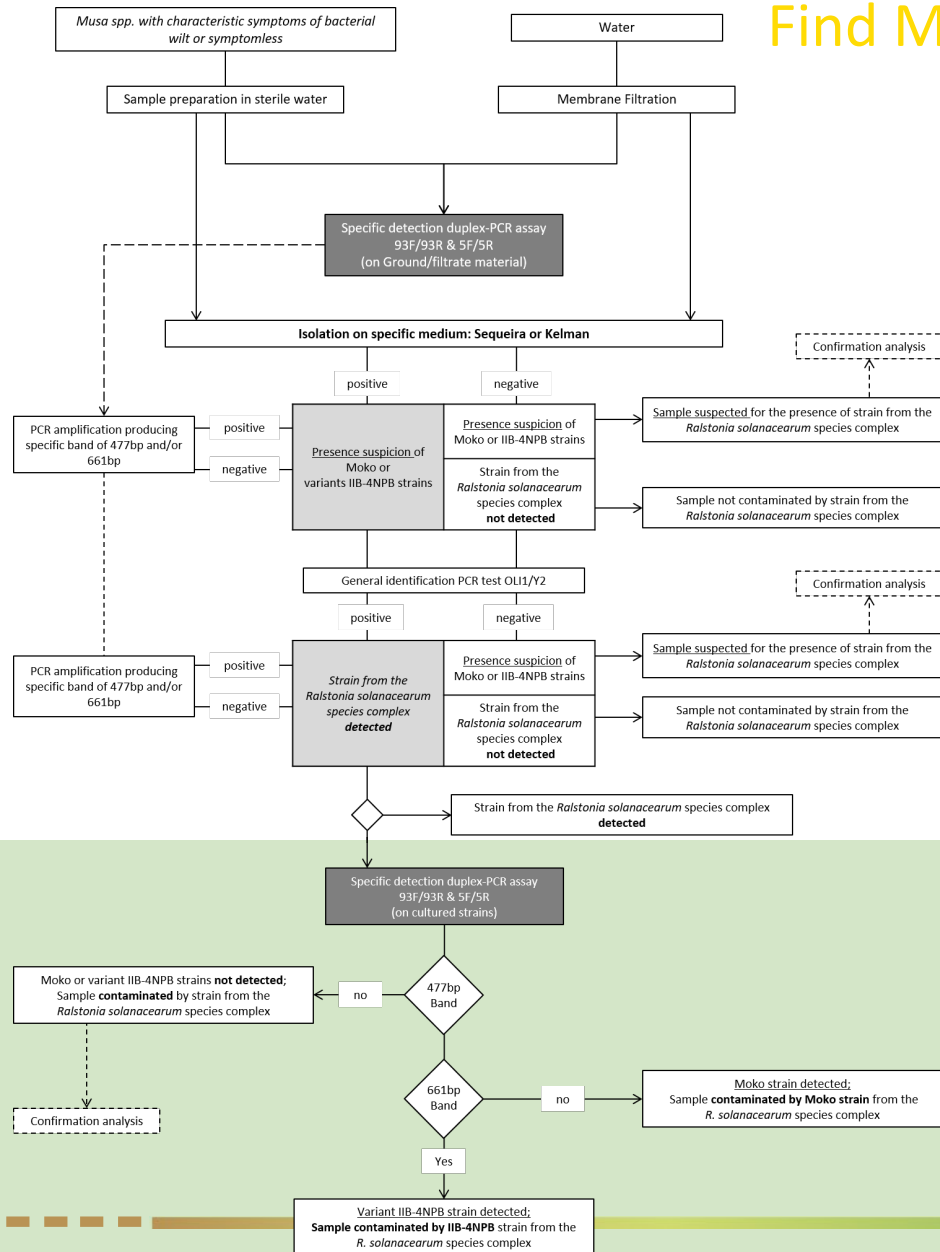
RALW3v1_108005: Binding protein KfrA like



Comprehensive genomic platform for the search of molecular markers

How genomics can improve the diagnostic field?

Find MOKO/NPB specific gene repertoires to design PCR primers



PLOS ONE

RESEARCH ARTICLE

A Duplex PCR Assay for the Detection of *Ralstonia solanacearum* Phylotype II Strains in *Musa* spp.

Gilles Cellier^{1*}, Aurélie Moreau¹, Aude Chabirand¹, Bruno Hostachy¹, Florent Ailloud^{1,2}, Philippe Prior³

1 French Agency for Food, Environmental and Occupational Health & Safety (ANSES), Plant Health Laboratory (LSV), Tropical Pests and Diseases Unit, Saint Pierre 97410, Reunion, France, **2** Centre de coopération internationale en recherche agronomique pour le développement (CIRAD), Unité Mixte de Recherche (UMR) Peuplements Végétaux et Bioagresseurs en Milieu Tropical (PVBMT), Saint Pierre 97410, Reunion, France, **3** Institut national de la Recherche Agronomique (INRA), Department of Plant Health and Environment (SP-E)—Centre de coopération internationale en recherche agronomique pour le développement (CIRAD), Unité Mixte de Recherche (UMR) Peuplements Végétaux et Bioagresseurs en Milieu Tropical (PVBMT), Saint Pierre 97410, Reunion, France

* gilles.cellier@anses.fr

anses
Agence nationale de sécurité sanitaire
alimentaire, environnementale, travail
Connaitre. Evaluer. Protéger.

Méthode d'analyse en santé des végétaux

RÉFÉRENCE : ANSES/LSV/MA 036 - Version 2
Juin 2018

Détection des souches responsables de la maladie de Moko et des variants IIB-4NPB dans le complexe d'espèces *Ralstonia solanacearum*, sur *Musa* spp. et dans l'eau

Wiley Online Library | **anses**

Bulletin OEPP
EPP0 Bulletin

Normes OEPP EPP0 Standards | [Free Access](#)

PM 7/21 (2) *Ralstonia solanacearum*, *R. pseudosolanacearum* and *R. syzygii* (*Ralstonia solanacearum* species complex)

First published: 17 April 2018 | <https://doi.org/10.1111/epp.12454> | Cited by: 3

Use of brand names of chemicals or equipment in these EPP0 Standards implies no approval of them to the exclusion of others that may also be suitable.

Laboratoire de la Santé des Végétaux
Laboratoire national de référence « Bactéries sur bananier, agrumes et plantes tropicales »

Le présent document est, sous sa forme électronique, mis à la disposition des utilisateurs en tant que méthode d'analyse. Ce document est la propriété de l'Anses. Toute reproduction, qu'elle soit totale ou partielle, n'est autorisée qu'à la condition expresse que la source soit citée, par exemple en faisant mention de sa référence (incluant sa version et année) et de son titre.

How genomics can improve the diagnostic field?

- *MOKO/NPB diagnostic PCR*
- *RSSC Research and diagnostic Microarrays*

How genomics can improve the diagnostic field?

Application through production of Research and Diagnostic microarrays

The microarray technology

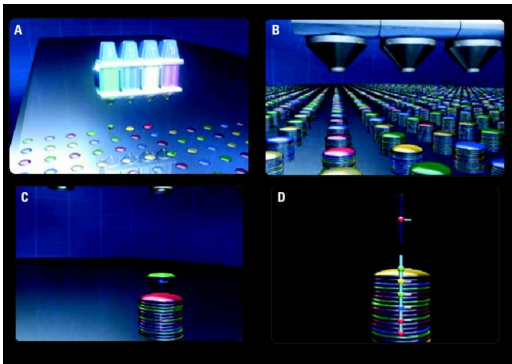
- Simultaneous detection and quantification of thousands of hybridization events
- Can test for a multitude of organisms in a single reaction: can screen for all biosecurity risk pathogens on specific hosts
- Great scope for miniaturization, high-throughput applications and development of integrated, automated systems

How genomics can improve the diagnostic field?

The Research pangenomic microarray (2010)

- Based on 6 fully sequenced strains within RSSC phylogeny
- 10,911 biological probes of 60mers
- 1 probe = 1 gene (and its orthologs)
- 3,500 control probes
- Agilent Technologies – *in situ* * Inkjet technology

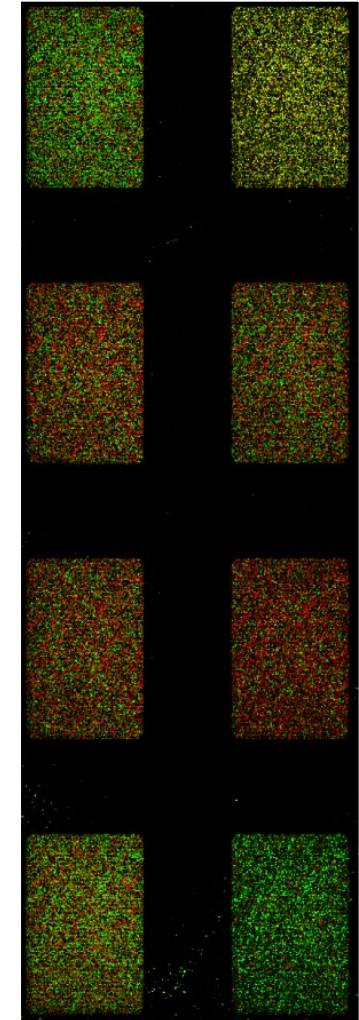
Inkjet technology



Oligo monomers are deposited uniformly onto specially-prepared glass slides

Gene Content

=



How genomics can improve the diagnostic field?

The Diagnostic Array Tube technology (2017)

Genomic data

in silico analysis (MaGe, NCBI)
target/non-target strains



Specific CDS Selection

17 groups (2-5 CDS / group)



Probe design

for validated CDS (50 mers)



Probes screening

Array glass slides (Genopole Toulouse)



Probes validation

Array Tube (Alere Technologies)

28 genomes for 17 major pathogenic and genetic subgroups including:

- 4 phylotypes
- Brown rot & Moko-causing strains
- Emerging strains IIB-4NPB



PCR Validation of CDS

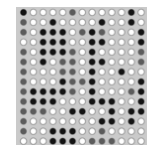
256 candidate probes

- 203 biological
- 53 control



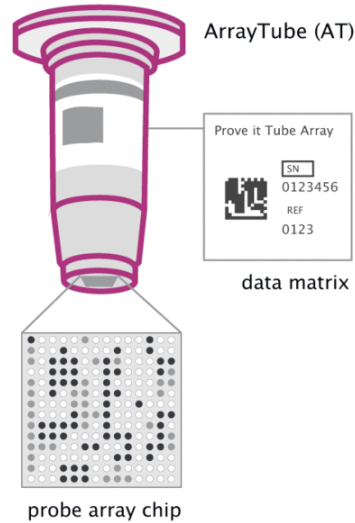
Validation on an 8-array glass slide

100 best probes in duplicate



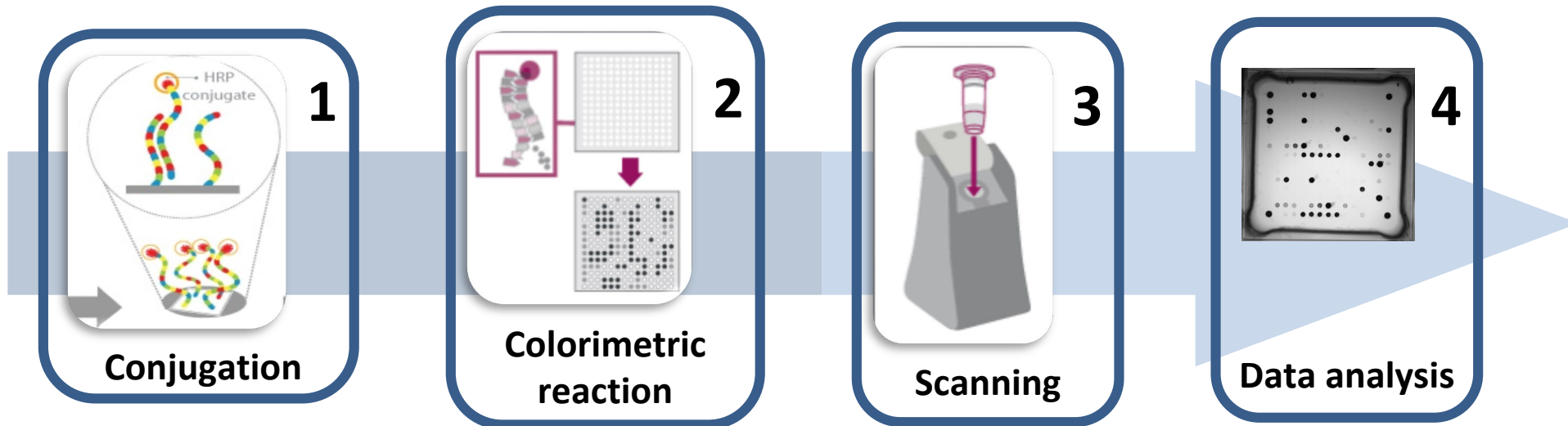
How genomics can improve the diagnostic field?

The Diagnostic Array Tube technology (2017)

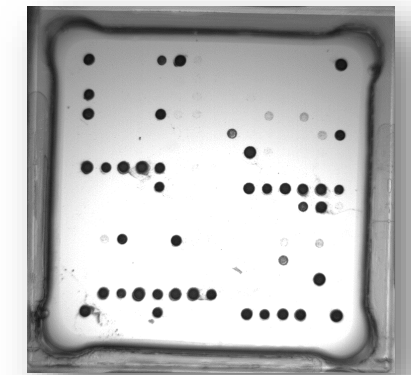
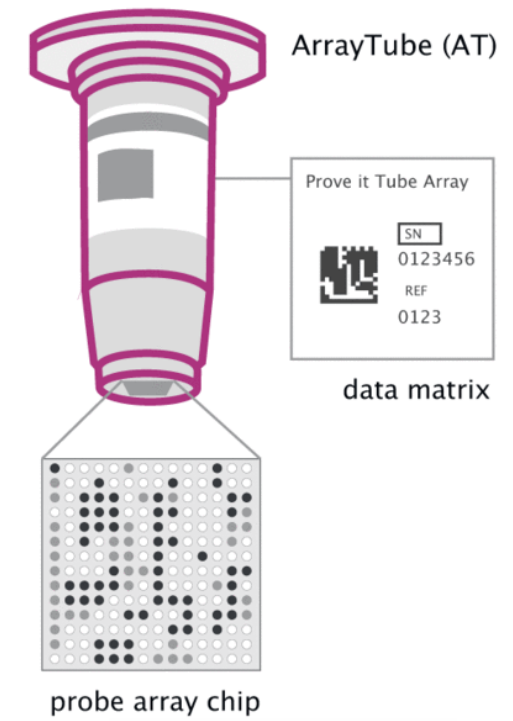
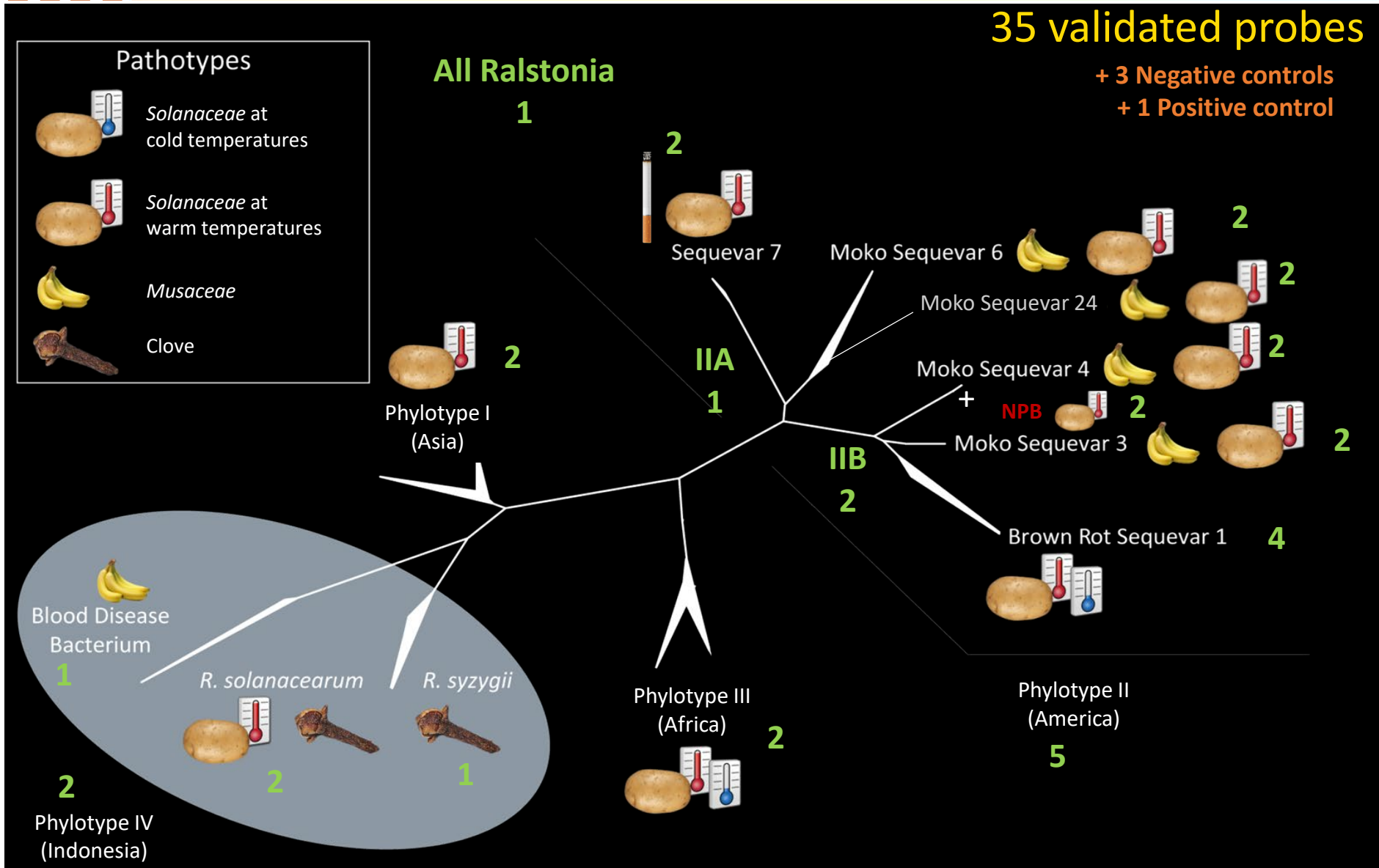


The **ArrayTube™** is a micro-reaction vial containing a probe based array chip at the bottom

- ✓ Custom probe arrays for serological and nucleic acid based formats
- ✓ Easy & user friendly processing with conventional lab equipment
- ✓ Cheap equipment compare to other microarrays or qPCR
- ✓ 3h protocol; no electrophoresis



How genomics can improve the diagnostic field?



How genomics can improve the diagnostic field?

Faster and more precise protocol development

- Production of dedicated diagnostic tools for both broad and specific strain tracking
- Creation of new applications for epidemiological purpose
- Increasing productivity and scientific developments toward a better and safer plant health

Genomics data allows to feed diagnostic applications that are useful for disease management, for producing diagnostic tools adapted to the most advanced scientific findings, and for providing guidance to decision makers

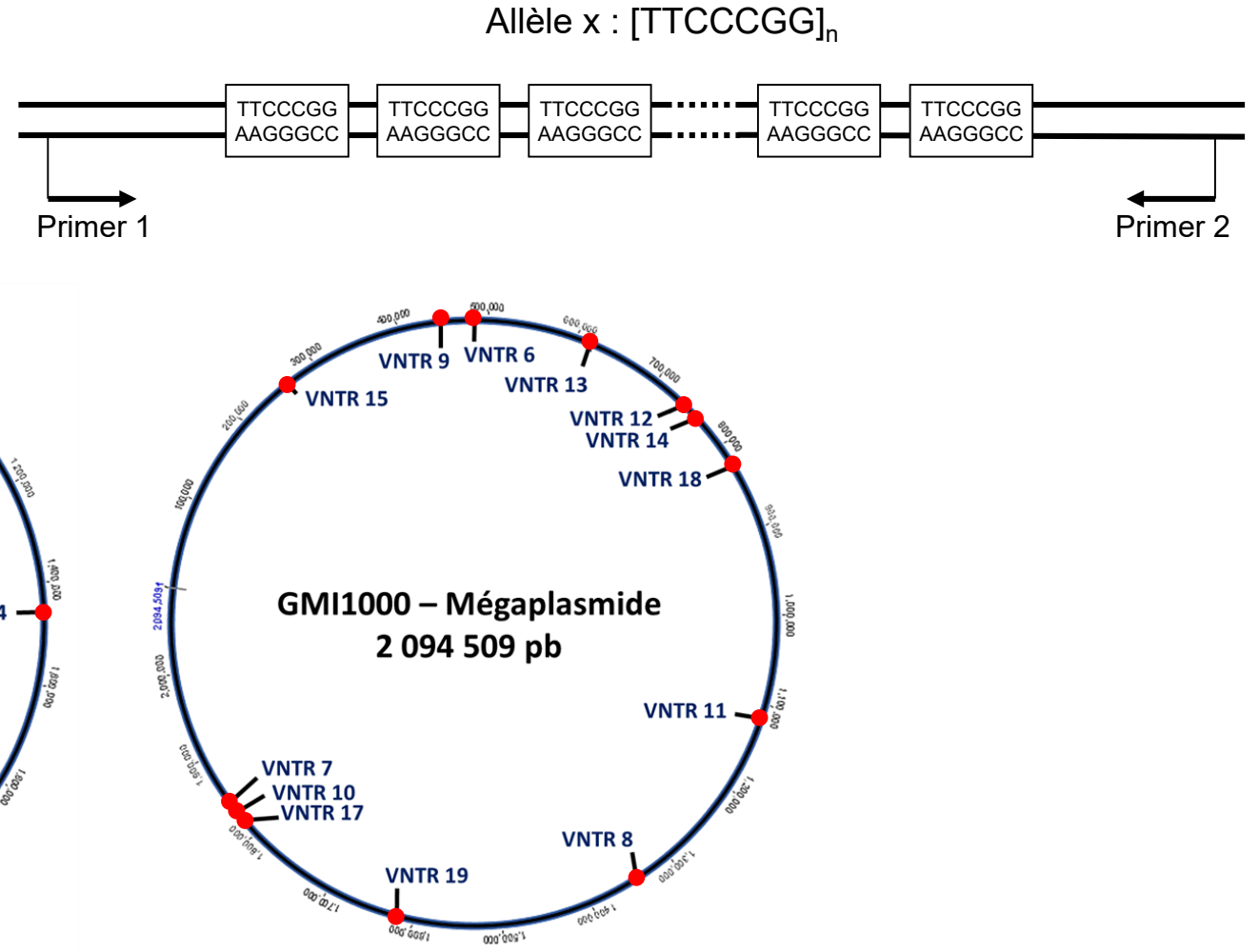
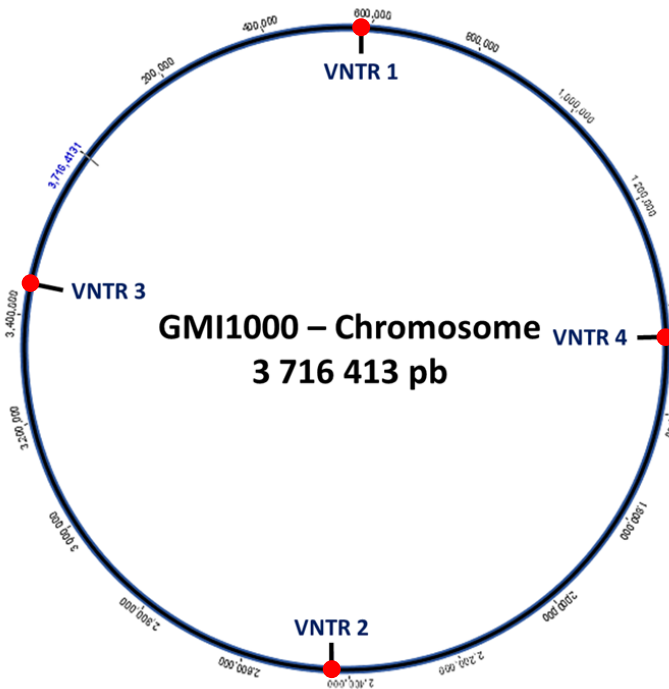
How genomics can improve Epidemiology?

How genomics can improve Epidemiology?

The Multiple-Locus Variable number tandem repeat Analysis

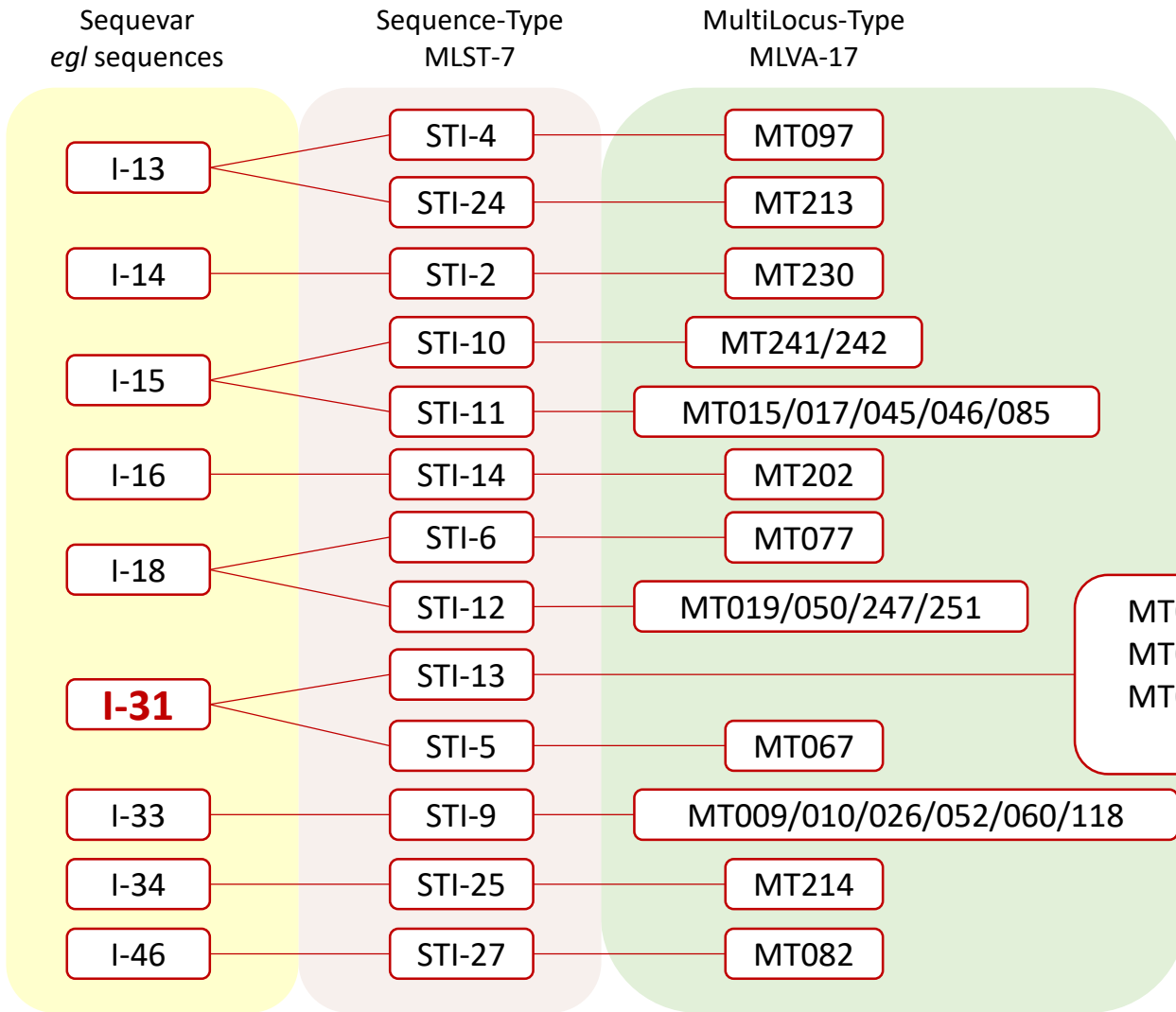
MLVA scheme

MT (Multilocus-Type= Haplotype): combination of alleles from different loci located on the same genome and transmitted together



How genomics can improve Epidemiology?

The Multiple-Locus Variable number tandem repeat Analysis



99 worldwide Phylotype I strains

MLST-7	MLVA-17
14 ST	45 MT
HGDI = 0,57	HGDI = 0,89

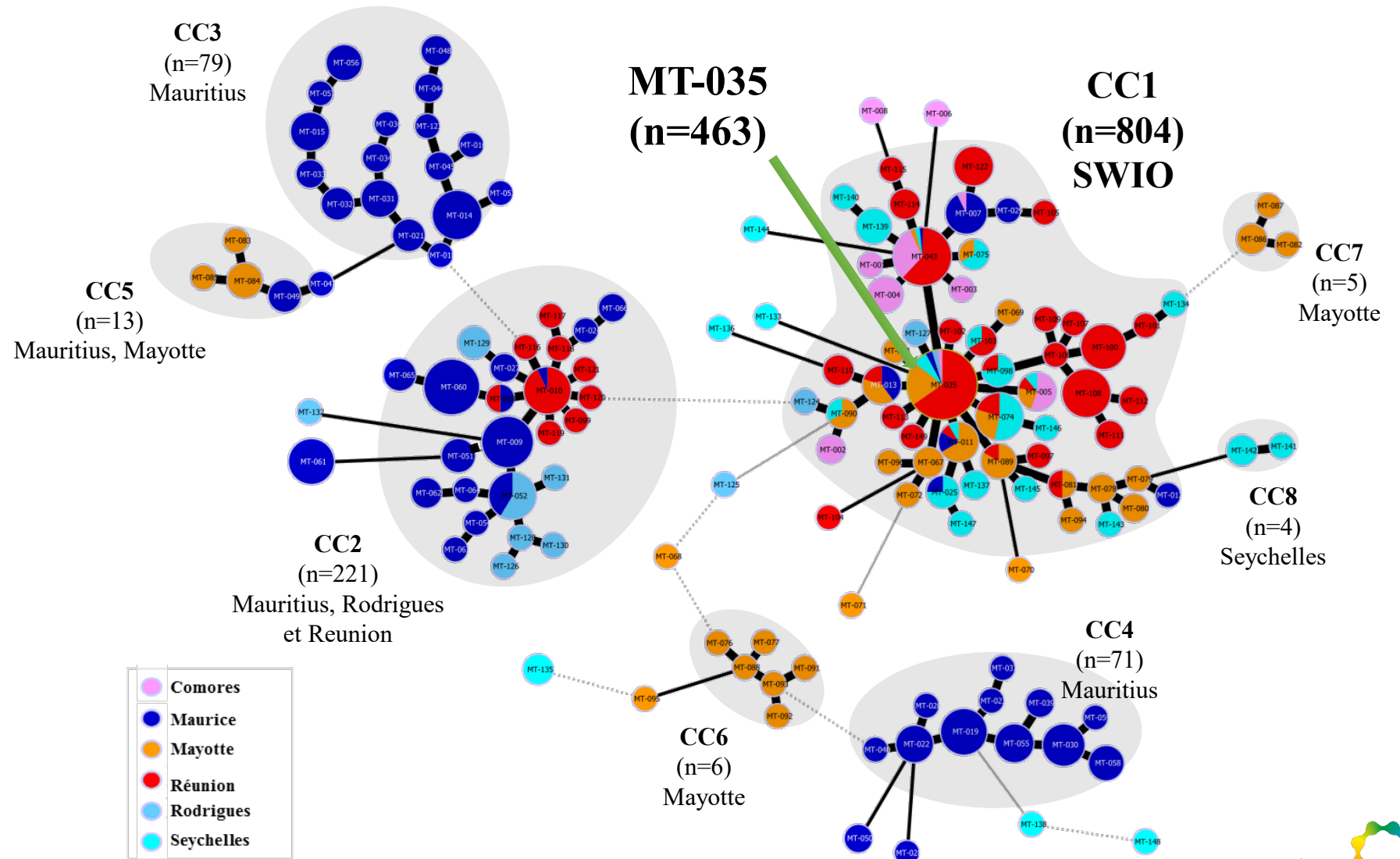
MT001 / MT002 / MT003 / MT004 / MT006 /
 MT011 / MT012 / MT025 / MT035 / MT043 /
 MT074 / MT075 / MT090 / MT103 / MT106 /
 MT108 / MT136 / MT138 / MT244

The phylogenetic signal is conserved and the resolving power has increased with the MLVA-17 scheme

How genomics can improve Epidemiology?

The Multiple-Locus Variable number tandem repeat Analysis

Genomic data allows to look deep inside pathogen key evolution markers. This helps for the identification of emerging clones that escape control strategies; and to trace bacterial strains that are important for the protection of agriculture



Final words

How did we manage genomics and other data ?

House made & ready to use pipelines

- Genomic data was produced by our team by sequencing full genomes and available at MaGe <http://www.genoscope.cns.fr/agc/microscope>, and NCBI
- Most of the genomic data was processed by personal scripts and pipelines, but are published and available upon request
- MLVA data was handled by Geneious, R statistical software, and Phyloviz, and are available in an international database and soon at NCBI
- Phylogenetic sequences are available at NCBI
- Data related to the RSSC microarray is available at the EMBL-EBI database
- RSSC strains are available at the CFBP database

An essential step toward scientific breakthroughs

What genomic data brings to Plant Health?

- i) Producing knowledge useful for biological characterization of plant pathogens (PRA)
- ii) Producing data for creating innovative and fitted diagnostics tools
- iii) Providing management strategies and guidance for decision makers

Importance of producing clean data, shared among international institutions

We need to think how to implement new metagenomics approaches in our daily work to improve the scientific knowledge and to produce strong scenario for disease management and containment

Practical use of metagenomics: the case study of the *Ralstonia solanacearum* species complex

Plant Health at the Age of Metagenomics
Paris, 26th September 2019