THAÍS RIBEIRO SANTIAGO

A DEEP ANALYSIS OF THE GENETIC STRUCTURE OF *Ralstonia solanacearum* IN BRAZIL REVEALS NOT MUCH SEX IN THE POPULATION!

Thesis presented to the Universidade Federal de Viçosa as part of the requirements for the degree of *Doctor Scientiae* in Plant Pathology.

VIÇOSA
MINAS GERAIS – BRAZIL
2014

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Câmpus Viçosa

T

S235d
2014

Santiago, Thaís Ribeiro, 1986-
A deep analysis of the genetic structure of *Ralstonia solanacearum* in Brazil reveals not much sex in the population! / Thaís Ribeiro Santiago. – Viçosa, MG, 2014.
ix, 74f. : il. (algumas color.) ; 29 cm.

Orientador: Eduardo Seiti Gomide Mizubuti.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Murcha bacteriana. 2. *Ralstonia solanacearum*. 3. Genética das populações. 4. Variabilidade. I. Universidade Federal de Viçosa. Departamento de Fitopatologia. Programa de Pós-graduação em Fitopatologia. II. Título.

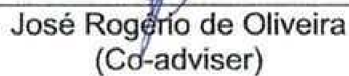
CDD 22. ed. 576.5

THAÍS RIBEIRO SANTIAGO

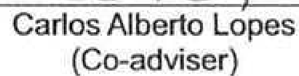
A DEEP ANALYSIS OF THE GENETIC STRUCTURE OF *Ralstonia solanacearum* IN BRAZIL REVEALS NOT MUCH SEX IN THE POPULATION!

Thesis presented to the Universidade Federal de Viçosa as part of the requirements for the degree of *Doctor Scientiae* in Plant Pathology.

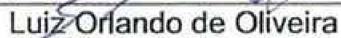
Approved: July 22nd, 2014.



José Rogério de Oliveira
(Co-adviser)



Carlos Alberto Lopes
(Co-adviser)



Luiz Orlando de Oliveira

Saulo Alves Santos de Oliveira

Eduardo Seiti Gomide Mizubuti
(Adviser)

To my mother Neide, my father and my sister Thássia, for all their love, sacrifice and constant support, always present in my heart.

ACKNOWLEDGMENTS

Thanks are due to the Universidade Federal de Viçosa, especially to the Plant Pathology Department and to Foundation for Research Assistance of Minas Gerais State (FAPEMIG) for granting a scholarship.

Thanks are also due to CNPq for providing financial support to the research.

Thanks to my parents Joaquim and Neide, my sister Thássia and my boyfriend Geraldo for their immense and unconditional support.

I am indebted to Professor Eduardo Seiti Gomide Mizubuti for his valuable advice, guidance, patience, friendship and opportunity.

Thanks are due to Professor José Rogério de Oliveira for his friendship and contributions to my professional education.

Thanks to Dr. Carlos Alberto Lopes for his contributions and suggestions for improving this work.

I am grateful to all my colleagues in the Graduate Program in Plant Pathology at UFV, especially my friends Robson, Saulo, Lahyre, Jaime, Miller, Sara, Jonas, Carlos, Alessandro, Maria, Adriana and Silvia. I am also grateful to the following members of the undergraduate programs at UFV: Sara, Lillian, Leonardo, Claudiney, Renato, Pollyany, Raquel, Luciano and Luana for their help and encouragement during this time.

Thanks to my friends Arshan, Fizza, Muhamad, Tânia and Kênia for their friendship and collaboration in my sandwich program.

Thanks are due to Professor Gustavo Caetano-Anolles for all his collaboration and guidance.

I am grateful to the Sara Moreira for her friendship and collaboration during this period.

Last but not least, thanks are due to all my colleagues who contributed to this work and were not mentioned here.

BIOGRAPHY

THAIS RIBEIRO SANTIAGO, daughter of Joaquim Santiago Sobrinho and Maria Rizoneide Ribeiro Rodrigues, was born in Brasiléia, AC, Brazil on November 10th, 1986.

In January of 2009, she graduated with a degree in Agronomy at Universidade Federal de Viçosa, Viçosa, MG.

In July 2010, she obtained her *Master Scientiae* degree in Plant Pathology from Universidade Federal de Viçosa, Viçosa, MG.

In July of 2010, she began her doctoral studies in the Plant Pathology Program at the Universidade Federal de Viçosa, Viçosa, MG, Brazil.

CONTENT

RESUMO ix
ABSTRACT xi
GENERAL INTRODUCTION........ 1
CHAPTER 1 - GENETIC STRUCTURE OF *Ralstonia solanacearum* CAUSING BACTERIAL WILT IN BRAZIL........ 4
1. Introduction 5
2. Material and methods........ 7
3. Results 12
4. Discussion........ 16
5. Conclusions........ 22
6. Acknowledgments 23
7. References........ 23
CHAPTER 2 - PARAMETRIZATION OF EVOLUTIONARY MECHANISMS THAT AFFECT THE POPULATION OF *Ralstonia solanacearum* IN BRAZIL USING GENE GENEALOGIES AND THE COALESCENT APPROACH........ 41
1. Introduction 41
2. Material and methods........ 43
3. Results 47
4. Discussion........ 49
5. Conclusion 54
6. Acknowledgements 54
7. References........ 54
GENERAL CONCLUSION 74

RESUMO

SANTIAGO, Thaís Ribeiro, D.Sc., Universidade Federal de Viçosa, agosto de 2014. **Uma análise profunda da estrutura genética de *Ralstonia solanacearum* no Brasil revela não muito sexo na população!** Orientador: Eduardo Seiti Gomide Mizubuti. Coorientadores: Carlos Alberto Lopes e José Rogério de Oliveira.

A murcha bacteriana, causada por *Ralstonia solanacearum*, causa perdas diretas e indiretas na produção de várias culturas. Apesar de o uso de variedades resistentes ser a melhor opção para o manejo da doença, nãoraro encontram-se relatos de suplantação da resistência. O primeiro relato de murcha bacteriana no Brasil é antigo, porém pouco se conhece sobre os biovares, filotipos, sequevares e variabilidade genética do patógeno. Por essa razão realizou-se o presente estudo. Isolados brasileiros de *R.solanacearum* foram caracterizados como biovar 1, 2 e 3, filotipo II, que está disperso por todo o país e filotipo I, que encontra-se predominantemente na região Norte. Sete sequevares foram identificados: 1, 4, 18, 27, 28, 41 e 50. Além disso, nós classificamos quatro novos sequevares no filotipo IIB como sequevar 53, 54, 55 e 56. Inicialmente, utilizou-se rep-PCR para estimar a variabilidade. Obtiveram-se 282 haplótipos. Uma possível explicação para o alto número de haplótipos é a ocorrência de recombinação. Isolados das regiões Sul, Sudeste e Central formam um grande grupo genético, enquanto aqueles das regiões Norte e Nordeste formam outro grupo. Tal fato evidencia fluxo gênico entre essas regiões, provavelmente mediado pelo transporte de tubérculos e mudas contaminadas. Detectou-se diferenciação genética entre isolados de tomate e batata coletados na região Sul-Sudeste-Central. Os mecanismos evolutivos foram parametrizados a partir de análises de genealogia de genes por processo coalescente utilizando sequências parciais de sete genes. Detectou-se evidência de subdivisão dos filotipos I, IIA e IIB, mas não subdivisão por hospedeiro. O fluxo gênico é mais intenso da região Sul para a região Norte. O filotipo II é uma linhagem ancestral que teve origem no Brasil e o filotipo I foi recentemente introduzido, possivelmente por ação antropogênica. O filotipo IIA deu origem ao filotipo IIB e essa diferenciação ocorreu possivelmente devido a

fatores ecológicos que precisam ser estudados com mais detalhes futuramente. Detectaram-se evidências de mutação, migração, recombinação e seleção na população de *R.solanacearum* no Brasil. Entretanto, mutação é mais importante que recombinação. Conclui-se que a utilização de variedades resistentes será uma grande desafio em todas regiões e hospedeiros.

ABSTRACT

SANTIAGO, Thaís Ribeiro, D.Sc., Universidade Federal de Viçosa, August, 2014. **A deep analysis of the genetic structure of *Ralstonia solanacearum* in Brazil reveals not much sex in the population!**Adviser: Eduardo Seiti Gomide Mizubuti. Co-advisers: Carlos Alberto Lopes and José Rogério Oliveira.

Bacterial wilt, caused by the *Ralstonia solanacearum*, cause direct and indirect losses in several crops. Planting of resistant varieties is the best option for disease management, but reports of resistance breakdown are commonly found in the literature. Although many years have passed after the first report of this pathogen in Brazil, information on biovars, phylotypes, sequevars and genetic variability of the pathogen is scarce. In the present study isolates were characterized as biovar 1, 2 and 3. Phylotype II is spread throughout Brazil and phylotype I is predominantly found in the North. Seven sequevars were identified: 1, 4, 18, 27, 28, 41 and 50. Moreover, we classified four new sequevars in the phylotype IIB as sequevar 53, 54, 55 and 56. Initially, we used rep-PCR to estimate the variability of the pathogen and 282 haplotypes were obtained. The high number of haplotypes could be due to the occurrence of recombination. Isolates of the South/Southeast/Central regions formed a genetically related cluster. Isolates from the North/Northeast regions formed another group. Gene flow occurs through the transportation of contaminated tubers and seedlings. Genetic differentiation was detected among isolates from tomato and potato collected in the South/Central/Southeast regions. In addition, gene genealogies based on the coalescent process were used to infer about evolutionary mechanisms. We detected evidence of subdivision of phylotypes I, IIA and IIB, but no subdivision by hosts. The gene flow is predominantly from the southern to the northern regions. We confirmed that phylotype II is an ancestral lineage that originated in Brazil and probably phylotype I was recently introduced by anthropogenic actions. Phylotype IIA originated phylotype IIB and this difference was probably due to ecological factors that need to be studied in more detail. Mutation, migration, recombination and selection occur in the population of *R. solanacearum* in Brazil, however mutation is more important

than recombination. We conclude that the use of resistant varieties is a major challenge in all regions and hosts.

GENERAL INTRODUCTION

Bacterial wilt caused by *Ralstonia solanacearum* is one of the most destructive diseases in various crops cultivated in the tropics. The inoculum can be dispersed by infested soil, irrigation water and tubers (Caruso et al., 2005; Denny, 2006). Several species of plants, distributed in more than 50 botanical families, are reported as hosts (Denny, 2006). Among the economically important crops, those belonging to the Solanaceae family are relevant, such as tomatoes, potatoes, peppers and eggplant (Hayward, 1994).

In an attempt to study the variability, different classification methods commonly used in bacteria species have been adopted. The main divisions are biovar, phylotype and sequevar (Buddenhagen et al., 1964; Fegan; Prior, 2005). Methods based on genome analysis, analysis of repeated sequences and sequencing of genomic regions are useful to elucidate the relationships within the species *R. solanacearum* (Strange, 2005). Although there is information on the genetic variability of *R. solanacearum* in Brazil and its geographical distribution, these studies were conducted for a reduced number of hosts and sampling areas (Costa et al., 2007; Pinheiro et al., 2011; Rodrigues et al., 2012; Santana et al., 2012). The knowledge about the pathogen population structure is useful for the understanding and development of control strategies and to guide breeding programs that aim to develop cultivars with durable resistance (Milgroom and Fry, 1997).

Despite several efforts, epidemics of bacterial wilt remain responsible for the low productivity in several regions. Under favorable temperatures, the appearance of wilt occurs initially in upper leaves and after a few days in all plant (Hayward, 1991; Akiew and Trevorrow, 1994). The difficulty of developing bacterial wilt control strategies have been caused by the lack of knowledge of ecology, variability and about the evolutionary mechanisms that affect the pathogen population (Strange, 2005). The limited knowledge of the variability in an area, makes selection of resistant cultivars in breeding programs very difficult. Resistant cultivars to bacterial wilt have been developed for various hosts such as potatoes, tomatoes and eggplant. However, cultivar resistance is often not efficient when planted in different areas (Lopes, 1994).

The objectives of this study were to: (1) characterize the isolates of *R. solanacearum* by biovar, phylotype and sequevar, (2) determine the genetic structure of pathogen population by geographic region and by host and (3) parameterize the evolutionary mechanisms affecting the population of *R. solanacearum* in Brazil.

References

Akiew, E. B. and Trevorrow, P. R. 1994.Management of bacterial wilt of tobacco.Pages 179-198.in Bacterial Wilt: The disease and its causative agent, *Pseudomonas solanacearum*. A. C. Hayward and G. L. Hartman, eds. CAB International, Wallingford, UK.

Buddenhagem, I.; Sequeira, L. and Kelman, A. 1962.Designation of races in *Pseudomonas solanacearum* (Abstr.).Phytopathology 52: 726.

Casuro, P.; Palomo, J. L. Bertolini, E., Alvarez, B., Lopez, M. M., Biosca, E. G. 2005. Seasonal variation of *Ralstonia solanacearum* biovar 2 populations in a Spanish river: recovery of stressed cells at low temperatures. Applied and Environmental Microbiology 71: 140-148.

Costa, S. B., Ferreira, M. A. S. V. and Lopes, C. A. 2007. Diversidade patogênica e molecular de *Ralstonia solanacearum* da região amazônica brasileira. Fitopatologia Brasileira 32:285-294.

Denny, T. P. 2006. Plant pathogenic *Ralstonia* species.Pages 573-644 in Plant-associated *Bacteria*. S. S. Gnanamanickan. The Nertherlands, Springer Publishing

Fegan, M. and Prior, P. 2005. How complex is the "*Ralstonia solanacearum* species complex"? Pages 449-461 in: Bacterial wilt disease and the *Ralstonia solanacearum* species complex. C. Allen, P. Prior, and A. C. Hayward. American Phytopathological Society Press, St. Paul, MN.

Hayward, A. C. 1991. Biology and epidemiology of bacterial wilt caused by *Pseudomonas solanacearum*. Annual Review Phytopathology 29: 65-87.

Hayward, A.C. 1994. The hosts of *Pseudomonas solanacearum*.Pages 9-24 in Bacterial Wilt: The disease and its causative agent, *Pseudomonas solanacearum*.A. C. Hayward and G. L. Hartman, eds. CAB International, Wallingford, UK.

Lopes, C. A. 1994. Situação da murcha bacteriana no Brasil. Pages 13-16 in Taller sobre enfermedades bacterianas de la papa. Embrapa CNPH CIP, Brasília, DF.

Milgroom, M. G. and Fry, W. E. 1997.Contribution of population genetics to plant disease epidemiology and management.Advances in Botanical Research 24:1-30.

Pinheiro, C. R., Amorim, J. A. E., Diniz, L. E. C., Silva, A. M. F., Talamini, V., Souza Júnior, M. T. 2011. Diversidade genética de isolados de *Ralstonia solanacearum* e caracterização molecular quanto a filotipos e sequevares. Pesquisa Agropecuária Brasileira46:593-602.

Rodrigues, L. M. R., Destéfano, S. A. L., Silva, M. J., Costa, G. G. L. and Maringoni, A. C. 2012. Characterization of *Ralstonia solanacearum* from Brazil using molecular methods and pathogenity tests.Journal Plant Pathology 94: 505-516.

Santana, B.G., Lopes, C. A., Alvarez, E., Barreto, C. C., Allen, C. and Quirino, B. F. 2012. Diversity of Brazilian biovar 2 strains of *Ralstonia solanacearum*. Journal of General Plant Pathology 78: 190-200

Strange, R. N. 2005.Plant Disease: A threat to global food security.Annual Review Phytopathology 43:83-116.

CHAPTER 1

GENETIC STRUCTURE OF *Ralstonia solanacearum* CAUSING BACTERIAL WILT IN BRAZIL

T. R. Santiago[1], C. A. Lopes[2], G. Caetano-Anollés[3], and E. S. G. Mizubuti[1,4]

Abstract: Understanding the genetic variability of populations of the plant pathogenic β-proteobacterium *Ralstonia solanacearum* (Rs)is key to the implementation of control measures, mainly breeding for disease resistance. In order to study the genetic structure of Rs in Brazil, 301 isolateswerecollectedfrom nine plant species and from all major geographic regions. Isolates were characterized according to biovar, phylotype, and sequevar to determine the amount and distributionof geneticvariability. Isolates were classified into biovars 1 (52%), 2 (32%) and 3 (16%) and phylotypes I (48 isolates) and II (253 isolates) with two subclusters, IIA and IIB. Phylotype II was found to be widespread in Brazil, whereas Phylotype I isolates were found in the Central, North and Northeast regions, the warmest areas. Based on the endoglucanase (*egl*)gene sequences, seven known sequevars (1, 4, 18, 27, 28, 41 and 50) and four new sequevars (54, 55, 56, 57) were identified. Large clonal fraction and high number of haplotypes (282)were observed inBrazil. Three groups were identified: group 1 made of phylotype I isolates; and two other groups comprised of phylotypes I and II isolates but split according to geographic regions; group 2 made of isolates collected in the North/Northeast and group 3 from the Central/Southeast/South regions. Absence of genetic separation of the South-Southeast-Central and North-Northeast regions supports the occurrence of gene flow at different latitudes probably due to long distance dissemination of planting materials or rivers. There was evidence of recombination in the population. Among the isolates collected in the South-Southeast-Central regions those from tomato were genetically distinctfrom the potato isolates. Thereishighgeneticvariabilityin theBrazilian population of*R. solanacearum*. Thismakesbacterialwiltcontrolusinghostresistance achallengingtask.

Key words: Brown rot,population genetics, molecular epidemiology, variability, diversity.

[1] Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil.
[2] Centro Nacional de Pesquisa de Hortaliças, Gama, Distrito Federal, Brazil.
[3] Department of Crop Sciences, University of Illinois, Urbana, Illinois, United States.
[4] e-mail: mizubuti@ufv.br.

1. Introduction

Ralstonia solanacearum [(Smith, 1896) Yabuuchi et al. 1995], previously known as *Pseudomonas solanacearum* and *Burkholderia solanacearum*, is the causal agent of bacterial wilt in many plant species (Yao & Allen, 2006). Despite its wide distribution, epidemics are more intense in the tropical regions of the world (Hayward, 1994). There are recordsof occurrenceof bacterial wilt inplantsfrom more than50botanical families, includingmonocots and dicots (Elphinstone, 2005; He et al., 1983). Economic lossesare difficult to quantify, but are high. While an evaluation of direct yield lossesispossible,estimatingindirect losses commonly associated with bacterial wiltis not straight forward. Specifically, the occurrenceof bacterial wiltin certain areasmay hamper the cultivation ofsusceptiblehosts, which imposes severerestrictions on the economic viabilityof land use. Manycrops of economic importance such as potato *(Solanum tuberosum*), tomato (*S. lycopersicum*), eggplant (*S. melongena*), banana (*Musa* spp.) and peanut (*Arachis hypogea*)are affected by the pathogen. In addition, recent reports show that the range ofplantsinfected by *R.solanacearum*is increasing, including in Brazil (Wilson et al., 2005; Malavolta, 2008).

*Ralstoniasolanacearum*has highphenotypicand genotypic variability and strains have been isolated from virgin jungle soils in Asia and the Americas (Hayward, 1991). Currently, *R. solanacearum* is considered a species complex and there is variation according to host range, geographic distribution, pathogenicity and physiological properties (Denny, 2006). The origin of the species complex is believed to predate the geological separation of the continents (Hayward, 1991).

Before thegenomic revolution, the characterization ofraces was used toseparate thestrains of*R.solanacearum* in relation to the range of hosts(Buddenhagem et al., 1962), whereas biovarswere defined asphysiologicalgroups that vary according to the abilityto usecarbohydrates(He et al., 1983). Although widely used, the classification based on phenotypic attributes does not permit the reliable categorization that is typically obtained with more powerful methods, such as those involving DNA sequences. Recently, new groups were proposed based on the nucleotide sequence of the rRNA interspacer ITS region and the Hrp B (hrpB) gene region and

endoglucanase (egl) genes (Fegan &Prior, 2005). Four phylotypes that make use of multiplex PCR and various sequevars derived from egl gene sequences variations were identified. Nowadays, this classification is the most accepted and is used in a number of studies of *R. solanacearum* strains from different regions, including the Americas.

Despite their considerable phylogenetic diversity, *R. solanacearum* strains are unified by their common pathology. All cause bacterial wilt, which is characterized by bacterial colonization of the xylem vessels to very high cell densities (10^9- 10^{10} CFU/mL xylem fluid), vascular browning, stunting, wilting, and often rapid death (Kelman, 1956; Buddenhagen et al., 1964; Peeters et al., 2013). The bacterium is dispersed by soil, water and infected propagation material. Strains infect plants most commonly through the roots, but sometimes through insect wounds (Hayward, 1991).

Bacterial wilt is difficult to control because the bacteria survive for years in infested soils and weed hosts. Breeding for host resistance is the best management strategy, but is complicated by the high genetic variability of the pathogen population. For example, tomatoes resistant to *R. solanacearum* in one region can be susceptible in another (Boshou, 2005). The lackof knowledge aboutthe variabilityof the pathogen, its magnitude and geographic distribution, make difficult the development and the strategic use of resistant varieties inseveralareas, particularly in Brazil.Thestrategic useof resistant cultivarsdepends on basic informationrelatingthe environmental conditions, the characteristics of the hostand genetic variabilityofthe pathogen population, which is often subject to changes. Therefore, successfulbreeding programsrequire knowledge of theevolutionary biologyof the pathogenand the mechanismsthat influencetheir pathogenicity(Strange, 2005). Since the variability of the pathogendirectly affects themanagement of the disease, epidemiological studies are also important.

While knowledge of the local pathogen population is required for successful breeding and integrated pest management programs, there is scarce information about the genetic variability of the Brazilian populations of *R. solanacearum*. It has been proposed that *R. solanacearum* which was putatively originated in Indonesia/Australia was possibly introduced to Brazil before the fragmentation of Gondwana (Castillo & Greenberg, 2007; Wicker et al., 2012).

Thus, the Brazilian population is very old and has had the opportunity to co-evolve with a high diversity of plant species throughout millions of years, but with few commercially grown hosts that were introduced and have been more intensively cultivated in the past 200 years. In one study, isolates from tomato and a few other hosts sampled in the Amazon region were found to be highly polymorphic and were classified into five groups that had no correlation with host range, biovar, ecosystem and sampling location of the pathogen (Costa et al. 2007). Similar results were observed for isolates from different hosts and different regions of the country (Rodrigues et al., 2012). However, the sampling limitations of both studies (only 22 isolates) restrict the scope of inferences. In another study, 53 *R. solanacearum* strains mainly from potato collected in seven Brazilian states were characterized using phylotype and sequevar (Santana et al., 2012). There was high genetic variation, but information on the amount and distribution of genetic variation in populations of *R. solanacearum* associated to other host species remains largely unknown. In this study, we are interested in answering the following questions: Which biovar, phylotype and sequevar are found in Brazil? Is the genetic diversity high? Is there spatial genetic differentiation or any population structure according to host? Is there evidence of recombination in the population of *R. solanacearum* in Brazil?

2. Material and methods

2.1. Bacterial strains and growth conditions

Bacterial strains used in this work are listed in Table 1 and Fig. 1. The isolates are deposited in the culture collection maintained by EMBRAPA Hortaliças (Brasília, Brazil) and Universidade Federal de Viçosa. A total of 301 isolates were analyzed. Isolates were obtained from different host plants, including *Solanumlycopersicum* (137 isolates), *S. tuberosum* (98), *Capsicum* spp. (4), *Piperhispidinervum* (20), *S. melongena* (19), *Eucalyptus* sp. (15), *Musa* spp. (2), *S. gilo* (3), and *Geranium* sp. (3), collected from 1987 to 2012, from different states in Brazil. When distributed according to geographic regions, 94 isolates were from the North, 37 from the Northeast, 53 from the Central, 41 from the Southeast and 76 from the South. These isolates were grown in

Kelman agar medium (Kelman, 1954) incubated for 48h at 28℃ and maintained in cryogenic storage tubes containing sterilized water.

2.2. DNA extraction and identification of bacterial strains

Bacteria were streaked on Kelman solid medium without tetrazolium chloride and incubated at 28℃ for 48 h. Bacterial growth from a single, isolated, and typical colony was transferred to another plate and incubated in the same way. Two mL of saline solution (0.01%) was added to plates with developed colonies that were scraped so as to prepare a cell suspension. The entire suspension was then pelleted by centrifugation and total genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega) following the manufacturer`s instructions. The quality of DNA was checked by gel electrophoresis. DNA was quantified with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) and adjusted to 25 ηg.

2.3. Species specific primers

The species-specific primer pair 759/760 designed to identify *R. solanacearum* was used in all isolates. PCR conditions were as described by Opina et al. (1997) in a total volume of 25 µL. Amplification was performed in a MJ PTC-100 thermocycler. PCR products were analyzed by eletrophoresis in 1% (wt/vol) agarose gels with 0.001 µg/mL GelRed (Biotium), and photographed under UV light. Fragments were compared with a 100 base pair (bp) DNA ladder. Control reactions without DNA were made to validate the results. A positive identification was based on the presence of a 282 bp amplicon.

2.4. Biovar characterization

Biovarswere determined using the physiological tests developed by Hayward (1964). Culture mediacontaining 1%filter-sterilized maltose, lactose, cellobiose, mannitol, sorbitol anddulcitolwereprepared and cast in Deepwell plates. Six wells containing different media were sown with suspensions of each isolate. The plates were kept at 28℃ and visual re adings were performed at 3, 5, 7 and 14 days after sowing. Color change of the culture medium from green

to yellow indicated carbohydrate utilization by bacteria and was regarded as positive. No differentiation of biovar 2A and 2B was tested.

2.5. DNA typing

The phylotypes were identified using multiplex PCR (Fegan & Prior, 2005). Four direct primers were used, Nmult 21:1F, Nmult 21:2F, Nmult 22:Inf, and Nmult 23:AF, and only one primer in the reverse direction, Nmult 22:RR. These phylotype-specific primers generate different products for each phylotype allowing strain differentiation (Fegan & Prior, 2005). Identification of phylotypes I, II, III and IV were based on the amplification of the expected 144 bp, 372 bp, 91 bp and 213 bp fragments, respectively. PCR amplifications were done in 25 µl reaction volumes containing 1X PCR buffer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 0.24 µl of each phylotype-specific primer (Nmult:21:1F, Nmult:21:2F, Nmult:22:InF), 0.4 µl Nmult22:RR, 0.72 µL Nmult 23:AF, 2 U of Taq DNA polymerase and 25 ηg of DNA template. Thermocycling conditions included an initial melting phase at 96° C for 5 min followed by 30 cycles of 94° C/15s, 59° C/30s, and 72° C/30s, and a final hold of 72° C for 10 min. All multiplex PCR amplifications were performed in a MJ PTC-100 thermocycler. For each PCR amplification, 5 µl PCR products were subjected to electrophoresis in 1.0% (wt/vol) agarose gels with 0.001 µg/mL GelRed (Biotium) and photographed under UV light. Theamplifiedfragment sizeswere estimatedby comparisonwith a 1kb markerladder when visualized in theultraviolet transilluminator. In each experiment, reference strains for the different phylotypes were used for comparison: GMI1000, UW443, UW386, K60, IPO1609, NCPPB3987, UW129, ANT307, CFBP3059, CFBP734, NCPPB1018, CFBP734, CRM39, R28, UW162, T1UY and *Ralstonia syzygii*.

2.6. Sequevar characterization

The genetic diversity and phylogenetic relationships of strains were characterized by comparative analysis of the partial nucleotide sequences of the *egl* gene. PCR amplification of a 750 bp region of *egl* was performed using the primer pair Endo-F and Endo-R (Fegan & Prior, 2005). The reaction mixture (25 µL, total volume) contained 12.5 µL of the 10X Dream Taq buffers, 4 pmol of each primer, 2 µL of template DNA and water to reach 25 µl total volume.

The solution was vortexed and quickly centrifuged. The thermocycling protocol included an initial denaturation step at 96℃ for 1 min, an amplification phase made of three steps repeated 30 times (a denaturation step at 95℃ for 1 min; an annealing step at around 70℃ for 2 min, an extension step at 72℃ for 2 min) and a final extension step at 72℃ for 10 min in the MJ PTC-100 thermocycler. PCR products were purified using Illustra GFX PCR DNA and the Gel Band Purification (GE Healthcare) kits and sequenced by Macrogen Services (Kumchun-ku, Seoul, Korea) using Endo-F and Endo-R primers (Poussier et al., 2000). The nucleotide sequences were edited with Sequencher 4.6 (Gene Codes Corp., Ann Arbor, Michigan) and aligned using Clustal W in MEGA 5.0 (Tamura et al., 2011) and were deposited in the GenBank database.

Two sets of aligned sequences were studied. The first set, comprised sequence alignments of all isolates obtained in the present study and representative isolates downloaded from GenBank, so as to include at leastonesequence ofeachsequevar of the *R. solanacearum* species complex. Phylogenies were reconstructed using three methods. The neighbor-joining was used with the evolutionary model of Jukes & Cantor (1969) and 1000 bootstrap resampling in MEGA 5.0 (Tamura et al., 2011). We also performed maximum likelihood (ML) in RAxML (Stamatakis, 2006). This analysis was done with the GTR+G model of evolution, based on the Akaike Information Criterion, running 1000 pseudo replicates. Bayesian phylogenetic analysis (MB) used MrBayes v. 3.2.2 (Ronquist & Huelsenbeck, 2003). The HKY+I nucleotide substitution model was selected using MrModeltest v. 2.2 (Nylander, 2004). The Markov Chain Monte Carlo (MCMC) analysis was conducted with four chains starting from a random tree topology and lasted 10 million generations. MB trees were saved at every 100 generations, resulting in 100.000 saved trees. A burn-in of 25% was applied and posterior probabilities (PP's) were calculated.

The second set was constructed after analyzing the first one. From each cluster of similar sequevars, one representative sequence was chosen for maximum parsimony (MP) analysis which was carried out using heuristic searches with 1000 random-addition sequence replicates and tree bisection reconnection (TBR) branch swapping in PAUP* (4.0 beta 10) (Swofford, 2002). Pairwise distance values were calculated using *p-distance* in MEGA. We assumed a sequence divergence values of less equal 1% for determining the

sequevar based on *egl* gene. Representative sequences of all phylotypes were used in the analyses. The ancestral and most divergent phylotype IV was used as outgroup (Wicker et al., 2012).

2.7. Genomic fingerprinting and genetic analysis

Aliquots of genomic DNA were used as templates to generate repetitive element PCR genomic fingerprints with the BOX A1R primer (Louws et al., 1994). PCR products were separated by electrophoresis on 2% agarose gels at 80 V, stained with GelRed (Biotium) and photographed under UV light. Digital images of the stained agarose gels were stored and the banding pattern was scored as presence or absence with the BioNumerics 4.5 software package (Applied Maths, Sint-Martens-Latem, Belgium).

The data set was analyzed to investigate potential subdivision of the population in relation to hosts and geographic regions. Additionally, clonality, genotypic diversity and genotypic differentiation were also quantified. Clonality of the population and subpopulations was estimated based on the frequency of clones in the population adjusting for the differences in sample size and the number of genotypes (Zhang et al.,2002). Genotypic diversity was assessed using the Stoddard and Taylor's G index (Stoddard & Taylor, 1988) using vegan and vegetarian packages of the R program (R Development Core Team 2011). Bootstrap standard error of the estimates was obtained with 1000 replicates. The index was scaled to allow for comparisons among different sample sizes. Rarefied genotypic richness estimates were calculated to study the proportion of genotypes in a population (Grünwald et al.,2003). The percentage of polymorphism and Nei's gene diversity (Nei, 1978) in each population was defined using Popgene 1.32 software (Francis & Yang, 2000).

Genetic differentiation between populations was estimated based on Θ values calculated with Multilocus 1.3 software (Agapow & Burt, 2001) using 1000 randomizations of the data set to generate the null distribution for region, host and phylotype. Linkage disequilibrium was tested with Multilocus to infer the potential occurrence of sexual recombination and to estimate the index of association and multilocus association, Ia and r_D, respectively. The frequency of clones within a population, clonal fraction, was estimated as $(N - G)/G$, where

G is the number of haplotypes within a population and *N* is the total number of isolates.

Populations were grouped using the unweighted pair group method with arithmetic mean (UPGMA) of Nei’s unbiased genetic distances and used to construct a phylogenetic tree with PopGen 1.32 (Yeh & Boyle, 1997). The SEQBOOT, GENDIST, CONSENSE and NEIGHBOR programs of PHYLIP 3.6 were used to calculate the bootstrap values for branches of the trees from 100 replicates (Felsenstein, 1993). Genotype relationships were studied using Nei's genetic distance matrices, UPGMA trees, and 1000 bootstrap replications generated with PAUP*. Analysis of molecular variance (AMOVA) based on Euclidean distances was used to test hypotheses regarding the partitioning of variability for *R. solanacearum* in order to assess the hierarchical population structure by host. All tomato isolates of the South, Southeast and Central regions were pooled and compared with the potato isolates distributed in the same region. The AMOVA was conducted in Arlequin 2.0 (Excoffier et al., 2005).

3. Results

3.1. Species-specific PCR confirmation of *R. solanacearum* species complex

The expected 281 bp amplicon was detected in all the 301 isolates under test, including the reference isolates. All isolates belong to the *R. solanacearum* species complex (Table 1).

3.2. Biovar classification of *R. solanacearum*

Biovar classification revealed bias in the geographic and host distribution of *R. solanacearum.* More than half of the isolates (52%) were assigned to biovar 1, while 32% and 16% of the isolates corresponded to biovars 2 and 3, respectively. Isolates of biovar 1 occurred in all regions. These isolates were more commonly associated with tomato, bell pepper and eucalyptus plants. Contrary to the observed for biovar 1 isolates, those of biovar 2 were commonly found in the South. No biovar 2 isolates were found in the North region. Isolates of biovar 2 were associated with potato, eucalypt and

eggplant. Isolates of biovar 3 also occurred in high frequency in the northern parts of the country, and few isolates in the Central and Southeast regions. No biovar 3 isolates were found in the South region. Biovar 3 was isolated from tomato, eggplant, bell pepper and capsicum. Most isolates (38) were found associated with tomato.

In the South, 39% and 61% of the isolates were of biovar 1 and biovar 2, respectively. In the Southeast region, 59%, 40% and 1% of the isolates were of biovars 1, 2 and 3, respectively. In Central Brazil: 55% were biovar 1, 43% biovar 2 and 2% biovar 3; in the Northeast: 57% biovar 1, 27% biovar 2 and 16% biovar 3; and in the North: 55% biovar 1 and 45% biovar 3. Biovar 3 was not found in the South region, but was highly correlated (100%) with phylotype I (see below). No isolates were assigned to biovars 4 or 5 (Table 2).

3.3. Phylotype classification and distribution

Phylotype-specific multiplex PCR analyses amplified an 114 bp fragment, typical of phylotype I, in 48 isolates and a 372 bp fragment, typical of phylotype II, in 253 isolates. Isolates belonging to phylotype I were found in different locations in at least three different states of the North and Northeast regions, as well as in the Distrito Federal, in the Central region and in Rio de Janeiro state, in the Southeast region. Phylotype I strains were obtained from tomato, eggplant, capsicum and bell pepper (Table 2). Phylotype II is widely distributed in the country and isolates were associated with tomato, potato, eggplant, bell pepper, eucalyptus, capsicum, geranium, banana and scarlet eggplant. In the South and Southeast regions, all isolates were classified as phylotype II. Isolates in the North, Northeast and Central regions belong to phylotypes I and II. Phylotype III and IV were not found.

3.4. Sequevar diversity

Partial sequencing of the endoglucanase (*egl*) gene showed restricted sequevar diversity. PCR amplification of a 750 bp region of the *egl* gene, which is a pathogenicity determinant, produced well resolved phylogenetic trees (Figure 2). Only reconstructed MB trees are presented, but the branching pattern using all methods were similar and isolates were grouped into four phylotype clades with high bootstrap support values. All strains were assigned a

phylogenetic position, which was entirely consistent with their phylotype determination on the basis of multiplex PCR. Phylogenetic analysis differentiated Brazilian isolates into phylotypes I, IIA and IIB. The first reconstructed tree included 301 Brazilian sequences and 52 reference sequences of different sequevars of isolates collected in different countries. In total, 108 variants were found. The variants were recorded as varying from 1 to 108.

Phylotype I was assigned to 48 Brazilian isolates and included 15 variants. All phylotype I isolates (48) were sequevar 18. Phylotype II isolates were divided into two subclusters: IIA and IIB, with 28 and 65 variants, respectively. The phylotype IIA subcluster was composed of isolates that infect different hosts, mainly tomato and eucalyptus, and encompasses sequevars 41 (16 isolates) and 50 (95 isolates). Subcluster IIB was comprised of 142 isolates and sequevars 1, 4, 27 and 28. Subcluster IIB was predominantly made of isolates from the South region and was more commonly associated with potato, eggplant, geranium, banana and bell pepper. The phylotype IIB isolates were split into two groups. One group was made of isolates from bell pepper, tomato, potato, eggplant and eucalyptus; while the other group consisted solely ofisolates frompotato of the South, Southeast and Central regions.

Sequevars 4, 28 and 41 were detected in isolates from different hosts and distinct regions. Sequevar 41 is reported in Brazil for the first time. Sequevar 50 was found in tomato isolates from all regions and sequevar 1 was found in the South, Southeast, Central and Northeast regions in potato isolates. Isolates of banana were sequevar 4. One isolate collected from eucalyptus was characterized as sequevar 27. Phylogenetic analysis showed that four new sequevars were found in phylotype IIB.

3.5. BOX-PCR cluster analysis

The BOX-PCR fingerprints of the 301 isolates produced 35 bands that ranged from 200 to 3500 bp (data not shown). A total of 282 haplotypes were derived from the multilocus pattern. The clonal fraction, the frequency of clones in the dataset, varied from 3% in the Northeast and Southeast regions to 12% in the South region. The average clonal fraction across different host species was 7.1% (with variation of 4.4 - 10%). Five isolates of tomato and six of potato of

the same haplotype were sampled from different regions. The most frequent genotype was observed three times in each state (data not shown).

Similar high levels of genotypic diversity were detected when analyses were partitioned by region or host (Table 3). The highest genotypic diversity was detected in the North region, followed by the South, Central, Southeast and Northeast. Genotypic diversity was highest in tomato and potato isolates, at least 5 times higher than in isolates from eggplant, eucalyptus and bell pepper. Since there was a limited number of isolates from capsicum, geranium, banana and scarlet eggplant, we did not estimate the genotypic diversity among isolates from these hosts. The genotypic richness in the Southeast and South regions was higher than in other regions. The tomato subpopulation had the highest genotypic richness while the bell pepper had the lowest. The highest gene diversity values were estimated for the North (H_E = 0.39, 92.9% polymorphic loci) and Northeast (H_E = 0.37, 91.4% polymorphic loci) regions and in the long pepper (H_E= 0.43, 100% polymorphic loci), eggplant (H_E= 0.42, 100% polymorphic loci) and tomato (H_E= 0.39, 94.6% polymorphic loci) subpopulations. The potato subpopulation from the South region had the lowest gene diversity. Relatively low values of H_E were estimated for phylotype I isolates (H_E = 0.21, 55.7% polymorphic loci) when compared with isolates from other phylotypes (data not shown).

In most tests the random mating hypothesis could not be rejected, thus there is evidence of recombination in the population of *R. solanacearum*. In all populations, except the eucalyptus population ($I_{A=}$ 0.28), the null hypothesis of recombination could not be rejected.

The population was moderately structured according to geographic region and host. The F_{ST} value between North and South populations was 0.19 and those between the Northeast populations and all other regions varied from 0.12 to 0.19. Apparently, the Northeast population differs from other populations and the populations from the Central, Southeast and South regions were more similar among each other since F_{ST} values varied from 0.007 to 0.095. The tomato and potato subpopulations differed significantly ($P<0.05$), however pairwise comparisons between other subpopulations were not significantly different (Table 4). AMOVA results support no geographic division regarding Central, Southeast and South regions. However, there was evidence of

population structuring according to host, especially between tomato and potato subpopulations. Molecular variance was mainly partitioned according to among (55.6%), followed by a similar variation level within subpopulations (44.4%) (Table 5).

Genotypes were sorted into three major and well-supported groups (Fig. 3). The first group included isolates of the Central, Southeast and South regions mainly those of phylotype IIA. The second group was comprised mainly by isolates from the North and Northeast regions and of phylotype IIA. The third group was comprised of phylotype I isolates. Haplotypes of phylotype IIB, mainly from potato in the South, Southeast and Central regions, were assigned to a separate cluster. Few isolates from eggplant and tomato were also clustered, though with little bootstrap support, into the latter group.

Gene flow seems to take place between the Southeast and Central regions, mainly among Minas Gerais, São Paulo, Goiás and Distrito Federal (Fig. 3). Evidence for the occurrence of migration was also detected between all regions of the country (Fig. 3). The UPGMA tree revealed distinct groups such as those formed by isolates of phylotype I and II, and phylotype II isolates from the South, Southeast and Central regions from populations of the Northeast and North regions.

4. Discussion

Brazil, a putative ancient center of diversity of *R. solanacearum*, has an extensive territory, approximately 8.5 million squared-kilometers, in which different ecosystems and a highly diverse community of plants can be found. This allows for the occurrence of different genoytpes of plant-associated organisms, including bacteria. Plants infected with *R. solanacearum* were detected in latitudes in Brazil as extreme as 00°02’20” N to 29°17’29S. Consequently, a diverse population of *R. solanacearum* was expected to be found in Brazil. High gene and genotypic diversity were estimated and two phylotypes of *R. solanacearum*, I and II, were detected. The population is weakly structured and there is evidence that human-mediated transportation of propagative material as well as natural dispersal by rivers can have contributed to spread the pathogen across regions and influence its genetic distribution. There was genetic differentiation between isolates from potato and tomato.

Mutation, recombination and gene flow seem to be the major evolutionary mechanisms that contribute to shape the *R. solanacearum* population.

4.1. Origin and natural history of the pathogen in South America

Worldwide, the different phylotypes of *R. solanacearum* are relatively distributed according to geographic region: phylotype I isolates most likely originated in Asia, phylotype II in America, phylotype III in Africa and phylotype IV in Indonesia, nevertheless more than one phylotype is commonly reported in a country due to effective dispersion processes (Horita et al., 2010; Sanchez-Perez et al., 2008; Xue et al., 2011; Ramsubhag et al., 2012). Most Brazilian isolates of *R. solanacearum* (84%) belong to phylotype II, although phylotype I isolates were also detected. The reasons for the absence of isolates of phylotypes III and IV in Brazil cannot be completely ascertained at the moment. Probably, phylotype III isolates have never been introduced in Brazil due to no transmission by true seeds. On the other hand, the insect-transmitted nature of phylotype IV isolates can impair its distribution due to the lack of vector/host in other countries (Remenant et al., 2011).

The occurrence of phylotype I isolates, which are originated in Asia, in different Brazilian regions can be attributed to human-mediated processes. We speculate that one possible pathway was through contaminated seeds or seedlings of black pepper brought to Brazil by Japanese immigrants in the early 20th Century. Black pepper is native to India and the crop was introduced by Japanese immigrants. It has been documented that 20 seedlings were brought from Singapore in 1933, but only two plants survived the trip (Yamauchi, 1992). The plants promptly adapted to warm regions in Brazil and within a 20 year-period the cultivated area quickly expanded (Yamauchi, 1992). Between 1940-1950, Japanese immigrants established new colonies in the Amazon region (Northern Brazil), Pernambuco state (Northeast), Minas Gerais, São Paulo and Rio de Janeiro states (Southeast), and in Brasília and Mato Grosso do Sul state (Central). In the present study, phylotype I isolates were distributed in different states, but predominate in Tocantins state in the North region. Additionally, two isolates were found in Brasília, in the Central region. Nevertheless, in other studies, phylotype I isolates were reported in Pernambuco, São Paulo, Rio de Janeiro and Amazonas states affecting bell pepper, tomato and long pepper

(Rodrigues et al., 2012; Silveira et al., 1998; Coelho Netto et al., 2004; Garcia et al., 2013).

Phylotype I isolates were identified as belonging to biovar 3 and phylogenetically associated with sequevar 18, which is also found in Guatemala (Sanchez Perez et al., 2008), French Guiana (Milling et al., 2009) and Trinidad (Ramsubhag et al., 2012), countries located from the equator up to 15° N. In Brazil, phylotype I isolates predominantly occurred in regions around the equator down to 15° S. Isolates of biovar 3 were more frequently associated with *Capsicum* species, as compared to tomatoes, as previously reported (Coelho Neto et al., 2004). These are the two most important solanaceous hosts affected by this variant in the northern and northeast Brazil. There are also interesting indication that breeding for resistance to biovar 3 is easier for tomatoes and more difficult to *Capsicum* (Lopes et al., 1993; Lopes & Boiteux, 2004). We speculate that biovar 3/ phylotype I isolates are well adapted to regions with higher temperatures.

The predominance of the American-originated phylotype II isolates in Brazil was expected (Wicker et al., 2012). However, we detected variation in the geographic distribution and host range of phylotype II isolates in the country. Phylotype IIA isolates were present in all regions and were obtained from different hosts, but most came from infected tomato plants cultivated in lowland areas. Isolates of phylotype IIB were pathogenic to potato and tomato, but most came from infected potato plants cultivated in highland areas (altitude ranging from 800m to 1600m). Additionally, phylotype and biovar were associated, as previously recorded (Hayward, 1991; Fegan & Prior, 2005).

Brazilian phylotype II isolates were phylogenetically diverse and included the known sequevars 1, 4, 27, 28, 41 and 51. In addition, four other sequences that did not cluster with any previously known sequevars were also found in phylotype IIB. There was predominant correlation with potato in the South, Southeast and Central regions. Curiously, sequevar 4 causes banana Moko disease in the Northeast and in wild tomato in the Southeast regions of Brazil. Wicker et al. (2007) described phylotype II-sequevar 4 isolates of different hosts that were unable to naturally infect banana (genotyped as phylotype II/4NPB). Similar genotypes were collected in Florida that caused wilting of tomato, anthurium and triploid banana (Norman et al., 1998). In Brazil,

sequevars 1, 4, 28 and 41 were found in several regions. Isolates collected in eucalyptus and potato were classified as sequevar 41, as previously described (Fonseca et al., 2013). The present study is the first formal report of sequevar 50 in Brazil. Using the proposed criterion of Fegan & Prior (2005) and Poussier et al. (2006) we classified four new sequevars in the phylotype IIB: sequevars 53, 54, 55 and 56.

4.2. Processes underlying the structure of the Brazilian population of *R. solanacearum*

Using BOX-PCR it was possible to detect three groups of *R. solanacearum* in Brazil: one formed by phylotype I isolates, a second group formed by isolates predominantly from the South, Southeast and Central regions, and a third group formed by isolates from the North and Northeast regions. We hypothesize that the spatial distribution of the population of *R. solanacearum* is largely affected by ecological differences between regions. The subpopulation of the Northeast region can be clearly differentiated from those of all other geographical regions, but the North. The difference between the Northeast and the non-North subpopulations can be explained by environmental properties such as temperature, humidity, altitude and absence of some host species (Tang, 2009).Further investigation on the influence of climate on *R. solanacearum* populationdiversity is required.

There is evidence of spatial population subdivision of Brazilian phylotype IIA between the North and South regions. Absence of regional structure in the Central, Southeast and South regions and North and Northeast regions suggested no isolation by distance and/or occurrence of gene flow between these large spatial groups. Previous studies used repetitive DNA elements to define groups for *R. solanacearum*.The population from Trinidad was found to be highly polymorphic and showed high gene diversity using BOX-PCR (Cruz et al., 2012; Ramsubhag et al., 2012). Contrary to what we found, no clustering by area was found. This can be explained by the relative close proximity of the sampled areas in Trinidad. In this island the most distant sampled areas were separated by approximately 60 km, whereas in Brazil, the two farthest areas were separated by 3.393 km. Short distances would allow gene flow by means of movement of people, equipments, and plant materials

(Ramsubhag et al., 2012). The decrease in the community similarity with geographic distance is a universal biogeographic pattern observed for all domains of life, mainly organisms with limited dispersal (Martiny et al., 2011). When analyzing the genetic structure at a smaller geographic scale such as the South, Southeast and Central regions and between North and Northeast regions we detected weak population structure. Using BOX-PCR analysis, regional differentiation was also detected when comparing populations of *R. solanacearum* in the Philippines (Ivey et al., 2007). In the Amazon region, no correlation with host and geographic origin was found (Coelho Neto et al., 2004). Additional studies using rep-PCR to characterize the Brazilian isolates may provide higher resolution data that allow the assessment of the relationship of genotypes to location and aggressiveness. This knowledge can be useful for breeding for resistance and better inspection of propagative materials.

Transportation of infected propagative materials as well as natural dispersal by rivers may have contributed to distribute variants of the pathogen across regions and influenced the genetic structure of the population. It was previously suggested that the occurrence of biovar 2 in potato fields in the Northeast and Central regions of Brazil may be associated with transport of latently infected tubers from colder southern regions to warmer northern regions (Santana et al., 2012). We detected biovar 2 potato isolates in the Central and Northeast regions of Brazil and haplotype 72 found in the South was detected in the Northeast regions (in Paraná and Bahia states, respectively). Similarly, haplotype 111 was found in the South and in Central regions (in Paraná and Goiás states, respectively). The lower genetic diversity among potato isolates than among tomato isolates also provides some support to the role played by transportation of infected or infested seed-tubers. The wide distribution can also be associated with fruit dissemination, but despite the fact that *R. solanacearum* can be spread by infected tomato fruits, this does not seem to occur very often (Sanchez Perez et al., 2008). In Europe and in the United States, *R. solanacearum* was detected in river water and the origins of introduction in cultivated fields have been traced to contaminated water used in irrigation (Hong et al., 2004; López & Biosca, 2005). The spatial distribution of sequevars 4, 28 and 41 over distinct regions may allegedly have occurred by water of the affluents of the rivers in distinct regions. Rivers can help disperse *R.*

solanacearum over short and long distances (Elphinstone et al.,1998). Biovar characterization also helped to understand the route of migration of *R. solanacearum* into Brazil.

The Brazilian population of *R. solanacearum* has higher genetic diversity than other organisms thought to reproduce predominantly asexually and with limited dissemination. In all regions and host subpopulations estimates of gene and genotypic diversity were high. High gene diversity provides indirect evidence for the occurrence of recombination (Maynard Smith, 1971). Furthermore, when isolates from all geographic regions were tested for linkage disequilibrium, there was strong evidence of occurrence of random mating and consequently that recombination is involved in bacterial diversity. A number of studies revealed the importance of recombination in *R. solanacearum* (Guidot et al., 2009; Coupat-Goutaland et al., 2011; Ramsubhag et al., 2012; Wicker et al., 2012). In a previous study, the multilocus sequence characterization of a world collection of *R. solanacearum* suggested that recombination is an active evolutionary mechanism that shapes population diversity (Wicker et al., 2012). Identification of recombination events provides insights into epidemiology and disease management and from a biological point of view has consequences for microbial diversification and adaptive evolution. For instance, recombination could facilitate adaptation through the acquisition of new metabolic capabilities, promoting speciation or diversification (Jain et al., 2003). A more thorough investigation using robust methods to identify recombination needs to be conducted with the Brazilian isolates of *R. solanacearum*.

Genetic divergence of potato and tomato occurs between populations that are geographically in close proximity. This process can lead to a pathogen adaptation to host plants (Stukenbrock & McDonald, 2008). The populations of *R. solanacearum* associated with potato and tomato in the Central, Southeast and South regions of Brazil are genetically different, but the highest fraction of molecular variation was attributed to be due to differences within host. Genetic variation is key to evolution. Selection and gene acquisition can be involved with host preference or differential aggressiveness. Host specialization changes may be caused by positively selected genetic variation, though no study on selection has been conducted to assess the influence of the host in the populations of *R. solanacearum*. A number of examples illustrate how bacteria adaptation is

influenced by horizontal gene transfer (de la Cruz & Davies, 2000). A single gene can be responsible by the change in host range with type III protein secretion gene product (*avr*A and *pop*P1) (Poueymiro et al., 2009). The influence of the host in the rhizosphere microbiome and its role in shaping pathogen populations is discussed by Berendsen et al. (2012). A complex study is necessary and other pathogenesis-related genes need to be studied to properly address speciation issues in the population of *R. solanacearum* in Brazil.

The striking genetic diversity of the Brazilian population of *R.* s*olanacearum* raises important issues related to both the applied and basic scenarios. From the applied perspective, the high genetic variability of the population is one of the reasons for the limited success of several management strategies aimed at controlling bacterial wilt in several hosts. It can be anticipated that this constitutes a major challenge to the control of bacterial wilt, particularly for the strategies associated with biological control and the deployment of resistant varieties. Additionally, the high variability associated with the positive selection exerted by the host can contribute to host adaptation and emergence of new ecotypes. The findings of the present study prompt the need for greater caution during the commercialization and use of propagation materials and the quality of irrigation water in clean areas. From the evolutionary biology perspective, the high genetic variability supports its categorization as an ancient population and the claim of South America as a center of origin of the pathogen: The high frequency of genotypes, high genotypic diversity, high gene polymorphism, the occurrence of bacterial wilt in native host plants, and the occurrence of the pathogen in virgin areas, including the Cerrado and the Amazon, suggest that *R. solanacearum* is native to these areas (Takatsu & Lopes, 1997). A more thorough investigation is being conducted to assess the likelihood of South America, Brazil in particular, be one of the centers of origin of *R. solanacearum*.

5. Conclusions

Our study represents the first analysis of the genetic structure of the highly diverse Brazilian *R. solanacearum* populationsthat is capable of separating the pathogen at host and regional scales. Previous studies

estimated genetic diversity by geographic distance, but were restricted to a limited number of isolates (Rodrigues et al., 2012; Cruz et al., 2012). Moreover, the fingerprinting methods that were used have been described as being of low resolution and incapable of reconstructing significant patterns of descent. In contrast, we used methods based on the scanning of entire genomes that are more powerful and appropriate for the comparative analysis of population diversity (Rademaker & De Bruijn, 1997), especially of microbial populations that are clonal. The spatial separation and high variability we detected in the populations we examined provided important information that could explain why a plant variety is only resistant in a given geographic region and susceptible to other bacterial populations. Our results explain why breeding programs for durable resistance appear more efficient at regional level and provide a basis to study environmental and historical constraints on the diversity and evolution of the pathogen. There is however much to learn about microbial diversification and pathogenesis. In future studies, we will use multi-locus sequence typing to study the influence of recombination, explore isolate variability in crop fields, the influence of the environment, and reconstruction of patterns of descent in genealogical analyses of Brazilian isolates of phylotype I and II. In addition, we will identify genomic regions and biological tests that can explain the wide range of plant hosts used by the pathogen.

6. Acknowledgments

This research was supported by CNPq grant 476940/2011-7. We gratefully acknowledge Maurício Rossato (Universidade de Brasília) and Saulo Oliveira (EMBRAPA Mandioca e Fruticultura) for technical support. We thank FAPEMIG for providing the scholarship to TRS and Caitilyn Allen, for providing the phylotype standards.

7. References

Agapow PM, Burt A, 2001. Indices of multilocus linkage disequilibrium.*Molecular Ecology Notes***1**, 101-2.

Berendsen RL, Pieterse CMJ, Bakker, PAHM, 2012. *Trends in Plant Science***17**, 478-86.

Bertolla F, Gijsegem FV, Nesme X, Simonet P, 1997. Conditions for natural transformation of *Ralstonia solanacearum.Applied Environmental Microbiology***63**, 4965-68.

Buddenhagem I, Sequeira L, Kelman A, 1962. Designation of races in *Pseudomonas solanacearum.Phytopathology***52**, 726.

Coelho Netto RA, Pereira BG, Noda H, Boher B, 2004. Murcha bacteriana no estado do Amazonas, Brasil. *Fitopatologia Brasileira* **29**, 21-7.

Costa SB, Ferreira MASV, Lopes CA, 2007. Diversidade patogênica e molecular de *Ralstonia solanacearum* da região amazônica brasileira. *Fitopatologia Brasileira***32**, 285-94.

Coupat-Goutaland B, Bernillon D, Guidot A, Prior P, Nesme X, Bertolla F, 2011. *Ralstonia solanacearum* virulence increased following large interstrain gene transfers by natural transformation. *Molecular Plant-Microbe Interactions***24**, 497-505.

Cruz L, Eloy M, Quirino F, Oliveira H, Tenreiro R, 2012. Molecular epidemiology of *Ralstonia solanacearum* strains from plants and environmental source in Portugal. *European Journal Plant Pathology***133**, 687-706.

de la Cruz F, Davies J, 2000. Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends in Microbiology* **8**, 128–33.

Denny TP, 2006.Plant pathogenic Ralstonia specie. In: Gnanamanickam SS, ed. *Plant-associated Bacteria*. Dordrecht, The Netherlands: Springer Publishing, 573-644.

Elphinstone JG, Stanford HM, Stead DE, 1998.Detection of *Ralstonia solanacearum* in potato tubers, *Solanum dulcamara* and associated irrigation water. In: Prior P, Allen C, Elphinstone JG, eds. *Bacterial Wilt Disease: Molecular and Ecological Aspects*. Springer-Verlag, Berlin, 133-9.

Elphinstone JG, 2005. The current bacterial wilt situation: a global overview. In: C. Allen C, Prior P, Hayward AC, eds. *Bacterial Wilt Disease and the Ralstonia solanacearum Specie Complex.* St. Paul, MN: American Phytopathological Society Press, 9-28.

Excoffier L, Laval G, Schneider S, 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online***1**, 47–50.

Fegan M, Prior P, 2005. How complex is the “*Ralstonia solanacearum* species complex?” In: C. Allen C, Prior P, Hayward AC, eds. *Bacterial Wilt Disease and the Ralstonia solanacearum Specie Complex*. St. Paul, MN: American Phytopathological Society Press, 449-61.

Felsenstein J, 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Seattle, WA: University of Washington.

Francis CY, Yang RC, 2000. Popgene version 1.32. [http//www. http://www.ualberta.ca/~fyeh/popgene.html]. Acessed 17 September 2014.

Garcia AL, Lima WG, Souza EB, Michereff SJ, Mariano RLR, 2013. Characterization of *Ralstoni*a *solanacearum* causing bacterial wilt in bell pepper in the state of Pernambuco, Brazil. *Journal of Plant Pathology* **95**, 237-45.

Guidot A, Prior P, Schoenfeld J, Carrère S, Genin S, Boucher C, 2007. Genomic structure and phylogeny of the plant pathogen *Ralstonia solanacearum* inferred from gene distribution analysis. *Journal of Bacteriology* **189**, 377-87.

Guidot A, Elbaz M, Carrère S, Siri MI, Pianzzola MJ, Prior P, Boucher C, 2009. Specific genes from the potato brown rot strain of *Ralstonia solanacearum* and their potential use for strain detection. *Phytopathology* **99**, 1105-12.

Hayward AC, 1964.Characteristics of *Pseudomonas solanacearum.Journal Applied Microbiology* **27**, 265-77.

Hayward AC, 1991.Biology and epidemiology of bacterial wilt caused by *Pseudomonas solanacearum.Annual Reviews Phytopathology.* **29**, 65-87.

Hayward AC, 1994.The hosts of *Pseudomonas solanacearum*. In: Hayward AC, Hartman GL, eds. *Bacterial Wilt: The disease and its causative agent, Pseudomonas solanacearum.* Wallingford, UK: CAB International, 9-24.

He LY, Sequeira L, Kelman A, 1983.Characteristic of strains of *Pseudomonas solanacearum* from China*. Plant Disease* **67**,1357-61.

Hong JC, Momol, M.T., Jones, J. B., Ji, P., Olson, S. M., Allen, C., Perez, A., Pradhanag, P. and Guven, K. 2008. Detection of *Ralstonia solanacearum* in irrigation ponds and aquatic weeds associated with ponds in North Florida. *Plant Disease* **92**,1674-82.

Peeters N, Guidot A, Vailleau F, Valls M, 2013. *Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era.*Molecular Plant Pathology* **14**, 651-62.

Horita M, Suga Y, Ooshiro A, Tsurchiya, K, 2010. Analysis of genetic and biological characters of Japanese potato strains of *Ralstonia solanacearum*. *Journal of General Plant Pathology* **76**, 196-207.

Jain R, Rivera MC, Moore JE, Lake JA, 2003. Horizontal gene transfer accelerates genome innovation and evolution. *Molecular Biology Evolution* **20**, 1598-1602.

Jukes TH, Cantor CR, 1969. Evolution of protein molecules. In: Munro HN, ed. *Mammalian protein metabolism*. New York, NY: Academic Press, 21-132.

Kelman A, 1954. The relationship of pathogenicity of *Pseudomonas solanacearum* to colony appearance on a tetrazolium medium.*Phytopathology* **44**, 693-5.

Lambert CD, 2002. Agricultural bioterrorism protection act of 2002: Possession, use, and transfer of biological agents and toxins; interim and final rule (7 CFR Part 331). *Fed Regist* **67**, 76908-38.

Lopes CA, Boiteux LS, 2004. Biovar-specific and broad-spectrum sources of resistance to bacterial wilt (*Ralstonia solanacearum*) in capsicum.*Crop Breeding and Applied Biotechnology* **4**, 350-5.

López MM, Biosca EG, 2005. Potato bacterial wilt management new prospects for an old problem. In: C. Allen C, Prior P, Hayward AC, eds. *Bacterial Wilt Disease and the Ralstonia solanacearum Specie Complex*. St. Paul, MN: American Phytopathological Society Press, 205-24.

Louws FJ, Fulbright DW, Stephens CT, de Bruijn FJ, 1994. Specific genomic fingerprints of phytopathogenic Xanthomonas and Pseudomonas pathovars and strains generated with repetitive sequences and PCR. *Applied and Environmental Microbiology* **60**, 2286-95.

Maynard Smith J, 1971. What use is sex? *Journal of Theoretical Biology* **30**,319-35.

Milling A, Meng F, Denny TP, Allen C, 2009. Interactions with hosts at cool temperatures, not cold tolerance, explain the unique epidemiology of *Ralstonia solanacearum* race 3 biovar 2. *Phytopathology* **99,** 1127-34.

Nei M, 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals.*Genetics* **89,** 583-90.

Nylander JAA, 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Opina N, Tavner F, Holloway G, Wang J-F, Li T-H, Maghirang R, Fegan M, Hayward AC, Krishnapillai V, Hong WF, Holloway BW, Timmis JN, 1997. A novel method for development of species and strain-specific DNA probes and PCR primers for identifying *Burkholderia solanacearum* (formerly *Pseudomonas solanacearum*).*Asia Pacific Journal of Molecular Biology and Biotechnology* **5,** 19–33.

Pinheiro CR, Amorim JAE, Diniz LEC, Silva AMF, Talamini V, Souza Júnior MT, 2011. Diversidade genética de isolados de *Ralstonia solanacearum* e caracterização molecular quanto a filotipos e sequevares. *Pesquisa Agropecuaria Brasileira* **46,** 593-602.

Poussier S, Prior P, Luisetti J, Hayward C, Fegan M, 2000. Partial sequencing of the *hrpB* and endoglucanase genes confirms and expands the known diversity within the *Ralstonia solanacearum* species complex. *Systematic and Applied Microbiology* **23**, 479-86.

Rademaker JLW, de Brujin FJ, 1997. Characterization and classification of microbes by rep-PCR genomic fingerprinting and computer assisted pattern analysis. In: Caetano-Anolles G, Gresshoff, eds. *DNA Markers: Protocols, Applications and Overviews*. New York, NY: John Wiley and Sons Inc, 151-71.

Ramsubhag A, Lawrence D, Cassie D, Fraser R, Umaharan P, Prior P, Wicker E, 2012. Wide genetic diversity of *Ralstonia solanacearum* strains affecting tomato in Trinidad, West Indies. *Plant Pathology* **61**, 844-57.

Rodrigues LMR, Destéfano SAL, Silva MJ, Costa GGL, Maringoni AC. 2012. Characterization of *Ralstonia solanacearum* from Brazil using molecular methods and pathogenity tests. *Journal of Plant Pathology* **94**, 505-16.

Ronquist F, Huelsenbeck JP, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19,** 1572-74.

Sanchez Perez A, Meija L, Fegan M, Allen C, 2008. Diversity and distribution of *Ralstonia solanacearum* strains in Guatemala and rare occurrence of tomato fruit infection. *Plant Pathology* **57,** 320-31.

Santana BG, Lopes CA, Alvarez E, Barreto CC, Allen C, Quirino BF, 2012. Diversity of Brazilian biovar 2 strains of *Ralstonia solanacearum*. Journal of General Plant Pathology **78,** 190-200.

Silveira, EB, Gomes AMA, Michereff SJ, Mariano RLR, 1998. Variability of *Ralstonia solanacearum* populations causing wilt on tomato in agreste of Pernambuco, Brazil. *Bacterial Wilt Newsletter* **15**, 8-10.

Smith JJ, Offord LC, Holderness M, Saddler GS, 1995. Genetic diversity of *Burkholderia solanacearum* (synonym *Pseudomonas solanacearum*) race 3 in Kenya. *Applied and Environmental Microbiology* **61,** 4263-8.

Stamatakis A, 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22,** 2688-90.

Stoddart JA, Taylor JF, 1988. Genotypic diversity: estimation and prediction in samples. *Genetics* **118,** 705-11.

Strange RN, Scott PR, 2005. Plant Disease: A Threat to Global Food Security. *Annual Review of Phytopathology* **43,** 83-116.

Stukenbrock EH, McDonald BA, 2008.The origins of plant pathogens in agro-ecosystems. *Annual Review Phytopathology* **46**, 75-100.

Swofford DL, 2002. Paup*: Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates. Sunderland, MA

Tamura K, Dudley J, Nei M, Kumar S, 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24,** 1596-9.

Yabuuchi E, Kosako Y, Yano I, Hotta H, Nishiuchi Y, 1995. Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. nov.: Proposal of *Ralstonia picketii* (Ralston, Palleroni and Doudoroff 1973) comb. nov. *Ralstonia solanacearum* (Smith 1896) comb.nov. and*Ralstonia eutropha* (Davis 1969) comb. nov. *Microbiology and Immunology* **39**, 897-904.

Yamauchi A, 1992. Uma epopéia moderna: 80 anos da imigração japonesa no Brasil. São Paulo, SP: Hucitec.

Yao J, Allen C, 2006. Chemotaxis is required for virulence and competitive fitness in the bacterial wilt pathogen *Ralstonia solanacearum*. *Journal of Bacteriology* **188**, 3697-708.

Yeh FC, Boyle TJB, 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits.*Belgian Journal of Botany* **129,** 157.

Wicker E, Grassart L, Coranson-Beaudu R, Mian D, Guilbaud C, Fegan M, Prior, 2007. *Ralstonia solanacearum* strain from Martinique (French West Indies) exhibiting a new pathogenic potential. *Applied and Environmental Microbiology* **73,** 6790-801.

Wicker E, Lefeuvre P, de Cambiaire JC, Lemaire C, Poussier S, Prior P, 2012. Contrasting recombination patterns and demographic histories of the plant pathogen *Ralstonia solanacearum* inferred from MLSA. *The ISME Journal* **6,** 961-74.

Wilson DJ, Falush D, McVean G, 2005.Germs, genomes and genealogies.*Trends in Ecology and Evolution* 20, 39-45.

Xue QY, Yin YN, Yang W, Heuer H, Prior P, Guo JH, Smalla K, 2011. Genetic diversity of *Ralstonia solanacearum* strains from China assessed by PCR-based fingerprints to unravel host plant- and site-dependent distribution patterns. *FEMS Microbiology Ecology* **75,** 507-19.

Zhan J, Kema GHJ, Waalwijk C, McDonald BA, 2002. Distribution of mating types alleles in the wheat pathogen *Mycosphaerella graminicola* over spatial scales from lesions to continents. *Fungal Genetics and Biology* **36,** 128-36.

FIGURES AND TABLES

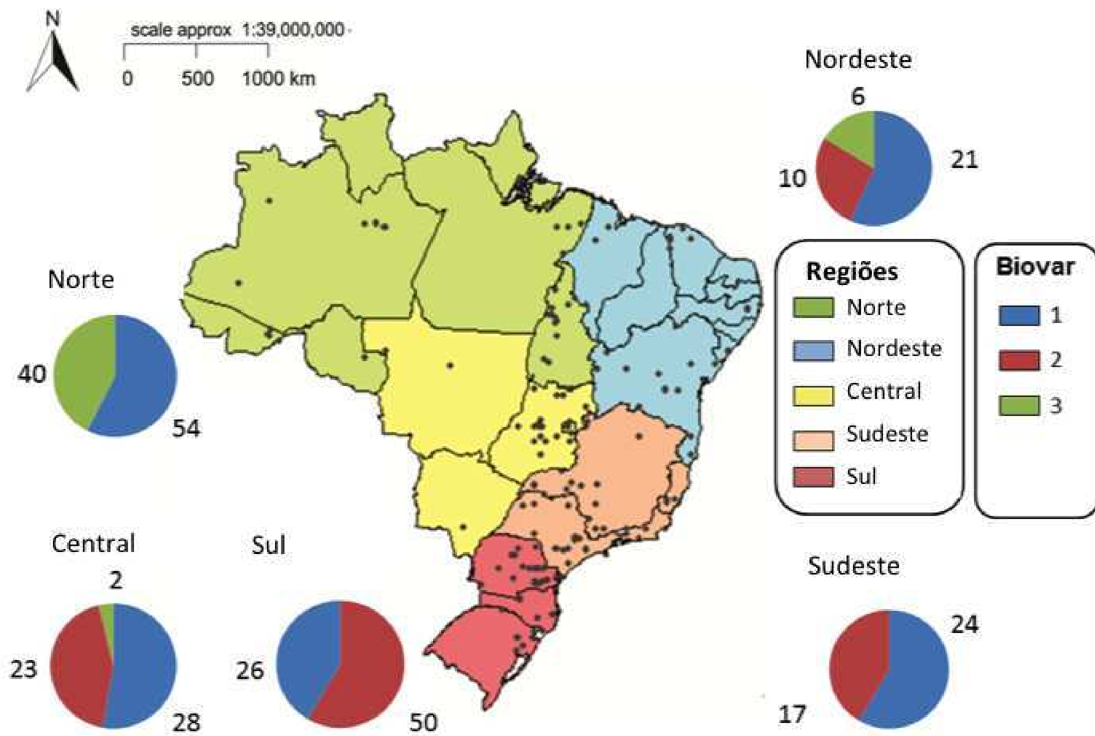
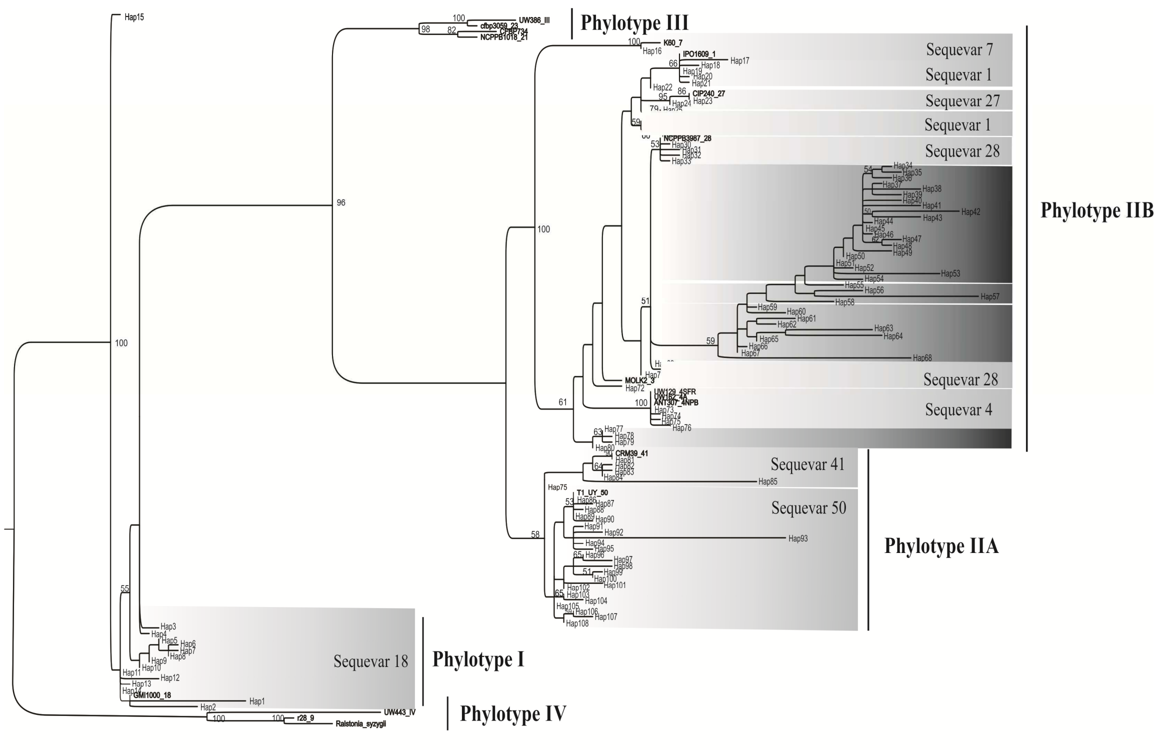
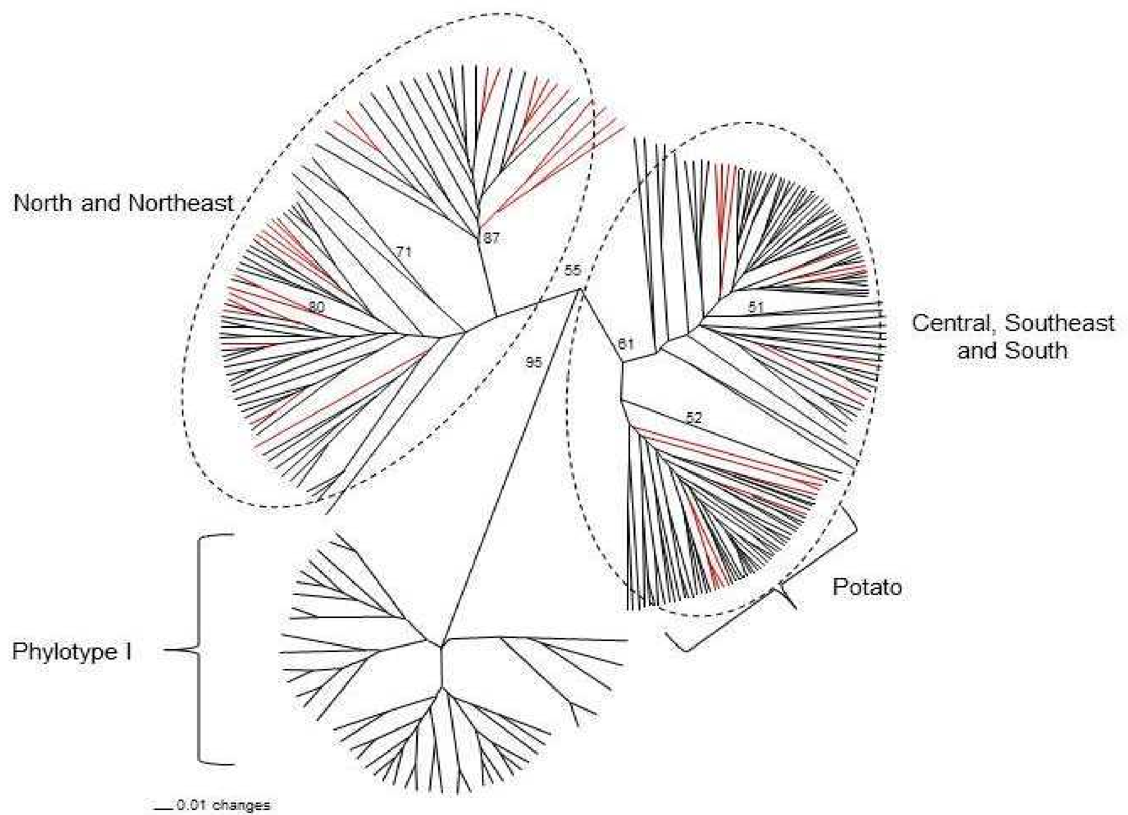


Figure 1. Map of Brazil showing the origin isolates and biovar distribution sampled in different regions of the country.



Bootstrap values >50% derived of 1000 replicates are reported at the nodes. The scale bar shows genetic distance; the bar represents 11 nucleotide differences.

Figure 2. Phylogenetic tree reconstructed using *egl* sequences with the PAUP program.



Bootstrap support values above 50% are shown. Migrations between the groups are indicated with red branches.

Figure 3. Dendrogram of *R. solanacearum* generated using the UPGMA clustering method and Nei's genetic distances.

Table 1. Brazilian and reference isolates of *R. solanacearum* used in phylogenetic analyses

Isolate	Host	Region	Biovar	Phylotype	Sequevar	GenBank Acession
RS01	Potato	South	1	IIA	Hap107/50	
RS02	Potato	Central	2	IIB	Hap41/54	
RS03	Potato	Central	1	IIB	Hap43/54	
RS04	Potato	Southeast	1	IIA	Hap105/50	
RS05	Potato	Southeast	1	IIA	Hap106/50	
RS06	Potato	South	1	IIA	Hap105/50	
RS07	Long pepper	North	3	I	Hap9/18	
RS08	Tomato	North	1	IIA	Hap81/41	
RS09	Potato	Southeast	1	IIA	Hap105/50	
RS12	Potato	Southeast	1	IIB	Hap61/56	
RS15	Eggplant	North	3	I	Hap6/18	
RS17	Long pepper	North	1	IIB	Hap62/56	
RS18	Tomato	North	3	I	Hap10/18	
RS21	Tomato	Northeast	1	IIA	Hap85/41	
RS23	Potato	Northeast	2	IIB	Hap62/56	
RS25	Potato	South	2	IIB	Hap63/56	
RS26	Potato	South	2	IIB	Hap63/56	
RS27	Potato	South	2	IIB	Hap64/56	
RS28	Potato	South	2	IIB	Hap65/56	
RS29	Potato	South	2	IIB	Hap63/56	
RS30	Potato	South	2	IIB	Hap66/56	
RS31	Tomato	North	1	IIB	Hap68/56	
RS33	Tomato	North	3	I	Hap8/18	
RS34	Tomato	North	3	I	Hap8/18	
RS35	Tomato	North	1	IIA	Hap96/50	
RS37	Tomato	North	1	IIA	Hap83/41	
RS41	Tomato	North	1	IIA	Hap96/50	
RS42	Tomato	North	1	I	Hap6/18	
RS43	Tomato	North	1	IIA	Hap107/50	
RS44	Potato	Central	1	IIB	Hap30/28	
RS53	Eggplant	Northeast	3	I	Hap1/18	
RS54	Eggplant	Central	1	IIA	Hap90/50	
RS56	Eggplant	Northeast	3	I	Hap1/18	
RS59	Potato	South	2	IIB	Hap33/28	
RS60	Potato	South	2	IIB	Hap32/28	
RS61	Potato	South	2	IIB	Hap31/28	
RS62	Tomato	Northeast	1	IIA	Hap84/41	
RS63	Tomato	Northeast	1	IIA	Hap84/41	
RS64	Tomato	Northeast	3	IIA	Hap98/50	
RS65	Potato	Northeast	2	IIB	Hap18/1	
RS66	Potato	Northeast	2	IIB	Hap18/1	
RS67	Potato	Northeast	2	IIB	Hap19/1	
RS68	Potato	Northeast	2	IIB	Hap22/1	
RS69	Potato	Northeast	2	IIB	Hap22/1	
RS70	Potato	Northeast	1	IIA	Hap99/50	
RS71	Eggplant	Central	2	IIB	Hap32/28	
RS72	Potato	Central	1	IIA	Hap98/50	
RS77	Eggplant	Central	2	IIB	Hap70/28	
RS81	Eggplant	Central	2	IIB	Hap70/28	
RS82	Eggplant	Central	2	IIB	Hap69/28	
RS83	Eggplant	Central	2	IIB	Hap71/28	
RS84	Eggplant	Central	2	IIB	Hap72/28	
RS86	Eggplant	Central	2	IIA	Hap98/50	
RS87	Eggplant	Central	2	IIA	Hap98/50	
RS88	Eggplant	Central	2	IIB	Hap30/28	
RS89	Eggplant	Central	2	IIB	Hap33/28	
RS90	Potato	Central	2	IIB	Hap58/55	
RS91	Potato	Central	2	IIB	Hap33/28	

Continuing...

Table 1.Continued

Isolate	Host	Region	Biovar	Phylotype	Sequevar	GenBank Acession
RS92	Potato	Central	2	IIB	Hap80/57	
RS93	Potato	South	1	IIA	Hap81/41	
RS94	Potato	South	2	IIB	Hap37/54	
RS95	Potato	South	2	IIB	Hap34/54	
RS96	Potato	South	2	IIA	Hap92/50	
RS97	Potato	South	2	IIB	Hap69/28	
RS98	Potato	South	2	IIB	Hap48/54	
RS99	Potato	South	2	IIB	Hap49/54	
RS100	Tomato	Central	1	IIA	Hap84/41	
RS101	Tomato	Central	1	IIA	Hap84/41	
RS102	Tomato	Central	1	IIA	Hap84/41	
RS103	Tomato	Central	1	IIA	Hap84/41	
RS107	Potato	South	2	IIB	Hap35/54	
RS111	Potato	South	2	IIB	Hap53/54	
RS113	Potato	South	2	IIB	Hap54/54	
RS115	Potato	South	2	IIB	Hap46/54	
RS116	Potato	Central	1	IIA	Hap86/50	
RS118	Potato	Central	1	IIA	Hap87/50	
RS119	Potato	Central	1	IIA	Hap87/50	
RS120	Potato	Southeast	2	IIB	Hap66/56	
RS121	Tomato	Central	1	IIB	Hap67/56	
RS122	Tomato	Central	1	IIA	Hap81/41	
RS127	Potato	Central	1	IIB	Hap60/56	
RS130	Potato	South	2	IIB	Hap68/56	
S131	Potato	South	2	IIB	Hap28/1	
RS132	Potato	South	2	IIB	Hap28/1	
RS134	Potato	South	2	IIB	Hap21/1	
RS139	Potato	South	2	IIB	Hap29/1	
RS140	Potato	South	2	IIB	Hap29/1	
RS141	Potato	South	1	IIA	Hap89/50	
RS142	Potato	Southeast	2	IIB	Hap63/56	
RS145	Potato	Southeast	2	IIB	Hap63/56	
RS146	Long pepper	North	1	IIB	Hap73/4	
RS147	Long pepper	North	1	IIB	Hap73/4	
RS153	Potato	South	2	IIB	Hap66/56	
RS154	Potato	South	2	IIB	Hap66/56	
RS156	Potato	Central	2	IIB	Hap43/54	
RS157	Potato	Central	2	IIB	Hap51/54	
RS158	Tomato	Southeast	1	IIA	Hap93/50	
RS160	Potato	South	2	IIB	Hap77/57	
RS161	Potato	South	2	IIB	Hap77/57	
RS164	Potato	South	2	IIB	Hap77/57	
RS165	Potato	South	2	IIA	Hap97/50	
RS167	Tomato	North	1	IIA	Hap100/50	
RS168	Tomato	North	1	IIA	Hap100/50	
RS170	Tomato	North	1	IIA	Hap100/50	
RS171	Tomato	North	1	IIA	Hap100/50	
RS172	Tomato	Central	1	IIA	Hap108/50	
RS175	Potato	Southeast	1	IIB	Hap66/56	
RS177	Long pepper	North	1	IIB	Hap74/4	
RS178	Long pepper	North	1	IIB	Hap74/4	
RS179	Long pepper	North	1	IIB	Hap76/4	
RS183	Tomato	Northeast	1	IIA	Hap95/50	
RS184	Tomato	Northeast	1	IIA	Hap95/50	
RS185	Tomato	Northeast	1	IIA	Hap95/50	
RS186	Tomato	Northeast	1	IIA	Hap95/50	
RS187	Tomato	Northeast	1	IIA	Hap95/50	
RS191	Long pepper	North	1	IIB	Hap75/4	
RS194	Long pepper	North	1	IIA	Hap81/41	

Continuing...

Table 1.Continued

Isolate	Host	Region	Biovar	Phylotype	Sequevar	GenBank Acession
RS195	Tomato	North	1	IIA	Hap103/50	
RS197	Eggplant	North	1	IIA	Hap82/41	
RS198	Long pepper	Central	3	I	Hap4/18	
RS199	Capsicum	Central	1	IIA	Hap81/41	
RS201	Capsicum	Central	2	IIB	Hap42/54	
RS202	Potato	South	2	IIB	Hap38/54	
RS203	Potato	South	2	IIB	Hap38/54	
RS204	Potato	South	2	IIB	Hap39/54	
RS205	Potato	Central	1	IIA	Hap94/50	
RS206	Potato	Central	1	IIB	Hap69/28	
RS207	Tomato	Central	1	IIB	Hap33/28	
RS208	Tomato	North	1	IIA	Hap91/50	
RS209	Tomato	North	1	IIA	Hap91/50	
RS210	Banana	North	1	IIB	Hap73/4	
RS212	Potato	Northeast	2	IIB	Hap20/1	
RS213	Potato	Northeast	2	IIB	Hap20/1	
RS214	Banana	North	1	IIB	Hap73/4	
RS215	Tomato	North	1	IIA	Hap81/41	
RS216	Tomato	North	1	IIA	Hap81/41	
RS221	Tomato	Central	1	IIA	Hap90/50	
RS225	Long pepper	Central	1	IIA	Hap104/50	
RS229	Long pepper	Central	1	IIA	Hap88/50	
RS230	Potato	Southeast	2	IIB	Hap55/55	
RS231	Potato	Southeast	2	IIB	Hap57/55	
RS232	Potato	Southeast	2	IIB	Hap17/1	
RS233	Eggplant	North	1	IIA	Hap81/41	
RS234	Eggplant	North	1	IIA	Hap81/41	
RS235	Tomato	North	1	IIA	Hap81/41	
RS236	Tomato	North	1	IIA	Hap81/41	
RS237	Potato	South	2	IIB	Hap36/54	
RS238	Potato	South	2	IIB	Hap52/54	
RS239	Potato	South	2	IIB	Hap50/54	
RS240	Potato	Central	2	IIB	Hap78/57	
RS241	Potato	Southeast	2	IIB	Hap44/54	
RS242	Potato	Southeast	2	IIB	Hap45/54	
RS243	Potato	Southeast	1	IIA	Hap101/NS	
RS244	Potato	Southeast	2	IIB	Hap40/54	
RS245	Potato	Central	2	IIB	Hap79/57	
RS246	Potato	Southeast	2	IIB	Hap46/54	
RS247	Potato	Southeast	2	IIB	Hap47/54	
RS248	Tomato	South	2	IIB	Hap22/1	
RS249	Tomato	South	2	IIB	Hap22/1	
RS250	Potato	South	2	IIB	Hap27/1	
RS251	Potato	South	1	IIA	Hap102/50	
RS252	Potato	South	2	IIB	Hap45/54	
RS253	Potato	South	2	IIB	Hap46/54	
RS254	Potato	South	2	IIB	Hap25/1	
RS255	Potato	South	2	IIB	Hap26/1	
RS256	Tomato	South	1	IIA	Hap107/50	
RS257	Tomato	South	1	IIA	Hap107/50	
RS258	Tomato	South	1	IIA	Hap89/50	
RS260	Potato	Northeast	2	IIB	Hap78/57	
RS261	Potato	Northeast	2	IIB	Hap78/57	
RS264	Eucalypt	Northeast	1	IIA	Hap95/50	
RS265	Eucalypt	Northeast	1	IIA	Hap95/50	
RS266	Eucalypt	Northeast	1	IIA	Hap95/50	
RS267	Eucalypt	Northeast	1	IIA	Hap83/41	
RS274	Tomato	South	1	IIA	Hap104/50	
RS277	Tomato	South	1	IIB	Hap56/55	

Continuing...

Table 1. Continued

Isolate	Host	Region	Biovar	Phylotype	Sequevar	GenBank Acession
RS278	Tomato	South	1	IIB	Hap56/55	
RS279	Tomato	South	1	IIB	Hap56/55	
RS280	Tomato	South	1	IIB	Hap34/54	
RS282	Tomato	North	1	IIB	Hap38/54	
RS283	Tomato	North	3	I	Hap2/18	
RS284	Tomato	North	3	I	Hap2/18	
RS285	Tomato	Northeast	1	IIA	Hap105/50	
RS286	Tomato	Northeast	1	IIA	Hap106/50	
RS287	Tomato	Northeast	1	IIA	Hap105/50	
RS288	Tomato	Northeast	1	IIA	Hap105/50	
RS289	Tomato	Northeast	1	IIA	Hap107/50	
RS290	Tomato	Northeast	1	IIA	Hap102/50	
RS293	Eucalypt	North	1	IIB	Hap23/27	
RS294	Eucalypt	North	1	IIB	Hap23/27	
RS295	Eucalypt	North	1	IIB	Hap23/27	
RS296	Eucalypt	North	1	IIB	Hap24/27	
RS297	Eucalypt	Southeast	1	IIA	Hap95/50	
RS298	Eucalypt	Southeast	1	IIA	Hap103/50	
RS300	Tomato	Southeast	1	IIB	Hap37/54	
RS301	Tomato	South	1	IIB	Hap37/54	
RS302	Tomato	South	1	IIB	Hap37/54	
RS308	Tomato	North	1	IIA	Hap92/50	
RS309	Tomato	North	3	I	Hap2/18	
RS310	Tomato	North	3	I	Hap2/18	
RS311	Tomato	North	3	I	Hap2/18	
RS312	Tomato	North	3	I	Hap2/18	
RS313	Tomato	North	3	I	Hap2/18	
RS314	Tomato	North	3	I	Hap2/18	
RS315	Tomato	North	3	I	Hap5/18	
RS316	Tomato	North	3	I	Hap518	
RS317	Tomato	North	3	I	Hap5/18	
RS318	Tomato	North	1	IIA	Hap88/50	
RS319	Tomato	North	3	I	Hap3/18	
RS320	Tomato	North	3	I	Hap3/18	
RS321	Tomato	North	3	I	Hap3/18	
RS322	Tomato	North	3	I	Hap8/18	
RS323	Tomato	North	3	I	Hap8/18	
RS324	Tomato	North	3	I	Hap8/18	
RS325	Tomato	North	3	I	Hap8/18	
RS326	Tomato	North	3	I	Hap8/18	
RS327	Tomato	North	3	I	Hap8/18	
RS329	Scarlet eggplant	North	1	IIA	Hap97/50	
RS330	Scarlet eggplant	North	1	IIA	Hap97/NS	
RS331	Scarlet eggplant	North	1	IIA	Hap97/NS	
RS333	Tomato	North	1	IIA	Hap96/50	
RS334	Tomato	North	3	I	Hap7/18	
RS335	Tomato	North	3	I	Hap7/18	
RS340	Tomato	North	1	IIA	Hap92/50	
RS341	Tomato	North	1	IIA	Hap92/50	
RS342	Tomato	North	1	IIA	Hap97/50	
RS343	Tomato	North	1	IIA	Hap92/50	
RS344	Tomato	North	3	I	Hap7/18	
RS345	Tomato	North	3	I	Hap12/18	
RS347	Tomato	North	3	I	Hap12/18	
RS348	Tomato	North	3	I	Hap12/18	
RS350	Tomato	North	1	IIA	Hap97/50	
RS358	Eggplant	North	3	I	Hap8/18	
RS359	Eggplant	North	3	I	Hap8/18	
RS363	Tomato	Southeast	1	IIA	Hap105/50	

Continuing...

Table 1.Continued

Isolate	Host	Region	Biovar	Phylotype	Sequevar	GenBank Acession
RS374	Tomato	Central	1	IIA	Hap101/50	
RS380	Capsicum	Northeast	3	I	Hap15/NS	
RS381	Tomato	Southeast	1	I	Hap15/NS	
RS382	Tomato	Southeast	1	IIA	Hap88/50	
RS385	Tomato	Northeast	1	IIA	Hap95/50	
RS386	Tomato	Southeast	1	IIA	Hap97/50	
RS387	Potato	Southeast	1	IIB	Hap16/7	
RS388	Long pepper	North	3	I	Hap5/18	
RS389	Long pepper	North	3	I	Hap5/18	
RS391	Tomato	North	3	I	Hap11/18	
RS392	Tomato	North	3	I	Hap11/18	
RS394	Tomato	Southeast	1	IIA	Hap91//50	
RS403	Tomato	North	3	I	Hap11/18	
RS404	Tomato	North	3	I	Hap14/18	
RS406	Tomato	South	1	IIA	Hap96/50	
RS407	Tomato	South	1	IIA	Hap96/50	
RS408	Tomato	South	1	IIB	Hap66/56	
RS409	Tomato	South	1	IIB	Hap66/56	
RS410	Long pepper	North	1	IIB	Hap76/4	
RS411	Long pepper	North	1	IIB	Hap76/4	
RS412	Long pepper	North	1	IIB	Hap74/4	
RS413	Long pepper	North	1	IIB	Hap76/4	
RS414	Long pepper	North	1	IIB	Hap76/4	
RS415	Long pepper	North	1	IIB	Hap76/4	
RS416	Tomato	South	1	IIB	Hap37/54	
RS417	Tomato	South	1	IIB	Hap37/54	
RS418	Tomato	South	1	IIB	Hap37/54	
RS419	Tomato	South	1	IIB	Hap37/54	
RS422	Tomato	South	1	IIB	Hap39/54	
RS423	Potato	South	2	IIB	Hap17/1	
RS424	Tomato	North	3	I	Hap13/18	
RS425	Tomato	North	3	I	Hap14/18	
RS430	Tomato	South	1	IIA	Hap104/50	
RS431	Tomato	South	1	IIB	Hap34/54	
RS432	Tomato	South	1	IIB	Hap34/54	
RS433	Tomato	South	1	IIB	Hap37/54	
RS443	Tomato	Central	1	IIA	Hap87/50	
RS444	Tomato	Central	1	IIA	Hap87/50	
RS445	Tomato	Central	1	IIA	Hap87/50	
RS446	Tomato	Central	1	IIA	Hap87/50	
RS448	Capsicum	Southeast	1	IIB	Hap35/54	
RS450	Potato	South	1	IIB	Hap46/54	
RS451	Tomato	Central	1	IIB	Hap35/54	
RS452	Tomato	Southeast	1	IIA	Hap90/50	
RS453	Tomato	Southeast	1	IIA	Hap89/50	
RS454	Tomato	Southeast	1	IIA	Hap89/50	
RS455	Tomato	Southeast	1	IIA	Hap89/50	
RS456	Tomato	Southeast	1	IIA	Hap89/50	
RS457	Tomato	Southeast	1	IIA	Hap89/50	
RS458	Tomato	Southeast	1	IIA	Hap89/50	
RS460	Tomato	South	1	IIB	Hap55/55	
RS461	Potato	Northeast	1	IIB	Hap57/56	
RS463	Potato	Southeast	2	IIA	Hap87/50	
RS465	Potato	Southeast	2	IIA	Hap88/50	
RS469	Tomato	Southeast	1	IIB	Hap37/54	
RS471	Tomato	Northeast	3	I	Hap6/18	
RS472	Tomato	Northeast	3	I	Hap6/18	
RS478	Geranium	Southeast	2	IIB	Hap38/54	
RS479	Geranium	Southeast	2	IIB	Hap38/54	

Continuing...

Table 1.Continued

Isolate	Host	Region	Biovar	Phylotype	Sequevar	GenBank Acession
RS480	Geranium	Southeast	2	IIB	Hap54/56	
RS481	Eucalypt	Central	2	IIB	Hap 59/56	
RS482	Eucalypt	Central	2	IIB	Hap 59/56	
RS483	Eucalypt	Central	1	IIA	Hap85/41	
RS484	Eucalypt	Central	2	IIA	Hap93/50	
RS485	Eucalypt	Central	2	IIA	Hap93/50	
RS486	Tomato	South	2	IIB	Hap 93/50	
GMI1000	Tomato	Fr Gwyana	3	I	18	
UW443	Banana	Indonesia	BDB	IV	10	
UW386	Tomato	Nigeria		III		
ICMP7963	Potato	Kenia		II	7	
K60	Tomato	N Carolina	1	II	7	
IPO1609	Potato	Netherlands		II	1	EF371814
JT516	Potato	Reunion		II	1	
UW163	Banana	Peru		II	4	
NCPPB3987	Potato	Brazil		II	28	AF295261
CRM39	Tomato	Cameroon		II	41	EF439726
CIP240	Potato	Brazil		II	27	
T1UY	Tomato	Uruguay		II	50	
JT525	Pelargonium	Reunion		III	19	
R. syzygii	Tomato	Indonesia		IV	10	

Table 2. Geographic and host distribution of biovar, phylotype and sequevar types of Brazilian *R. solanaceaum* isolates

(A) Geographic distribution

Origin	Biovar			Phylotype			Sequevar
	1	2	3	1	2A	2B	
North	52	0	42	42	31	21	4,28,50
Northeast	21	10	6	6	19	12	1,41
Central	29	23	1	1	26	26	1,4,28,41
Southeast	24	16	1	1	20	20	1,27
South	30	46	0	0	14	62	1,28,50

(B) Host distribution

Origin	Biovar			Phylotype			Sequevar
	1	2	3	1	2A	2B	
Tomato	96	3	38	38	74	25	4,18,28,41,50
Potato	24	74	0	0	17	81	1,4,28
Eggplant	5	9	5	5	7	10	4,18,28,41
Long pepper	16	0	4	4	4	12	4,18,41
Eucalypt	11	4	0	0	11	4	27,41
Capsicum	1	1	2	2	0	2	4
Geranium	0	3	0	0	0	3	-
Banana	2	0	0	0	0	2	4
Scarlet eggplant	3	0	0	0	3	0	-

Table 3. Genetic diversity indices and linkage disequilibrium estimates for region and host subpopulations of Brazilian *R. solanacearum* isolates

(A) Region

Population	North	Northeast	Central	Southeast	South
N	94	37	53	41	76
Nº haplotype	89	36	50	40	67
G	2.34 (1.40 – 5.32)	0.97 (0.13 – 4.05)	1.33 (0.58 – 4.5)	1.08 (0.27 – 4.19)	1.76 (1.15 – 5.06)
E(g15)	36.23	36	35.55	36.18	36.06
H_E	0.39	0.37	0.36	0.37	0.35
% polym	92.86	91.43	80	86	77
Ia	0.08	-0.14	-0.101	-0.067	-0.186
Rd	0.031	-0.006	-0.004	-0.002	-0.007

(B) Host

Population	Tomato	Potato	Eggplant	Eucalyptus	Long pepper
N	137	98	19	15	20
Nº haplotype	131	89	18	14	18
G	7.1 (2.03 – 8.96)	5.17 (1.99 – 5.91)	1.19 (-0.49 – 3.47)	0.94 (-0.64 – 3.28)	1.27 (-0.38 – 3.54)
E(g15)	14.9	14.69	14.38	14	13.89
H_E	0.39	0.33	0.42	0.37	0.43
% polym	94.57	82.57	100	94	100
Ia	0.036	0.011	-0.281	0.272	0.073
Rd	0.001	0.0003	-0.01	0.013	0.057

N, sample size; G, Genotypic diversity - Scaled Stoddart and Taylor index and confidence intervals (in parentheses); E(gn), Genotypic richness estimated the smallest N; H_E, Gene diversity within populations calculated based on clone-corrected data; Ia and rd; index and multilocus associated to estimate of linkage disequilibrium.
* p-value significant at $P<0.05$.

Table 4. Genetic differentiation between for region and host populations of Brazilian *R. solanacearum*[a] and number of migrants (Nm)[b]

(A) Region

Population	North	Northeast	Central	South	Southeast
North		5.584[b]	3.27	1.1	3.4
Northeast	0.041[a]		1.06	1.8	1.8
Central	0.071	0.19*		2.38	35.46
South	0.185*	0.12*	0.095		2.87
Southeast	0.067	0.12*	0.007	0.08	

(B) Host

Population	Tomato	Potato	Eggplant	Eucalyptus
Tomato				
Potato	0.091*			
Eggplant	0.017	0.024		
Eucalyptus	0.007	0.037	-0.02	
Bell pepper	-0.004	0.014	0.012	0.006

* Tetha value was significant at $P<0.05$.

Table 5. Analysis of molecular variance (AMOVA) of *R. solanacearum* tomato and potato populations in the Central, Southeast and South regions of Brazil for 100 individuals using BOX-PCR

Source of variation	d.f.	Sum of squared deviations	Variance components	Proportion of variance components (%)
Among population	1	53.5	0.8	55.6
Among individuals within population	98	54.6	0.6	44.4
Total	99	108.0	1.4	

CHAPTER 2

PARAMETRIZATION OF EVOLUTIONARY MECHANISMS THAT AFFECT THE POPULATION OF *Ralstonia solanacearum* IN BRAZIL USING GENE GENEALOGIES AND THE COALESCENT APPROACH

T. R. Santiago[5], C. A. Lopes[6], G. Caetano-Anollés[7], and E. S. G. Mizubuti[1,8]

Abstract: Parametrization of evolutionary processes that affect populations of plant pathogenic bacteria can be useful to understand adaptation, mainly host adaptation, and ultimately to seek effective management strategies. *Ralstonia solanacearum* (Rs) causes bacterial wilt, a disease that affects several plant species and that is particularly severe to many crops in the tropics. Four phylotypes of Rs are characterized and their origins are geographically related. Phylotype II is an “ancestral lineage” originated in Brazil, but its evolutionary history has not been thoroughly investigated. The genealogies of 7 genes (5 in the chromosome and 2 in the megaplasmid) of 189 Brazilian isolates of Rs were reconstructed to study its evolutionary history. Four lineages were identified using a nonparametric multivariate method. The estimated mutation rate varied from 10^{-5} to 10^{-4}. Mutation was more important than recombination in the Brazilian population of Rs. Population expansion was detected for phylotype IIA. Recent introduction of phylotype I in Brazil was also identified. We anticipate that the development of resistant cultivars to bacterial wilt will be a major challenge in all host crops.
Key words:bacterial wilt, evolution, migration, molecular epidemiology, variability, resistance.

1. Introduction

Bacteria have complex lifestyles and the comprehension of the mechanisms that drive evolution can provide insights on how adaptation takes place (Stukenbrock and McDonald, 2008). Adaptation is influenced by genetic changes and environmental interactions. *Ralstonia solanacearum* [(Smith, 1896) Yabuuchi et al. 1995] is a highly variable bacterial species capable of adapting to different environments and causes disease in a broad range of host plants from different families. Additionally, the pathogen overcome plant resistance and adapt to new hosts with unprecedented facility (Peeter et al., 2013).

[5] Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil.
[6] Centro Nacional de Pesquisa de Hortaliças, Gama, Distrito Federal, Brazil.
[7] Department of Crop Sciences, University of Illinois, Urbana, Illinois, United States.
[8] E-mail: mizubuti@ufv.br.

Ralstonia solanacearum is a species complex that includes strains with probably different geographic origins, defined as phylotypes: I, originated in Asia; II, in America; III, in Africa; and IV, in Indonesia (Fegan and Prior, 2005). It has been demonstrated that different ecological and selective forces operating at each habitat can exert large influence in pathogen populations (Kinkel et al. 1987; Whipps et al. 2008; Pozo et al, 2012). The average nucleotide diversity for whole genomes of different *R. solanacearum* isolates was estimated to investigate the genetic distances among lineages of the pathogen (Remenant et al., 2010). The differences provided strong support to separate the complex into three species: *Ralstonia pseudosolanacearum* comprised of phylotypes I and III, *R. solanacearum* comprised of phylotypes IIA and IIB, and *R. syzygii* that includes the phylotype IV (Safni et al., 2014).The repeated convergent evolution of *R. solanacearum* in phylotypes has brought much attention to the mechanisms governing adaptation and driving the separation of this species. Apparently geographic isolation plays a major role is species sorting (Fegan and Prior, 2005). Separation according to host has been recorded for strains of *R. solanacearum* that cause Moko disease in banana/heliconia and brown-rot infecting potato/tomato (Peeters et al., 2013). Modern agroecossystem provide dense and uniform host populations and a homogeneous environment that favors selection. In turn, host specialization, adaptation to new environment and management practices can take place relatively rapidly (Balter, 2007; Stukenbrock and McDonald, 2008).

Parametrization of evolutionary mechanisms that affect the population of *R. solanacearum* has been done using a multilocus sequence analysis (MLSA) of isolates from several countries (Wicker et al., 2012). In this collection, isolates of the four phylotypes were detected and recombination seems to have played an important role in the evolution of the population. The largest fraction of clonality was detected in the IIB phylotype subpopulation. Parametrization of evolutionary mechanisms of local populations is important to support breeding programs aimed at developing resistant crop varieties to bacterial wilt. Understanding how genetic variation is generated, maintained and distributed in the pathogen population can help orient breeding strategies as well as the use of resistant varieties.

The first study to characterize the population of *R. solanacearum* was conducted using MLSA and there was evidence of a clonal population (Castillo and Greenberg, 2007). However, the representativeness of the samples analyzed was limited, since most isolates were of phylotype IIB associated with potato. Therefore, processes that account for the current distribution of phylotype IIA and IIB remain unresolved in Brazil. Additionally, there is no detailed information related to evolutionary mechanisms and the demographic and ecological processes that may have affected the Brazilian population of *R. solanacearum*. In this study the gene genealogies and the coalescent process of the main lineages of *R. solanacearum* present in Brazil were used to estimate parameters of evolutionary mechanisms that affect the pathogen population.

2. Material and methods

2.1. *Ralstonia solanacearum* isolates

Isolates were collected in various regions in Brazil, from different host plants and years (Table 1). Bacterial cultures were storedintubescontainingsterilizedwater or cryogenically in 40% glycerol at 80℃ (Table 1). A total of 189 isolates of *R. solanacearum* were characterized. Isolates of the three phylotypes, I, IIA and IIB, that occur in Brazil were analyzed (Fig.1) (Santiago et al., unpublished data).

2.2. DNA extraction of isolates

Isolates were grown in Kelman medium (Kelman, 1954) with agitation at 180 rpm for 48 h at 28℃. Two mL of culture were ce ntrifuged at 12500 rpm for 7 min and genomic DNA was extracted according to the instructions of the Wizard Genomic DNA Purification Kit (Promega). The DNA was kept at -20°C after treatment with RNAse and when necessary the DNA concentration was checked and adjusted using NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) and adjusted.

2.3. Genetic characterization

Partial sequences of seven genomic regions, *gdha* - glutamate dehydrogenase; *mutS* - methyl-directed mismatch repair; *ppsa* -

phosphoenolpyruvate synthase; *leuS* - leucyl-tRNA synthetase; *gyrB* - DNA gyrase, subunit B; *egl* – endoglucanase; and *flic* – flagellin were generated for each isolate using primers previously described (Prior and Fegan, 2005; Castillo and Greenberg, 2007; Ait Tayeb et al., 2008; Wicker et al., 2012). The PCR was carried out with DreamTaq™ Green PCR Master Mix (containing DreamTaq™ DNA polymerase, optimized DreamTaq™ Green buffer, MgCl2 and dNTPs), 4 pmol of each primer and 2 µL of template DNA in a total reaction volume of 25 µl for one sample (Thermo Scientific Fermentas Ltd.). Initial denaturation was at 96℃ for 2 min, followed by an amplification phase repeated 30 times (a denaturation step at 95℃ for 1 min; an annealing step for 2 min at a gene specific temperature according to Wicker et al. (2012), an extension step at 72℃ for 2 min); and a final extension step at 72℃ for 10 min in a MJ Research PTC-100 thermocycler. To confirm amplifications all samples were subjected to electrophoresis in 1% (wt/vol) agarose gel and fragment sizeswere estimatedby comparisonwith the marker. Samples that contained the expected amplicon size were incubated with ExoSAP-I at 37°C for 30 min, followed by 80°C for 10 min to inactivate the enzymes. The amplified PCR products were sequenced by Macrogen Services (Kumchun-ku, Seoul, Korea) using reverse and forward primers. The sequences were edited and aligned with other reference sequences (Table 1) using Sequencher 4.6 (Gene Codes Corp., Ann Arbor, Michigan) and Clustal W in MEGA 5.0 software (Tamura et al., 2011). Newly identified sequences were deposited in Genbank (Table 1). Concatenated sequences were organized in the same order as previously suggested (Wicker et al., 2012).

2.4. Population structure analyses

A nonparametric multivariate clustering method was also used to analyze the population structure. The discriminant analysis of principal components (DAPC) was processed in ADEGENET package (Jombart et al., 2010), in R program.DAPC was used because it does not assumed Hardy-Weinberg equilibrium. The number of clusters was determined based on the bayesian information criterion (BIC). Detecting population subdivision is important to select an appropriate model, panmictic or subdivided, in later coalescent analyses.

2.5. Genetic diversity

Genetic diversity based on each gene separately and in concatenated arrangement were estimated using haplotype number and nucleotide diversity calculated with DnaSP 5.0 (Librado and Rozas, 2009). Selection was inferred with neutrality tests (Tajima's D, Fu & Li's and F*) for each gene, for each population and phylotype, using DnaSP. 5.0.

Occurrence of intragenic and intergenic recombination was analyzed using the concatenated dataset. First, the neighbor-net method implemented in the SplitsTree4 program was used to identify reticulate regions (Hudson and Bryant, 2006). Indication of recombination breakpoints was then investigated using seven non-parametric programs: RDP, GENECONV, BOOTSCAN, MAXIMUM CHI SQUARE, CHIMAERA, SISCAN and 3SEQ in RDP v. 3.44. For optimal recombination detection, we selected the default settings for all methods and a Bonferroni corrected *P* value cut-off of 0.05. Tests to detect recombination breakpoints were also carried out using GARD (genetic algorithm for recombination detection) (Pond et al., 2006a, 2006b). Recombination was accepted when detected by at least four methods. Tests were made with concatenated genes.

The LDhat software was used to obtain the Watterson θ estimator (1975) and the composite likelihood method was used to estimate the recombination rate (ρ) (McVean et al. 2002). The relative frequency of occurrence of recombination over mutation (ρ/θ) was calculated using the gene conversion model.Tests were made for each of the seven genes separately and for the concatenated sequences.

The BEAST v. 1.8.1 software (Drummond and Rambaut, 2007) was used to infer the mutation rate and the tip dates were defined as the year of isolation. The definition of the best substitution model was based on the Akaike Information Criterion (AIC, Akaike, 1974) implemented in jModelTest 0.1 for each gene (Posada, 2008). The runs were performed using the lognormal relaxed molecular clock model with partitioned nucleotide data by codon position. We used 1 million MCMC iterations with a sampling frequency of 10000 generations and 20% of the sample was discarded as burn-in.

The IMa2 software was used to analyze asymmetric migration, to estimate the effective population size and the time of divergence between

phylotypes (Hey, 2010). All simulations were done assuming subdivision based on the phylotypes of Brazil. All loci were included in this analysis and the HKY substitution model was used. The chain length was set to 50 million generations, sampled every 1000 generation with the first 25% discarded as burn-in. The significance of migration was estimated using likelihood ratio test (LRT) (Nielsen and Wakeley, 2001).

2.6. Demographic analyses

To estimate population size dynamics in geological time, demography statistic inferences were made using the Extended Bayesian Skyline Plot (EBSP) implemented in BEAST v. 1.8 for three phylotypes using multilocus data. The Lognormal relaxed clock model was used for all partitions, substitution models and Bayesian MCMC were the same as those used to calculate the mutation rate. We used the mutation rate previously estimated with BEAST and the divergence time was converted to thousand of years.

2.7. Coalescent simulation

Coalescent gene genealogies were inferred to elucidate the ancestral history of the Brazilian population of *R. solanacearum*. Using the Mobyle SNAP Workbench for population genetic analyses (Monacell and Carbone, 2014). After sequence alignments, single genes were analyzed. Initially, the sequences were collapsed with phenotype (phylotype group) into unique haplotypes with SNAP Map (Aylor et al., 2006), with insertions, deletions and infinite-sites violations excluded. The resulting sequences were inspected for site incompatibilities, which were removed using SNAP CladeEx (Bowden et al., 2008). Neutrality, migration and Hudson's test were performed as described above for each gene and concatenated sequence. The estimation of θ and the asymmetrical migration rates were used in the likelihood estimation. Coalescence with population subdivision, assuming no recombination, was implemented in Genetree (Bahlo and Griffiths, 2000; Griffiths and Tavare, 1994) to infer all possible mutational pathways and ancestry of isolates of *R. solanacearum* for each locus. Ten million simulations of the coalescent were performed for each run. Coalescent trees were viewed with Treepic (Bahlo and Griffiths, 2000; Griffiths and Tavare, 1994).

3. Results

3.1. Genetic diversity of Brazilian isolates of *R. solanacearum*

The total length of concatenated sequences was 4240 bp, which corresponds to 929 bp for *gdha*, 414 bp for *gyr*, 662 bp for *leus*, 692 bp for *muts*, 460 bp for *ppsa*, 333 bp for *flic* and 750 bp for *egl*. The number of haplotypes ranged from 16 (*ppsa*) to 46 (*gdha*) and the number of segregating sites varied from 31 (*flic*) to 64 (*gdha*) (Table 2). The mean haplotype diversity was higher in phylotype IIB (0.74) and IIA (0.71) compared to phylotype I (0.30). Haplotype diversity approached its maximum, ranging from 0.83 to 0.91 (Table 2).

3.2. How many populations?

Nonparametric analyses pointed to four DAPC genetic clusters (Fig. 2). The groups were strongly associated with phylotype: (1) A group comprised of phylotype I and III individuals (green in Fig. 2); (2) a group of phylotype IV individuals (blue), (3) phylotype IIB predominantly with potato isolates (red); and (4) a large group of phylotype IIA isolates mainly from tomato (orange color) (Fig. 2). However, when isolates of phylotype IIB were analyzed separately, they could be separated in sub-groups IIB1 and IIB2 (Fig. S2). There were single groups of phylotype I and IIA isolates (data not shown).

3.3. Population suddivision

The pattern of population subdivision described above was confirmed by pairwise Hudson's test using single genes and concatenated alignments data not shown). In all subsequent analyses, isolates were grouped according to the phylotype clusters identified in DAPC.

3.4. Recombination

The first evidence for the occurrence of recombination was gathered from the reticulate tree of the neighbor-net reconstruction using the aligned concatenated sequences dataset (Fig. S1). Using RDP3 and GARD programs six recombination events were detected in the concatenated dataset. Recombination affected 106 isolates, from all phylotypes. Apparently only the

ppsa region is free of recombination. Estimates of the population recombination rate, ρ, ranged of 0 to 10 per site (Table 3).

3.5. Mutation

The Watterson θ estimate based on the number of segregating sites varied from 0.74 and 8.97 per site: these estimates were consistent within loci across the three phylotypes (Table 3). The mutation rates per gene, per year, were on average approximately 10^{-5}. However, the mutation rates for the *egl* of phylotype IIA isolates was one order of magnitude higher, approximately 10^{-4} substitution/site/year, than the other regions (Table 4).

The recombination and mutation estimates were used to calculate the recombination to mutation ratio (ρ/θ). Considering all isolates, the ratio varied from 0 to 0.55. Only for phylotype IIA/*fliC* the recombination rate was higher than the mutation rate: ρ/θ = 2.13.

3.6. Neutrality

The value of Tajima's D for each phylotype is shown in Table 2. Across loci, for all phylotypes, there was a trend for negative statistics, indicating that there is high frequency of variants. However, Tajima's D was only significant for *gdha* (Table 2). When isolates were organized into phylotypes, all populations had significant negative D values that can be due to population growth or selective sweeps in at least one gene. Synonymous substitutions (Ks) were more frequent than nonsynonymous (Ka) (from 5.6 to 17 times more frequent), except for the *flic* gene (Table 2). The Ka value for *flic* was more frequent than Ks. Ks is indicative of positive selection, which acts on the prevalence of nonsynonymous changes.

3.7. Population demography

Evidence of population growth was inferred from the Extended Bayesian skyline plot (EBSP). Particularly, there was clear expansion of population IIA that may have occurred in the last 10000 years. On the other hand, the phylotype IIB subpopulation appeared to have maintained constant size (Fig. 4).

3.8. Isolation-with-migration

Using a mutation rate of 10^{-7}, estimates of directional migration between regions were obtained. Isolates of North/Northeast are ancestral to those of the South/Southeast/Central. Thirteen migration rates were significantly different from zero. Migration occurs mainly from the South to the North of Brazil.

3.9. Coalescent analysis

The history of phylotypes in Brazil was investigated using housekeeping genes. The coalescence-based gene genealogies of six loci distinguished groups that were congruent with the DAPC analysis. For *ppsA*, *mut*S and *leuS*, there is a subdivision of phylotype IIA and IIB. Phylotype IIA is ancestral to phylotype II and differentiated recently in the phylotype IIB (Fig. 6). For *fliC*, *gdhA* and *gyrB*, the populations IIA and IIB were not divided. Most of the mutations that were accumulated in *leuS* and *ppsA* originated only recently, at the tip of the trees.

4. Discussion

There are not many control measures effective against bacterial wilt and planting resistant varieties is, by far, the most promising approach to reduce yield losses. However, it is well known that the durability of resistance is largely influenced by the amount and distribution of genetic variation in the pathogen population. Determining how genetic variation is generated, maintained and distributed in populations of plant pathogens can be useful to establish management strategies. Therefore, parametrization of evolutionary mechanisms has become a cornerstone in molecular epidemiology of plant diseases. Estimating gene genealogies based on the coalescent process allows the investigation of different scenarios and better parametrization of evolutionary mechanisms. The present study was designed to allow parametrization of the main evolutionary mechanisms considering subdivision in the population of *R. solanacearum*. Mutation, recombination, migration and selection act upon the Brazilian population of *R. solanacearum*. Nevertheless, mutation seems to occur more frequently and plays a major role in shaping the populations.

4.1. Origin the *R. solanacearum* in Brazil

It has been previously proposed that the origin of the species complex predate the geological separation of the continents (Hayward, 1991). After the continental drift, allopatric processes contributed to the origin of the phylotypes. According to coalescent genealogy reconstruction, phylotype IV was the first phylotype to diverge, which suggest that *R. solanacearum* originated in Australia/Indonesia. Apparently, phylotype II originated in the Americas and phylotypes I and III derived from phylotype IV. The subdivision of phylotypes IIA and IIB probably occurred at the same time and most likely this took place in Brazil (Wicker et al., 2012). In the present work wide distribution and high variability in phylotype II isolates was detected, which corroborates the hypothesis of an old lineage. Additionally, when analyzing the haplotype network there is support for Brazil as the origin of phylotype II.

Several wild species of *Solanum* such as *S. chacoense*, *S. calvescens*, *S. muelleri*, *S. commersonii* and *S. dulcamara* are commonly found in the South, Central and Southeast of Brazil (Hijmans and Spooner, 2000) and the endemic presence of the pathogen may evidence co-evolution between the host plant and the pathogen. Thus, it can be claimed that the origin of phylotype II is in Brazil. There is sufficient evidence for an overlap in the center of origin of the pathogen and of the host crop. This also occurs in other plant pathogenic bacteria as *Xanthomonas oryzae* pv. *oryzae*and *X. axonopodis* pv. *citri*, causal agents of bacterial blight of rice, and citrus canker, respectively. Both seem to have originated in south Asia, as their hosts (Scortichini, 2005).

4.2. Strong population subdivision despite no geographic separation

Strong population subdivision within phylotype IIA and IIB was detected with no correlation to their geographic origin despite the low capacity for dissemination. What led to the subdivision remains speculative, but ecological specialization could have contributed to divide phylotypes IIA and IIB. This pattern stands in contrasting to the worldwide geographic separation of phylotype (Fegan and Prior, 2005). Although sympatric adaptation was also observed in phylotype I and III, phylotype I is more adapted to lowland environment and III to highland. Ecological specialization has also been

allegedly involved in the separation of phylotypes of *R. syzygii* and the Blood Disease Bacterium (Remenant et al., 2011).

Based on gene genealogies phylotype IIA originated first from phylotype II, then it was followed by phylotype IIB. Phylotype differentiation could be attributable to host selection combined with cultivation in different ecozones: one in the cool mountainous highlands versus cultivation in the tropical lowlands results from minor genetic factor differentially affecting pathogen fitness. *R. solanacearum* race 3 biovar 2, phylotype IIB, found in cooler areas in Brazil remained viable longer than other strains of phylotypes IIA when associated with host tissue (Scherf et al., 2010). We found indirect evidence of positive selection in all phylotypes, with Ka:Ks ratio >1 for three genes (*flic*, *egl* and *leuS*). Moreover, F_{st} analysis using the same isolates showed high difference in the tomato and potato (Santiago et al., unpublished data). Difference in fitness can influence population dynamics by allowing the persistence of an individual over many generations.

4.3. Proposed scenario for historical migration of *R. solanacearum*

According to the isolation-with-migration model, Brazilian isolates from the North/Northeast regions are more ancestral than those from the South, Southeast, and Central regions. Migration is occurring between geographic regions, mainly from the south to the north. Apparently, this migration pattern is preventing more accentuated geographic separation of phylotypes IIA and IIB.

Groups of Brazilian isolates of *R. solanacearum* can be influenced by the interaction of migration and selection in the phylotype II. Changes in temperature, absence of host and easy/efficient dispersion by potato tubers made phylotype IIB differentiate into phylotype IIB.1 and IIB.2. It probably happened recently, after the 1950s, with the intensification of potato seed commercialization to the North and Northeast regions. The topology of the neighbor-net phylogeny method and DAPC supports the separation of groups, with subpopulation with long branch lengths separating phylotype IIB.1 and IIB.2.

Human activities can also affect the dispersal pattern and the current distribution of individuals. Our findings suggest that bacteria dispersal occurs at larger distances (greater than hundreds of meters), than expected for natural

dispersal of a soilborne pathogen. Such long-distance dispersal event is well known for plant pathogenic fungi. It is possible that transportation of infected seeds or natural dispersal by rivers can contribute to long-distance dispersal of *R. solanacearum*.

4.4. Demographic history of phylotype II

Both phylotypes IIA and IIB have similar (high) diversity and recombination seems to occur, however phylotype IIA is expanding and IIB not. Host specialization can interfere in the population size (Nielsen, 2005). The constant size of the phylotype IIB population can be due to high availability of tomato and potato plants, the two most planted and commercialized vegetable crops in Brazil.

It is extremely difficult to construct any scenario in which large allelic differences could have accumulated in *R. solanacearum* via point mutation. Demographic processes of the pathogen population suggest that plant domestication by humans in America had a dramatic and lasting impact on the expansion of phylotype IIA in the last 10000 years (Fig. 3). Independent centers of origin and domestication were identified in America. Evidences of domestication were obtained indirectly by radiocarbon dating of organic seed (Erickson et al., 2005). The seeds were collected in dry caves and rock shelters and earliest dates are estimated between 7,000 and 10,000 years ago (Erickson et al., 2005). Several studies have retraced pathogen evolution during domestication (Couch et al., 2005; Munkacsi et al., 2008; Zaffarano et al., 2008). Similar processes could have happen with bacterial wilt in several crops. Modern agricultural practices and cultivar development contributes to growth of plant pathogen populations. Large areas cultivated with genetically uniform and high density planting of plant species favor plant disease epidemics (Yarwood, 1970; Hansen, 1987; Stukenbrock and McDonald, 2008).

4.5. Introduction of phylotype I

Phylotype I isolates, originated in Asia, are adapted to tropical conditions, have wide host range and can spread over long distance very rapidly (Wicker et al., 2012). The analysis of the datasets of the present study points to a recent introduction of phylotype I in Brazil. We hypothesized that a

bottleneck contributed to the lack of genetic variation and an excess of rare alleles. Wicker et al. (2012) found high levels of recombination in phylotype I isolates, however Brazilian isolates are nearly clonal. Moreover, the negative Tajima'D values was detected, which measures the difference between nucleotide diversity and a weighted estimate of nucleotide polymorphism (Tiffin and Moeller, 2006). In the case of *R. solanacearum* higher variability was found in the *leuS* gene and recent mutations were mapped to phylotype I in the coalescent tree. For phylotype I, population bottlenecks may have occurred recently. Together these results support a recently introduction in Brazil and phylotype I isolates have probably been dispersed by human activities related to agriculture.

4.6. Population level of recombination

Evidence of recombination within related subgroups of *Ralstonia* has been found, but we inferred that mutation contributed more to the genetic diversity than recombination in the population of *R. solanacearum* in Brazil. High mutation rates per gene/year were found to be around 10^{-5}, however rates for phylotype IIB were lower. This rate of mutation is similar to those estimated for other plant pathogenic bacteria (Mhedbi-Hajri et al. 2013). Recombination lead to important variation in population structure of pathogenic bacteria that need to be considered when predicting the ways the pathogens may respond to new habitats and hosts (Didelot and Maiden, 2010). Recombination is a major evolutionary mechanism acting upon the populations of two plant pathogenic bacteria that cause leaf spots. The emergence of two lineages of *Xanthomonas fastidiosa* subsp. *pauca*, a pathogen of citrus and coffee plants, is attributed to be due to recombination event (Nuney et al., 2012). For *X. axonopodis*pv. *campestris*MLST studies revealed that recombination occurred as often as point mutation (Fargier et al., 2007; Mhedbi-Hajri et al., 2013). Analysis of *R. solanacearum* collected worldwide found that recombination is important to the diversification of this species, regardless of the high clonality in phylotype IIB population (Wicker et al., 2013). Difference in the results probably is related to the larger number of isolates of phylotype IIB in Brazil, the absence of phylotype III, and the large number of phylotype I isolates that showed high level of recombination. Additionally, it is not clear in the previous analysis if all the

assumptions of mutation models were checked before the recombination analysis was conducted. Contrary to what has been reported, Castillo and Greenberg (2007) concluded that *R. solanacearum* is a clonal complex, a finding that is quite distinct from the results of Wicker et al. (2012). This discrepancy could be partially explained by the nature of the data set that had high frequency of phylotype IIB isolates and small sample sizes for phylotype I, III and IV.

5. Conclusion

The Brazilian population of *R. solanacearum* is recombining and there is high mutation rate. These contribute to the establishment of genetically diverse populations and it is expected that cultivar resistance will be easily overcome (Vinatzer et al., 2014). Breeding programs should take this into account and strategies for developing resistant cultivars may need to be re-oriented. More investigation need be conducted in order to answer new questions raised by this work: (1) when did the division of phylotype II occur? (2) Which ecological factor is determining the dynamic of phylotype II? (3) What are the major dispersal agents of *R. solanacearum* in Brazil. Moreover, the deeper analysis of the genome of *R. solanacearum* needs to be conducted to investigate adaptive differences at the molecular level and how it ranges across such a wide variety of habitats.

6. Acknowledgements

This research was supported by grants of the CNPq (476940/2011-7). The authors would like to express special thanks to Dr. Ignazio Carbone and José Cleydson da Silva for their help with the analysis.

7. References

Ahmed, H. U., Finckh, M. R., Afonso, R. F., Mundt, C. C. 1997. Epidemiological effect of gene deployment strategies on bacterial blight of rice. Phytopathology 87: 66-70.

Ait Tayeb, L., Lefevre, M., Passet, V., Diancourt, L., Brisse, S., Grimont, P. A. D. 2008. Comparative phylogenies of *Burkholderia*, *Ralstonia*, *Comamonas*, *Brevundimonas* and related organism derived from *rpoB*, *gyrB* and *rrs* gene sequences. Research in Microbiology 159:169-177.

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions Automatic Control 19:716-723.

Aylor, D.L., Price, E.W., Carbone, I. 2006. SNAP: combine and map modules for multilocus population genetic analysis. Bioinformatics 22:1399-1401.

Bahlo, M. and Griffiths, R. C. 2000.Inference from gene trees in a subdivided population. Theoretical Population Biology 57:79–95.

Balter, M. 2007.Seeking agriculture's ancient roots.Science316:1830-35.

Bowden, L.C., Price, E.W., Carbone, I. 2008. SNAP Clade and Matrix, Version 2. Available from: http://www.cals.ncsu.edu/plantpath/faculty/carbone/home.

Castillo, J. A. and Greenberg, J. T. 2007.Evolutionary dynamics of *Ralstonia solanacearum*. Applied and Environmental Microbiology 73:1225-1238.

Couch, B. C., Fudal, I., Lebrun, M-H., Tharreau, D., Valent, B., van Kim, P., Nottéghem, J-L., Kohn, L. M. 2005. Origins of host- specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice.Genetics.170:613-630.

Didelot, X. and Maiden, M. C. J. 2010.Impact of recombination on bacterial evolution. Trends in Microbiology 18: 315-322.

Drummond, A. J. and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling tree. BMC Evolutionary Biology 7: 214-222.

Erickson, D. L., Smith, B. D., Clarke, A. C., Sandweiss, D. H., Tuross, N. 2005. An Asian origin for a 10,000-year-old domesticated plant in Americas. PNAS 102:18315-18320.

Excoffier, L., Laval, G. and Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary Bioinformatics 1: 47-50.

Falush, D., Stephens, M., Pritchard, J. K. 2003. Inference of population structure using multilocus genotype data linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Fargier, E. and Manceau, C. 2007. Pathogenicity assays restrict the species *Xanthomonas campestris* into three pathovars and reveal nine races within *X. campestris* pv. *campestris*. Plant Pathology 56:805-818.

Fegan, M. and Prior, P. 2005. How complex is the "*Ralstonia solanacearum* species complex"? Pages 449-461 in: Bacterial wilt disease and the *Ralstonia solanacearum* species complex. C. Allen, P. Prior, and A. C. Hayward. American Phytopathological Society Press, St. Paul, MN.

Feil, E. J. and Spratt, B. G. 2001.Recombination and the population structures of bacterial pathogens. Annual Reviews of Microbiology 55:561-590.

Gonzalez, A., Ronce, O., Ferriere, R., Hochberg, M. E. 2014. Evolutionary rescue: an emerging focus at the intersection between ecology and evolution. Philosophical Transactions of the Royal Society B 368:1-8.

Goss, E. M., Potnis, N., Jones, J. B. 2013. Grudgingly sharing their secrets: new insight into the evolution of plant pathogenic bacteria. New Phytologist 199:630-632.

Griffiths, R. C. and Tavare, S. 1994. Ancestral inference in population genetics. Statistical Science 9:307-319.

Hayes, B., Tesar, B., Zuraw, K. 2003. OTSoft: Optimality Theory Software, Version 2.1. Available from: <http://www.linguistics.ucla.edu/people/hayes/otsoft/>

Hayward, A. C. 1991. Biology and epidemiology of bacterial wilt caused by *Pseudomonas solanacearum*. Annual Review Phytopathology 29: 65-87.

Hayward, A.C. 1994. The hosts of *Pseudomonas solanacearum*. Pages 9-24 in: Bacterial Wilt: The disease and its causative agent, *Pseudomonas solanacearum*. A. C. Hayward and G. L. Hartman. CAB International, Wallingford, UK.

Hey J. 2010. Isolation with migration models for more than two populations. Molecular Biology and Evolution 27:905-920.

Hijmans, R. J. and Spooner, D. M. 2001. Geographic distribution of wild potato species. American Journal Botany 11:2101-2112.

Hijmans, R. J., Gavrilenko, T., Stephenson, S., Bamberg, J., Salas, A., Spooner, D. M. 2007. Geographical and environmental range expansion through polyploidy in wild potatoes (Solanum section Petota). Global Ecology of Biogeography 16:485-495.

Hudson, R. R., Boos, D. D., Kaplan, N. L. 1992. A statistical test for detecting geographic subdivision. Molecular Biology and Evolution 9: 138-151.

Huson, D. H. and Bryant, D. 2005. Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution 23: 254-267.

Hutchinson, G. E. 1957. Concluding Remarks. Cold Spring Harbor Symposia on Quantitative Biology 22: 415-427.

Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403-1405.

Kinkel, L. L., Andrews, J. H., Berbee, F. M., Nordheim, E. V. 1987. Leaves as islands for microbes. Oecologia 71: 405-408.

Pond, S. L. K., Posada, D., Gravenor, M. B., Woelk, C. H., Frost, S. D. W. 2006a.Automated phylogenetic detection of recombination using a genetic algorithm. Molecular Biology and Evolution 23:1891-1901.

Pond, S. L. K., Posada, D., Gravenor, M. B., Woelk, C. H., Frost, S. D. W. 2006b. GARD: a genetic algorithm for recombination detection. Bioinformatics 22:3096-3098.

Librado, L. and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuvre, P. 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462-2463.

Matute, D. R., Quesada-Ocampo, L. M., Rauscher, J. T., McEwen, J. G. 2008. Evidence for positive selection in putative virulence factors within the *Paracoccidioides brasiliensis* species complex. PloS Neglected Tropical Diseases 2: e296.

McDonald, B. A. and Linde, C. 2002.The population genetics of plant pathogens and breeding strategies for durable resistance. Euphytica 124:163-180.

McVean, G., Awadalla, P., Fearnhead, P. 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics 160:1231-1241.

Mhedbi-Hajri, N., Hajri, A., Boureau, T., Darrasse, A., Durand, K., Brin, C., Saux, F., Manceau, C., Poussier, S., Pruvost, O., Lemaire, C., Jacques, M-A. 2013. Evolutionay history of the plant pathogenic bacterium *Xanthomonas axonopodis*. PLoS One 8: e58474.

Monacell, J. T. and Carbone, I. 2014. Mobyle SNAP Workbench: a web-based analysis portal for population genetics and evolutionary genomics. Bioinformatics : btu055.

Munkacsi, A. B., Stoxen, S., May, G. 2008.*Ustilago maydis* populations tracked maize through domestication and cultivation in the Americas. Proceedings of the Royal Society B: Biological Sciences 275: 1037-1046.

Nielsen, R. 2005. Molecular signatures of natural selection. Annual Review of Genetics 39:197-218.

Nielsen, R. and Wakeley J. 2001.Distinguishing migration from isolation: A Markov chain Monte Carlo approach.Genetics158:885-896.

Nunney, L., Vickerman, D. B., Bromley, R. E., Russell, S. A., Hartman, J. R., Morano, L. D., Stouthamer, R. 2013. Recent evolutionary radiation and host plant specialization in the *Xylella fastidiosa* subspecies native to the United States. Applied Environmental Microbiology 79: 2189-2200.

Peeters, N., Guidot, A., Vailleau, F., Valls, M. 2013.*Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era. Molecular Plant Pathology 14: 651-662.

Posada, D. 2008.jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25:1253-1256.

Pozo, M. I., Lachance, M-A., Herrera, C. M. 2012. Nectar yeasts of two southern Spanish plants: the roles of immigration and physiological traits in community assembly. FEMS Microbiology Ecology 80: 281-293.

Pritchard, J., Stephens, M., Pritchard, J. 2000.Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Remenant, B., Coupat-Goutaland, B., Guidot, A., Cellier, G., Wicker, E., Allen, C., Fegan, M., Pruvost, O., Elbaz, M., Calteau, A., Salvignol, G., Mornico, D., Mangenot, S., Barbe, V., Medigue, C., Prior, P. 2010. Genomes of three tomato pathogens within the *Ralstonia solanacearum* species complex reveal significant evolutionary divergence. BMC Genomics 11:379.

Remenant, B., de Cambiaire, J. C., Cellier, G., Jacobs, J. M., Mangenot, S., Barbe, V., Lajus, A., Vallenet, D., Medigue, C., Fegan, M., Allen, C., Prior, P. 2011. The genomes of the insect-transmitted plant pathogens *Ralstonia syzygii* and Blood Disease Bacterium reveal a recent evolutionary origin within the *R.solanacearum* species. PLoS One 6: e24356.

Safni, I., Cleenwerck, I., Vos, P. D., Fegan, M., Sly, L., Kappler, U. 2014. Polyphasic taxonomic revision of the *Ralstonia solanacearum* species complex: proposal to emend the descriptions of the *R. solanacearum* and *R. syzygii* and reclassify current *R. syzygii* strain as *Ralstonia syzygii* subsp. *syzygii*, *R. solanacearum* phylotype IV strains as *Ralstonia syzygii* subsp. indomesiensis subsp. nov., banana blood disease bacterium strains as *Ralstonia syzygii* subsp. celebesensis subsp. nov. and*R. solanacearum* phylotype I and III strains as *Ralstonia pseudosolanacearum* sp. nov. International Journal of Systemic and Evolutionary Microbiology 64:3087-3103.

Scherf, J. M., Milling, A., Allen, C. 2010. Moderate temperature fluctuations rapidly reduce the viability of *Ralstonia solanacearum* race 3, biovar 2, in infected geranium, tomato, and potato plants. Applied Environmental Microbiology 76:7061-7067.

Scortichini, M. 2005. The population structure of some plant pathogenic bacteria: an ecological and adaptative perspective. Journal Plant Pathology 87:5-12.

Strukenbrock, E. H. and McDonald, B. A. 2008.The origins of plant pathogens in agro-ecosystems. Annual Review Phytopathology 46:75-100.

Swanson, J. K., Montes, L., Mejia, L., Allen, C. 2007. Detection of latent infections of *Ralstonia solanacearum*race 3 biovar 2 in geraniums. Plant Disease 91:828-834.

Tamura, K., Dudley, J., Nei, M., Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24:1596-1599.

Tiffin, P. and Moeller, D. A. 2006.Molecular evolution of plant immune system genes. Trends in Genetics 22: 662-670.

Toukam, G. M. S., Cellier, G., Wicker, E., Guilbaud, C., Kahane, R., Allen, C., Prior, P. 2009. Broad diversity of *Ralstonia solanacearum* strains in Cameroon. Plant Disease93: 1123-1130.

Vinatzer, B., Monteil, C. L., Clark, C. R. 2014. Harnessing population genomics to understand how bacterial pathogens emerge, adapt to crop hosts and disseminate. Annual Review of Phytopathology 52:19-43.

Watterson, G. A. 1975. On the number of segregating sites in genetic models without recombination. Theoretical Population Biology 7:256-276.

Whipps, J. M., Hand, P., Pink, D., Bending, G. D. 2008. Phyllosphere microbiology with special reference to diversity and plant genotype. Journal Applied Microbiology 105:1744-1755.

Wicker, E., Lefeuvre, P., Cambiaire, J. C., Lemaire, C., Poussier, S., Prior, P. 2012. Contrasting recombination patterns and demographic histories of the plant pathogen *Ralstonia solanacearum* inferred from MLSA. International Journal of Systemic and Evolutionary Microbiology.6:961-974.

Zaffarano, P. L., McDonald, B. A., Linde, C. C. 2008. Rapid speciation following recent host shifts in the plant pathogenic fungus *Rhynchosporium*. Evolution 62:1418-1436.

FIGURES AND TABLES

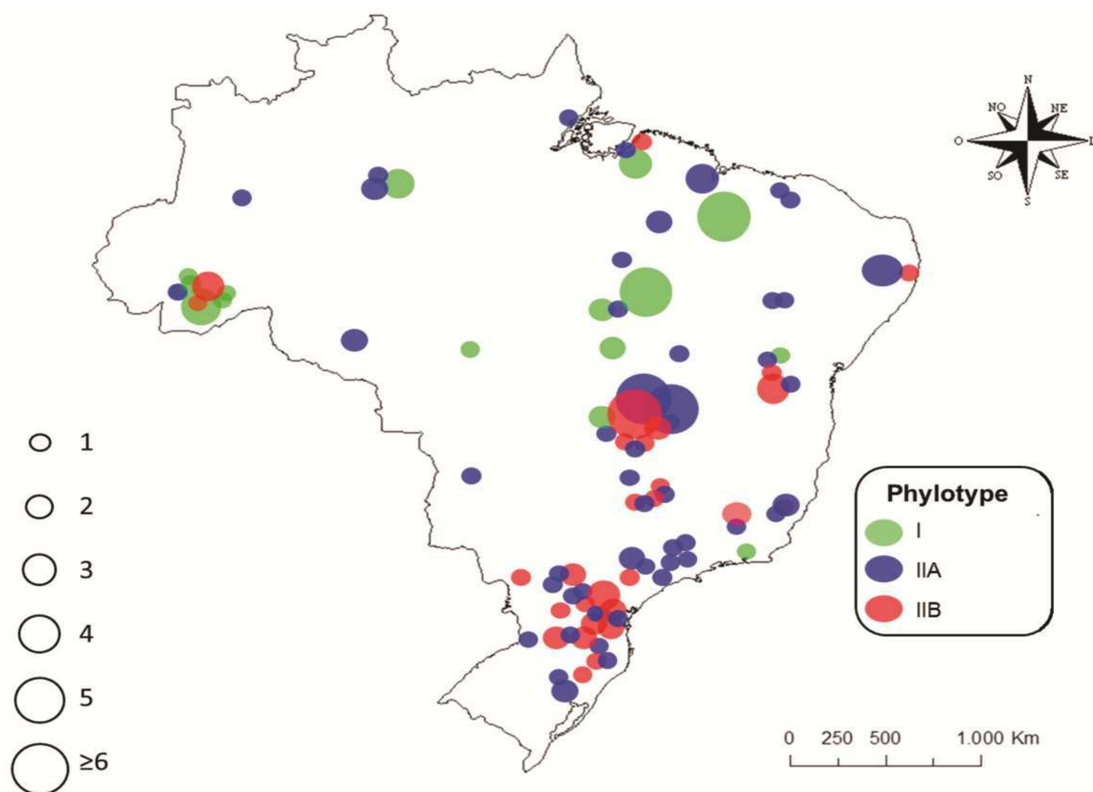


Figure 1. Sampling locations of phylotypes I, IIA and IIB of *Ralstonia solanacearum* in Brazil. Circles are drawn such as to represent the number of isolates.

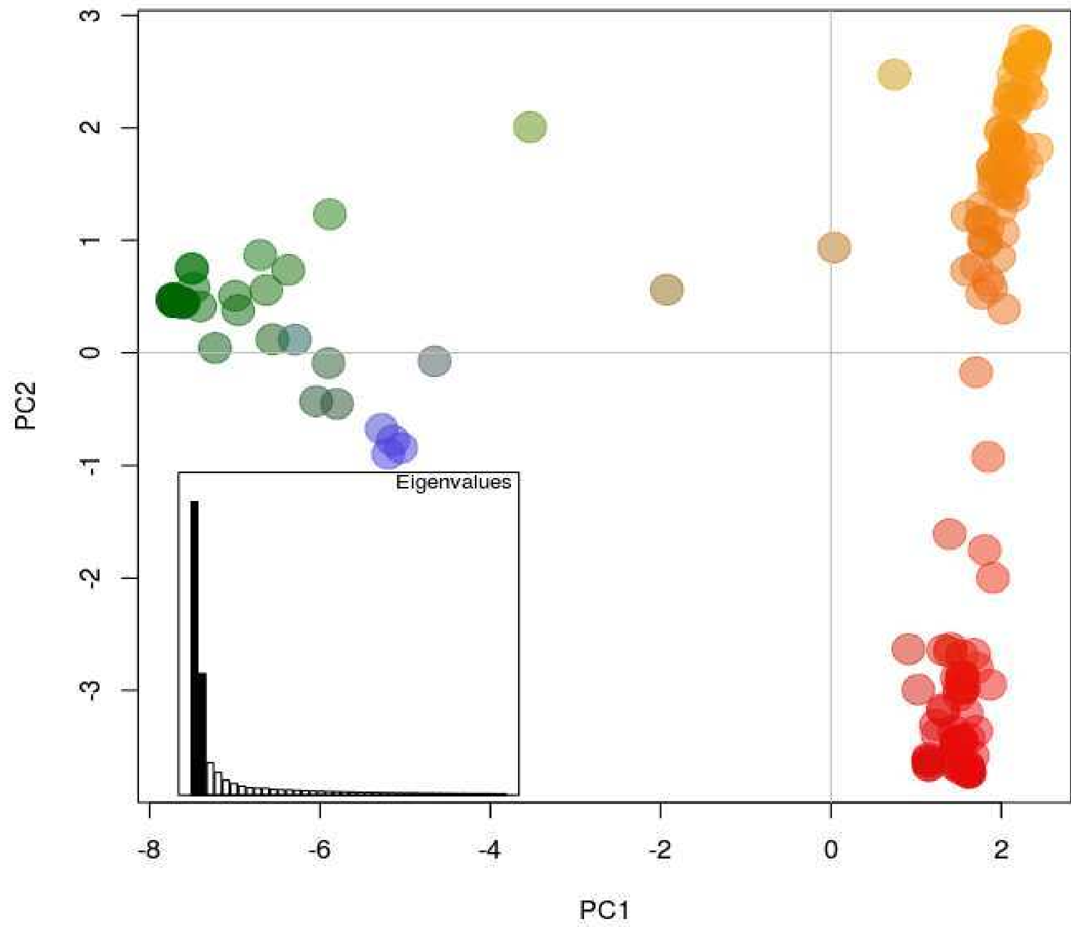


Figure 2. Scatter-plot of the phylotype distribution of *Ralstonia solanacearum* isolates into four genetic subgroups (1) Phylotype I and III with green color, (2) Phylotype IV with blue color, (3) Phylotype IIB with red color and (4) Phylotype IIA with orange color.

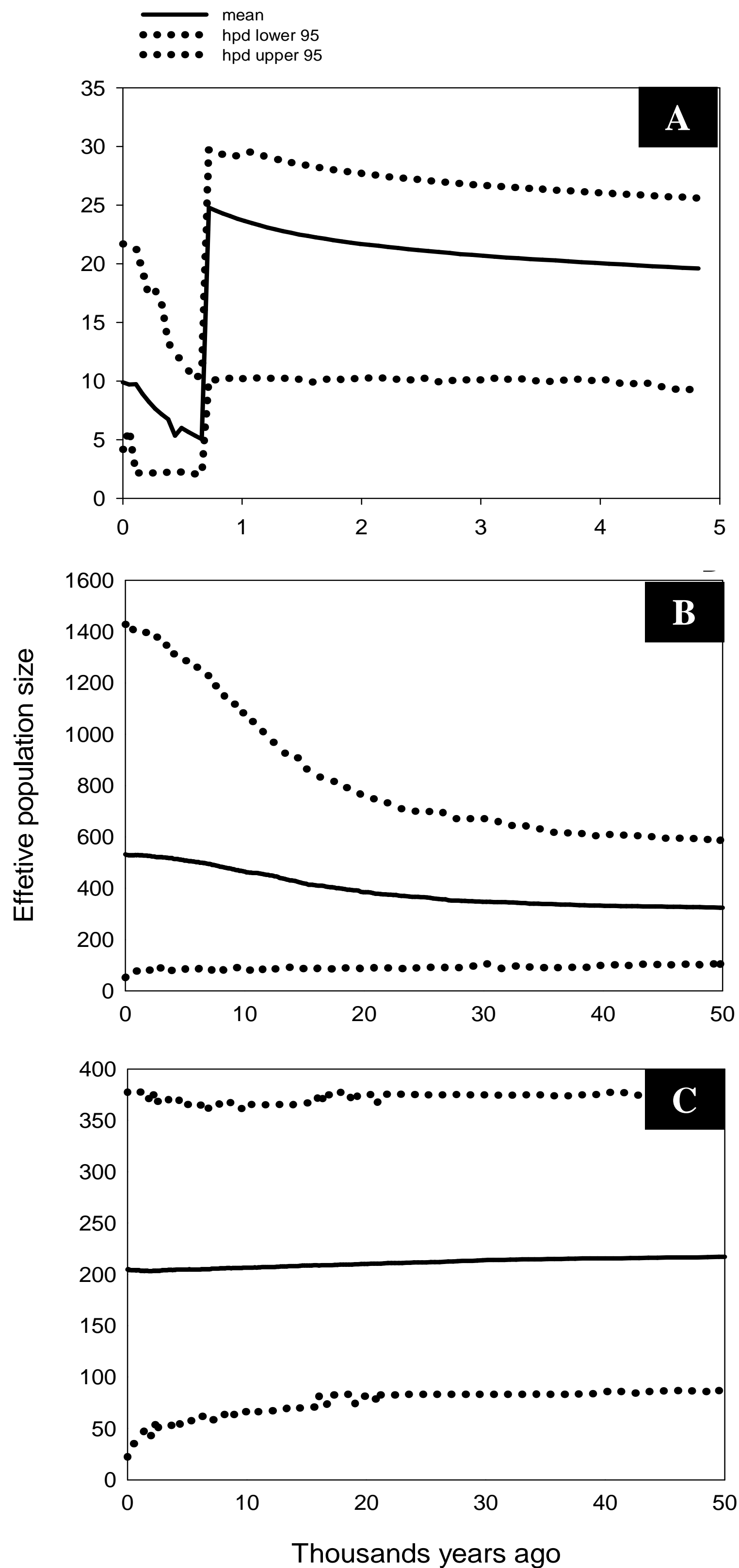


Figure 3. Bayesian Skyline Plots (EBSP) representing population size changes over time. The X-axis is time in millions of years. Y-axis is the mean effective population size in millions of individuals for: (A) Phylotype I, (B) Phylotype IIA and (C) Phylotype IIB. Dotted lines represent the upper and lower bounds of the 95% highest posterior density (hpd).

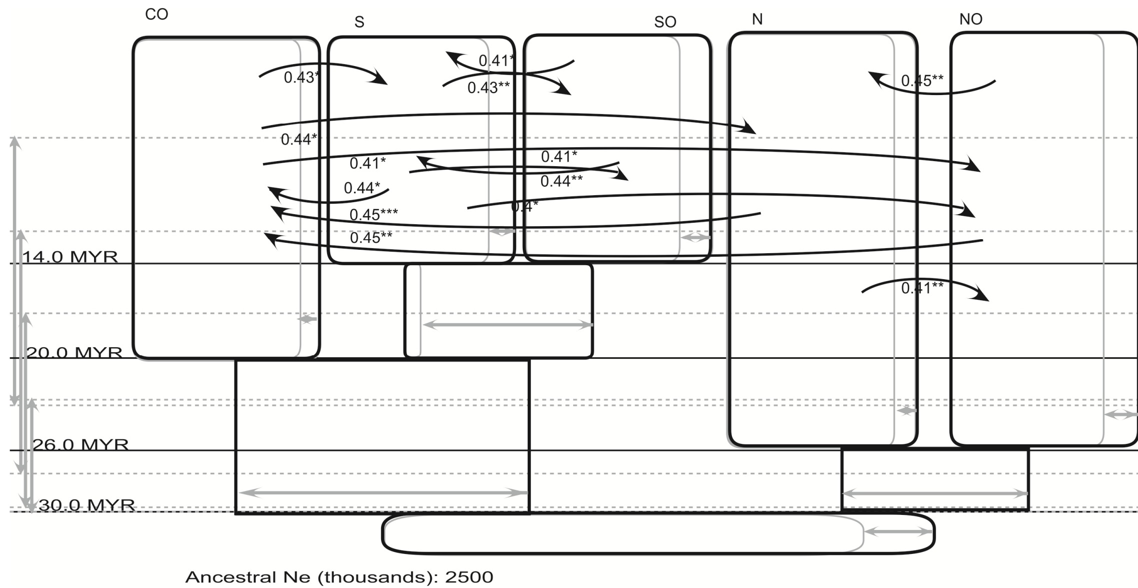


Figure 4. Migration of 189 isolates of *R. solanacearum* among different geographic regions in Brazil. Time divergence was estimated in millions of years (MYR). Direction of migration are represented by red arrows, and the numerical values represent the effective number of migrants that were significantly different from zero (p<0.05).

A

Haplotype	B	C	D	J	H	I	K	M	L	A	E	F	G	N	O	P
Frequency	33	1	1	1	1	43	6	1	9	28	2	3	1	3	49	7
Phylotype IIA	.	.	.	1	1	42	6	1	9	.	.	.	.	.	.	1
Phylotype IIB	3	.	.	.	.	1	.	.	.	28	2	3	1	3	49	6
Phylotype I	30	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.

B

	B	E	S	A	F	G	I	Q	R	J	O	N	D	K	H	L	M	P	T	C	U
	31	2	1	36	1	3	1	49	1	1	7	1	1	21	1	7	1	20	1	2	1
	1	.	1	36	.	3	1	49	1	.	.	.	.	.	.	.	.	.	1	2	1
	.	.	.	.	1	.	.	.	.	1	7	1	1	21	1	7	1	20	.	.	.
	30	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

C

	E	F	G	B	D	H	I	J	K	A	C	O	P	Q	U	M	N	L	R	S	T	V	W	X	Y
	1	1	1	1	26	1	1	1	2	2	1	1	1	1	3	6	1	1	16	1	12	16	1	1	90
	.	.	.	.	3	.	.	.	.	2	1	1	1	1	.	.	.	.	.	.	.	1	1	1	84
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	6	1	1	16	1	12	15	.	.	6
	1	1	1	1	23	1	1	1	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

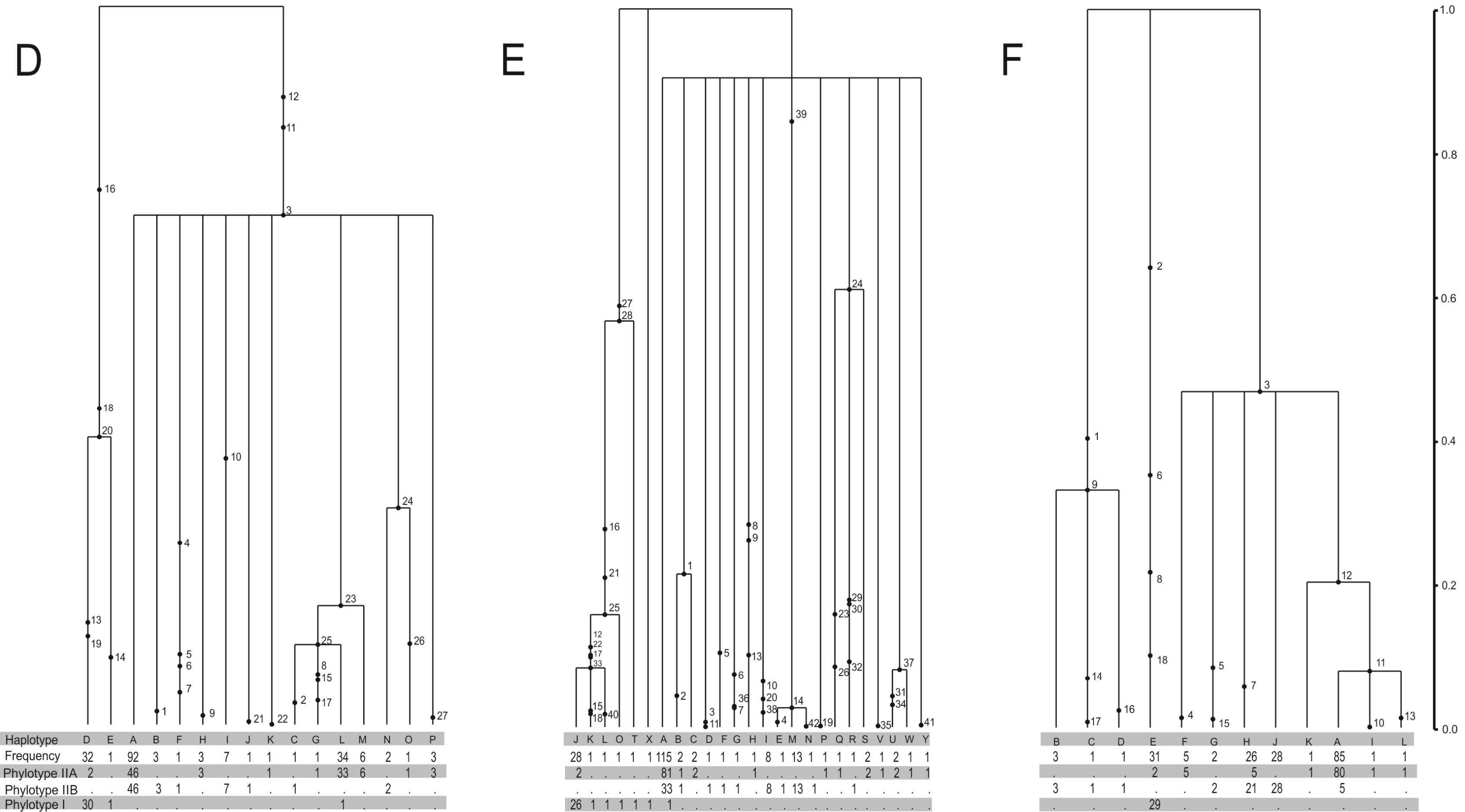


Figure 6. Colescence-base gene genealogies of the *mutS* (a), *leuS* (b), *ppsA* (c), *gyrB*(d), *gdha* (e) and *fliC* (f) show the distribution of mutation in phylotype IIA, IIB and I of *R. solanacearum*. The number of haplotypes represented in each branch is shown below the trees.

Table 1. Origin and supporting information on the *R. solanacearum* isolates used in this study

Isolate code	Phylotype	Host	City/State	Year
7	I	*Capsicum annuum*	Manaus/AM	1987
18	I	*Solanum lycopersicum*	Manaus/AM	1987
33	I	*Solanum lycopersicum*	Macapá/AP	1988
34	I	*Solanum lycopersicum*	Macapá/AP	1988
42	I	*Solanum lycopersicum*	Macapá/AP	1988
198	I	*Capsicum annuum*	Brasília/DF	2000
291	I	*Solanum lycopersicum*	Guaraí/TO	2008
310	I	*Solanum lycopersicum*	Fortaleza do Tabocão/TO	2008
311	I	*Solanum lycopersicum*	Fortaleza do Tabocão /TO	2008
316	I	*Solanum lycopersicum*	Fortaleza do Tabocão /TO	2008
319	I	*Solanum lycopersicum*	Fortaleza do Tabocão /TO	2008
322	I	*Solanum lycopersicum*	Fortaleza do Tabocão /TO	2008
324	I	*Solanum lycopersicum*	Fortaleza do Tabocão /TO	2008
326	I	*Solanum lycopersicum*	Fortaleza do Tabocão /TO	2008
352	I	*Solanum lycopersicum*	Guaraí/TO	2008
358	I	*Solanum melongena*	Gurupi/TO	2008
359	I	*Solanum melongena*	Gurupi/TO	2008
376	I	*Solanum lycopersicum*	Sinop/MT	2008
378	I	*Solanum lycopersicum*	Manaus/AM	2008
381	I	*Solanum lycopersicum*	Nova Friburgo/RJ	2008
388	I	*Capsicum* sp.	Rio Branco/AC	2009
390	I	*Capsicum* sp.	Rio Branco/AC	2009
391	I	*Solanum lycopersicum*	Rio Branco/AC	2009
424	I	*Solanum lycopersicum*	Bujari/AC	2010
425	I	*Solanum lycopersicum*	Porto Acre/AC	2010
449	I	*Capsicum* sp.	Brasília/DF	2011
470	I	*Solanum lycopersicum*	Dom Pedro/MA	2011
471	I	*Solanum lycopersicum*	Dom Pedro/MA	2011
472	I	*Solanum lycopersicum*	Dom Pedro/MA	2011
476	I	*Solanum lycopersicum*	Dom Pedro/MA	2011
477	I	*Solanum lycopersicum*	Dom Pedro/MA	2011
508	I	*Solanum lycopersicum*	Utinga/BA	2013
AC8	I	*Capsicum* sp.	/AC	2012
2	IIB	*Solanum tuberosum*	Brasília/DF	1987
3	IIB	*Solanum tuberosum*	Brasília/DF	1987
12	IIB	*Solanum tuberosum*	Itapetininga/SP	1987
26	IIB	*Solanum tuberosum*	Castro/PR	1988
28	IIB	*Solanum tuberosum*	Castro/PR	1988
30	IIB	*Solanum tuberosum*	Umuarama/PR	1988
31	IIB	*Solanum lycopersicum*	Macapá/AP	1988
44	IIB	*Solanum tuberosum*	Brasília/DF	1989
50	IIB	*Solanum melongena*	Brasília/DF	1982
53	IIB	*Solanum melongena*	São Luiz/MA	1984
56	IIB	*Solanum melongena*	Recife/PE	1989
68	IIB	*Solanum tuberosum*	Ibicoara/BA	1990
69	IIB	*Solanum tuberosum*	Ibicoara/BA	1990
71	IIB	*Solanum melongena*	Brasília /DF	1990
77	IIB	*Solanum melongena*	Brasília/DF	1991
81	IIB	*Solanum melongena*	Brasília/DF	1992
83	IIB	*Solanum melongena*	Brasília/DF	1992
84	IIB	*Solanum melongena*	Brasília/DF	1992
91	IIB	*Solanum tuberosum*	Brasília/DF	1992
92	IIB	*Solanum tuberosum*	Brasília/DF	1992
97	IIB	*Solanum tuberosum*	Piraí do Sul/PR	1992

Continuing...

Table 1. Continued

Isolate code	Phylotype	Host	City/State	Year
99	IIB	*Solanum tuberosum*	Piraí do Sul/PR	1992
113	IIB	*Solanum tuberosum*	Mariental/PR	1992
115	IIB	*Solanum tuberosum*	Lapa/PR	1992
120	IIB	*Solanum tuberosum*	/PR	1992
121	IIB	*Solanum lycopersicum*	V. Bonita/SC	1992
129	IIB.2	*Cucumis sativus*	Brasília/DF	1995
143	IIB	*Solanum tuberosum*	Viçosa/MG	1997
148	IIB.2	*Solanum tuberosum*	Coleção do IB/SP	1977
149	IIB	*Solanum tuberosum*	Coleção do IB/SP	1976
154	IIB	*Solanum tuberosum*	Guarapuava/PR	1997
156	IIB	*Solanum tuberosum*	Cristalina/GO	1997
177	IIB.2	*Piper hispidinervum*	Belém/PA	1998
191	IIB.2	*Piper hispidinervum*	Rio Branco/AC	2000
192	IIB.2	*Piper hispidinervum*	Rio Branco/AC	2000
193	IIB	*Piper hispidinervum*	Rio Branco/AC	2000
202	IIB	*Solanum tuberosum*	Castro/PR	2001
210	IIB.2	*Musa* sp.	Manaus/AM	2002
232	IIB	*Solanum tuberosum*	Buritis/MG	2003
238	IIB	*Solanum tuberosum*	Guarapuava/PR	2003
245	IIB	*Solanum tuberosum*	Luziânia/GO	2004
246	IIB	*Solanum tuberosum*	Uberaba/MG	2005
248	IIB	*Solanum tuberosum*	São José dos Pinhais/PR	2005
249	IIB	*Solanum lycopersicum*	São José dos Pinhais/PR	2005
253	IIB.2	*Solanum tuberosum*	Contenda/PR	2005
254	IIB	*Solanum tuberosum*	Contenda/PR	2005
261	IIB	*Solanum tuberosum*	Mucugê/BA	2005
269	IIB	*Solanum tuberosum*	Cristalina/GO	2006
405	IIB	*Solanum lycopersicum*	Lebon Regis/SC	2010
411	IIB.2	*Piper hispidinervum*	Rio Branco/AC	2010
420	IIB	*Solanum tuberosum*	Bom Jesus/RS	2010
427	IIB	*Solanum tuberosum*	Guarapuava/PR	2010
429	IIB	*Solanum tuberosum*	Araxá/MG	2010
431	IIB	*Solanum lycopersicum*	Faxinal/PR	2010
432	IIB	*Solanum lycopersicum*	Faxinal/PR	2010
442	IIB	*Solanum lycopersicum*	Urupema/SC	2011
461	IIB	*Solanum lycopersicum*	Ibicoara/BA	2011
481	IIB	*Eucalyptus* sp.	Brasília, DF	2011
482	IIB	*Eucalyptus* sp.	Brasília/DF	2011
512	IIB	*Solanum tuberosum*	Serra do Salitre/MG	2013
513	IIB	*Solanum tuberosum*	Serra do Salitre/MG	2013
W40	IIB	*Solanum tuberosum*		1979
W81	IIB	*Solanum tuberosum*		1997
W299	IIB	*Pelargonium* sp.		2002
13	IIA	*Solanum tuberosum*	Brasília/DF	1987
23	IIA	*Solanum tuberosum*	Luis Eduardo Magalhães/BA	1988
103	IIA	*Solanum lycopersicum*	Brasília/DF	1992
130	IIA	*Solanum tuberosum*	Araucária/PR	1995
158	IIA	*Solanum lycopersicum*	Bragança Paulista/SP	1997
167	IIA	*Solanum lycopersicum*	Cacoal/RO	1997
170	IIA	*Solanum lycopersicum*	Cacoal/RO	1997
201	IIA	*Capsicum annuum*	Brasília/DF	2001
206	IIA	*Solanum lycopersicum*	Goiânia/GO	2001
214	IIA	*Musa* sp.	Manaus/AM	2002
221	IIA	*Solanum lycopersicum*	Brasília/DF	2003
255	IIA	*Solanum tuberosum*	Contenda/PR	2005

Continuing...

Table 1. Continued

Isolate code	Phylotype	Host	City/State	Year
257	IIA	*Solanum lycopersicum*	Santo Amaro/SC	2005
277	IIA	*Solanum lycopersicum*	Caxias do Sul/RS	2007
278	IIA	*Solanum lycopersicum*	Caxias do Sul /RS	2007
285	IIA	*Solanum lycopersicum*	São Benedito/CE	2007
299	IIA	*Solanum lycopersicum*	Brasília/DF	2007
300	IIA	*Solanum lycopersicum*	Araguari/MG	2007
313	IIA	*Solanum lycopersicum*	Fortaleza do Tabocão/TO	2008
351	IIA	*Solanum lycopersicum*	Fortaleza do Tabocão/TO	2008
406	IIA	*Solanum lycopersicum*	Caçador/SC	2010
418	IIA	*Solanum lycopersicum*	Baixo Pinheiral/SC	2010
419	IIA	*Solanum lycopersicum*	Baixo Pinheiral/SC	2010
426	IIA	*Solanum lycopersicum*	Brasília/DF	2010
433	IIA	*Solanum tuberosum*	Santa Juliana/MG	2010
436	IIA	*Solanum lycopersicum*	Seabra/BA	2011
443	IIA	*Solanum lycopersicum*	Brasília/DF	2011
444	IIA	*Solanum lycopersicum*	Brasília/DF	2011
445	IIA	*Solanum lycopersicum*	Brasília/DF	2011
446	IIA	*Solanum lycopersicum*	Brasília/DF	2011
448	IIA	*Capsicum* sp.	Alto Caxixe/ES	2011
453	IIA	*Solanum lycopersicum*	Sumaré/SP	2011
459	IIA	*Solanum lycopersicum*	Reserva/PR	2011
460	IIA	*Solanum lycopersicum*	Borrazópolis/PR	2011
462	IIA	*Solanum lycopersicum*	São Benedito/CE	2011
463	IIA	*Solanum tuberosum*	Camanducaia/MG	2011
464	IIA	*Solanum tuberosum*	Ipuiuna/MG	2011
467	IIA	*Solanum lycopersicum*	Maria da Fé/MG	2010
468	IIA	*Solanum lycopersicum*	Coimbra/MG	2010
469	IIA	*Solanum lycopersicum*	Coimbra/MG	2010
514	IIA	*Solanum lycopersicum*	Santa Maria de Jetibá/ES	2013
525	IIA.2	*Solanum lycopersicum*	Brasília/DF	2013
526	IIA.2	*Solanum lycopersicum*	Brasília/DF	2013
527	IIA.2	*Solanum lycopersicum*	Brasília/DF	2013
528	IIA.2	*Solanum lycopersicum*	Brasília/DF	2013
529	IIA.2	*Solanum lycopersicum*	Brasília/DF	2013
530	IIA.2	*Solanum lycopersicum*	Brasília/DF	2013
531	IIA.2	*Solanum lycopersicum*	Brasília/DF	2013
532	IIA.2	*Solanum lycopersicum*	Brasília/DF	2013
1	IIA.1	*Solanum tuberosum*	Itajai/SC	1987
4	IIA.1	*Solanum tuberosum*	Maria da Fé/MG	1987
8	IIA.1	*Solanum lycopersicum*	Manaus/AM	1987
9	IIA.1	*Solanum tuberosum*	Venda Nova/ES	1987
10	IIA.1	*Solanum lycopersicum*	Elias Fausto/SP	1987
16	IIA.1	*Solanum lycopersicum*	Coari/AM	1987
21	IIA.1	*Solanum lycopersicum*	Camocim de São Felix/PE	1987
37	IIA.1	*Solanum lycopersicum*	Macapá/AP	1988
62	IIA.1	*Solanum lycopersicum*	Petrolina/PE	1990
63	IIA.1	*Solanum lycopersicum*	Petrolina/PE	1990
64	IIA.1	*Solanum lycopersicum*	Petrolina/PE	1990
72	IIA.1	*Solanum tuberosum*	Brasília/DF	1990
86	IIA.1	*Solanum melongena*	Brasília/DF	1992
87	IIA.1	*Solanum melongena*	Brasília/DF	1992
93	IIA.1	*Solanum tuberosum*	Piraí do Sul/PR	1992
119	IIA.1	*Solanum tuberosum*	Ibiá/MG	1992
122	IIA.1	*Solanum lycopersicum*	Brasília, DF	1992
165	IIA.1	*Solanum tuberosum*	Nova Prata/RS	1997

Continuing...

Table 1. Continued

Isolate code	Phylotype	Host	City/State	Year
172	IIA.1	*Solanum lycopersicum*	Planaltina/GO	1997
190	IIA.1	*Capsicum annuum*	Tabatinga/AM	2000
194	IIA.1	*Piper hispidinervum*	Rio Branco/AC	2000
197	IIA.1	*Solanum melongena*	Belém/PA	2000
199	IIA.1	*Capsicum annuum*	Brasília/DF	2000
205	IIA.1	*Solanum tuberosum*	Cristalina/GO	2001
223	IIA.1	*Solanum lycopersicum*	Maracãs/BA	2003
225	IIA.1	*Capsicum annuum*	Brasília/DF	2003
233	IIA.1	*Solanum melongena*	Belém/PA	2003
234	IIA.1	*Solanum melongena*	Belém/PA	2003
235	IIA.1	*Solanum lycopersicum*	Rondon do Pará/PA	2003
236	IIA.1	*Solanum lycopersicum*	Rondon do Pará/PA	2003
243	IIA.1	*Solanum tuberosum*	Saturno/SP	2004
251	IIA.1	*Solanum tuberosum*	Araucária/PR	2005
263	IIA.1	*Solanum tuberosum*	Santo André/SP	2005
318	IIA.1	*Solanum lycopersicum*	Fort. Do Tabocão/TO	2008
333	IIA.1	*Solanum lycopersicum*	Aragominas/TO	2008
342	IIA.1	*Solanum lycopersicum*	Matinha/MA	2008
364	IIA.1	*Solanum lycopersicum*	Itajai/SC	2008
385	IIA.1	*Solanum lycopersicum*	Camocim de São Felix/PE	2009
456	IIA.1	*Solanum lycopersicum*	Sumaré/SP	2011
458	IIA.1	*Solanum lycopersicum*	Sumaré/SP	2011
485	IIA.1	*Solanum lycopersicum*	Faxinal/PR	2011
502	IIA.1	*Solanum lycopersicum*	Ibiapina/CE	2012
503	IIA.1	*Solanum melongena*	Ubajara/CE	2012
W22	IIA	*Musa* sp.		1998
W42	IIA	*Solanum tuberosum*		1979
W301	IIA	*Solanum tuberosum*		2000
50	IIA.1	*Solanum melongena*	Brasília/DF	1982
15	IIA.1	*Solanum melongena*	Macapá/AP	1987
380	IIA.1	*Capsicum annuum*	Camocim de São Felix/PE	2008

Table 2. Polymorphism in *R. solanacearum* inferred for six loci using 189 isolates (a) and to each phylotype (b)

a)

Locus	Sites	Hap	Hd	S	Ks	Ka	CG%	D	D*	F*	K_a/K_s
gdhA	929	46	0.90	64	0.04	3	65.3	0.04	-2.72*	-1.79	0.09
mutS	692	25	0.86	50	0.06	6	70.2	0.84	-1.82	-0.79	0.12
ppsA	460	16	0.83	50	0.09	5	66.7	0.76	-0.52	0.05	0.06
leuS	662	51	0.86	44	0.10	17	64.4	0.64	0.08	0.39	0.18
gyrB	414	40	0.91	40	0.06	8	65.1	0.51	-1.66	-0.88	0.13
fliC	333	37	0.87	31	0.01	22	61.1	0.37	-1.37	-0.77	1.49
Egl	750	36	0.86	47	0.07	4	62.5	0.74	-1.63	-0.39	0.06
Total	4240	251	0.87	326	0.04	10	65.8	0.50	-1.81	-0.81	0.27

b)

Phylotype	Locus	Site	Hap	Hd	S	Ks	Ka	CG%	D	D*	F*	K_a/K_s
I	*gdhA*	929	9	0.42	34	16	2	65.8	-1.32	0.32	-0.26	0.13
I	*mutS*	692	2	0.05	19	3	1	71.1	-2.57*	-4.83*	-4.83*	0.27
I	*ppsA*	460	4	0.17	33	17	1	67.8	-2.65*	-4.85*	-4.86*	0.08
I	*leuS*	661	17	0.78	27	5	7	64.2	-1.23	0.45	-0.11	1.33
I	*gyrB*	414	4	0.17	21	12	3	66.3	-1.94*	0.09	-0.67	0.29
I	*Egl*	750	8	0.24	30	14	3	63.0	-2.47*	-1.32	-1.72	1.29
I	*fliC*	333	5	0.27	14	2	3	59.5	-2.36*	-3.66*	-3.82*	1.45
IIA	*gdhA*	929	27	838	46	15	0.0009	65.0	-1.70	-3.97*	-3.67*	0.06
IIA	*mutS*	692	11	651	19	7	0.0004	70.1	-1.94*	-4.19*	-4.00*	0.05
IIA	*ppsA*	460	9	656	40	19	0.0007	66.4	-2.21*	1.01	-0.31	0.04
IIA	*leuS*	661	21	0.55	38	2	0.0038	64.5	-2.31*	-0.68	-1.59	1.81
IIA	*gyrB*	414	27	0.83	31	45	2	64.8	-0.65	-3.33*	-2.74*	0.04
IIA	*Egl*	750	21	0.76	30	32	0.0121	67.2	-2.84	-3.56*	-3.42	1.04
IIA	*fliC*	333	22	719	24	7	0.0103	61.3	-1.31	-0.51	-0.98	1.46
IIB	*gdhA*	929	13	688	31	9	0.0006	65.3	-2.07*	-3.11*	-3.26*	0.06
IIB	*mutS*	616	12	789	25	21	0.0013	69.9	-0.87	-2.48	-2.26	0.06
IIB	*ppsA*	460	6	485	16	12	0.0004	66.6	-1.68	-2.07	-2.30	0.04
IIB	*leuS*	661	23	918	22	2	0.0091	642	-0.09	0.35	0.23	3.64
IIB	*gyrB*	414	12	768	23	28	0.0005	65.1	-1.26	-1.75	-1.87	0.02
IIB	*Egl*	750	22	0.83	18	6	0.0032	68.6	-1.34	-1.65	-1.62	1.02
IIB	*fliC*	333	12	719	19	8	0.0095	61.5	-0.88	0.53	0.03	1.09

Number of analyzed sites (Sites); number of haplotypes (Hap); haplotype diversity (Hd, Nei 1987); number of polymorphic sites (S); Number of synonymous substitutions per synonymous site (Ks) and number of nonsynonymous substitutions per nonsynonymous site (Ka) substitution rate calculated by Nei and Gojobori (1986) method; GC content (GC%); neutrality tests of Tajima (1989) and Fu and Li (1993) (Tajima's D andFu and Li's D* and F*) and associated P-value: * *p<0.05**

Table 3. Estimates of mutation (θ) and recombination (ρ) rates using Ldhat 2.2 program

Gene	□				□				□/□			
	I	IIA	IIB	Whole set	I	IIA	IIB	Whole set	I	IIA	IIB	Whole set
egl	0	3	2	2.62	4.68	5.89	5.23	6.13	0	0.50	0.38	0.42
flic	0	10	0	3.06	3.67	4.68	4.50	5.50	0	2.13	0	0.55
gyrB	0	1	3	1.02	5.17	6.04	4.91	6.87	0	0.16	0.61	0.14
gdha	3	1	1	1.02	8.38	8.97	7.05	11.00	0.35	0.11	0.14	0.09
leus	2	0	4	2.04	7.15	7.21	5.76	9.45	0.27	0	0.69	0.21
muts	0	1	0	1.02	0.74	7.21	7.47	8.59	0	0.13	0	0.11
ppsa	0	0	0	0	0.74	7.80	4.06	8597	0	0	0	0

Table 4. Mutation rate per gene and per year inferred in BEAST for each gene in each subdivision within Brazilian *R. solanacearum*

Gene	Mutation rate per year		
	I	IIA	IIB
gyrB	4.3x10-5	2.1x10-5	1.7x10-6
fliC	8.5x10-5	9.2x10-5	0.9x10-5
Gdha	4.3x10-5	0.7x10-5	3.6x10-5
mutS	5.2x10-5	4.1x10-5	7.7x10-6
leuS	3.5x10-5	5.1x10-5	5.8x10-5
Ppsa	2.4x10-5	5.0x10-5	4.6x10-5
egl	6.4x10-5	8.5x10-4	1.5x10-5

Supporting information

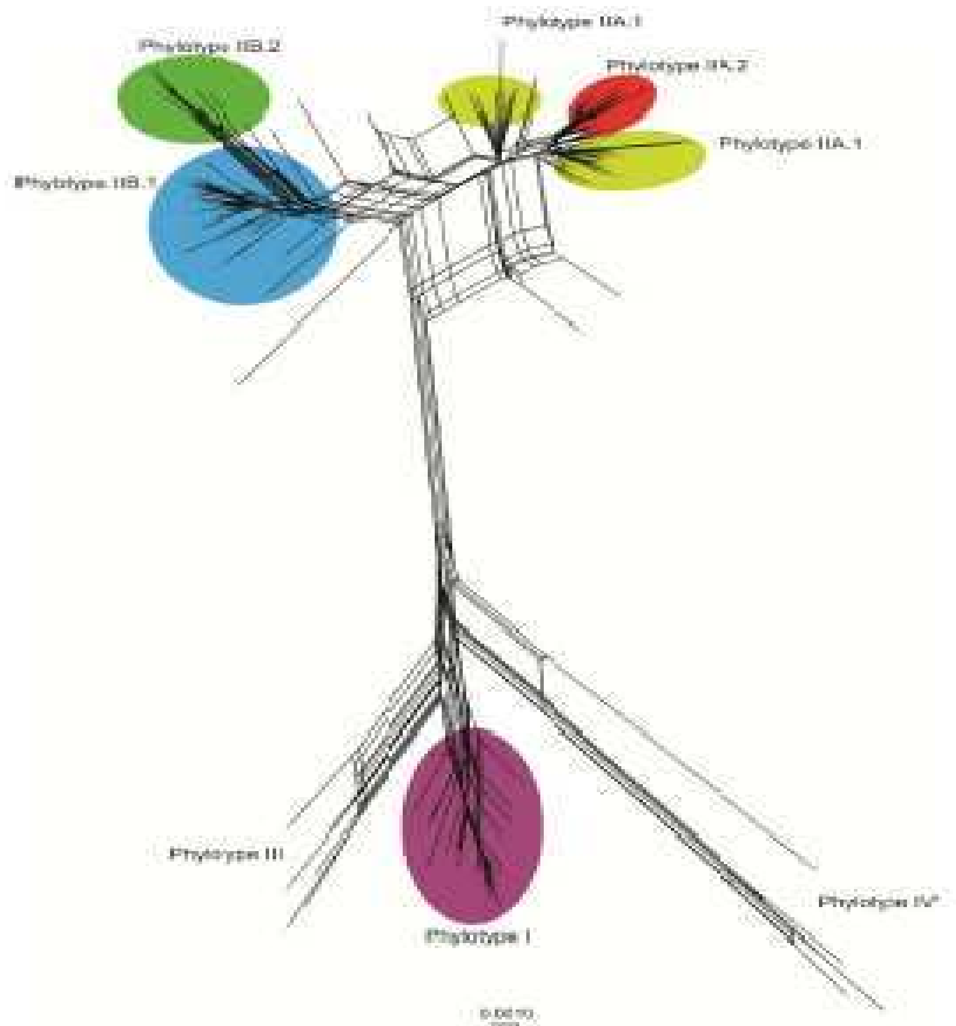
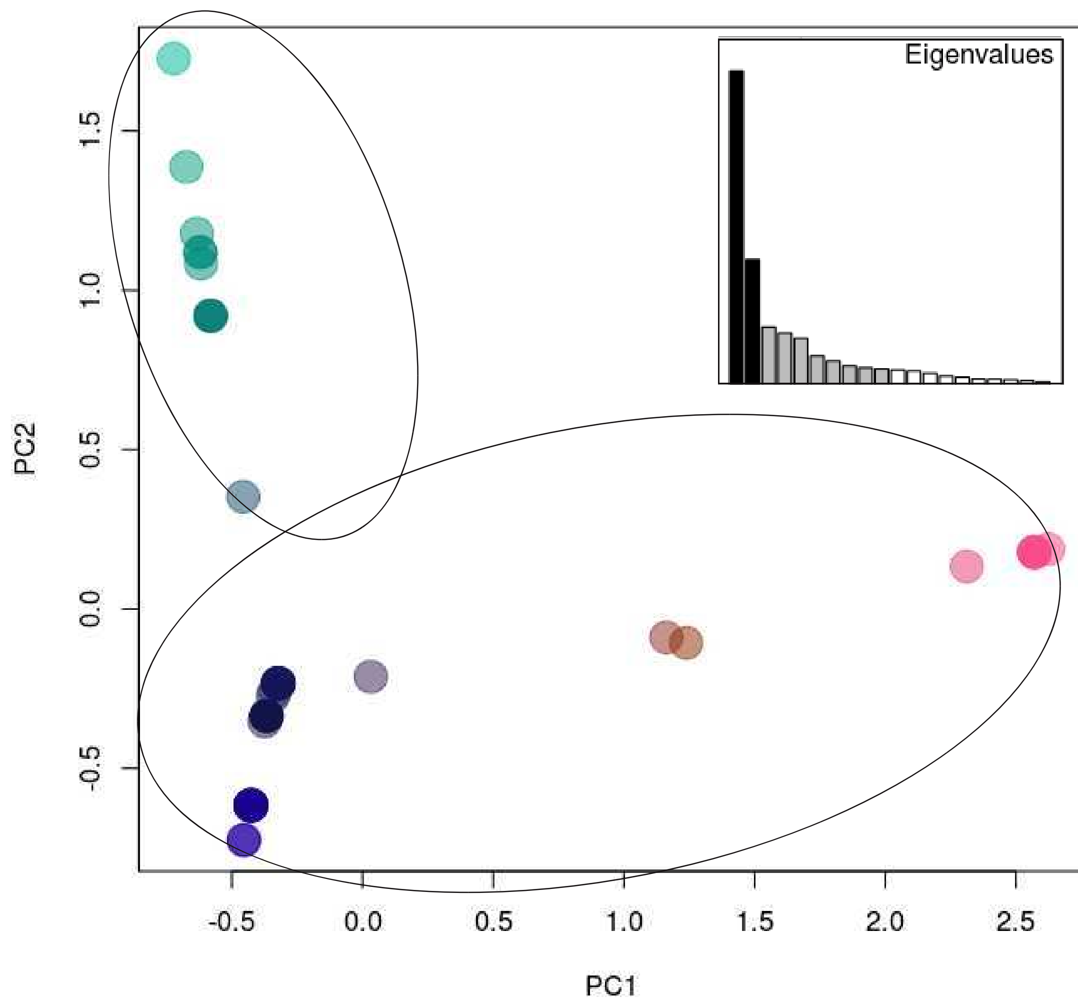


Figure S1. Neighbor net analysis revealed a complex network with recombination (shown as parallelograms) between isolates of *R. solanacearum.*

Figure S2. Scatter-plot of the phylotype distribution of phylotype IIB of *Ralstonia solanacearum* isolates into two genetic subgroups (1) Phylotype IIB1 with green color and (2) Phylotype IIB2 with blue color.



GENERAL CONCLUSION

The Brazilian population of *Ralstonia solanacearum* has high genetic variability and no evidence of genetic groups geographically separated;

Isolates were detected in biovars 1, 2 and 3, the phylotypes I and II and sequevares 1, 4, 18, 27, 28, 41 and 50;

For new sequevar were described: 54, 55, 56 and 57;

There is a greater genetic differentiation based on host (potato and tomato) than by region;

There is strong evidence supporting the origin of phyilotipo II in Brazil, additionally, the phylotipo IIA originated from filotipo IIB;

Apparently isolated of phylotipo I were recently introduced in Brazil;and migration, selection, mutation and recombination are evolutionary mechanisms that affect the population of *Ralstonia solanacearum* in Brazil, however the mutation is more effective in generating variability than recombination.