



HAL
open science

A minisatellite-based MLVA for deciphering the global epidemiology of the bacterial cassava pathogen *Xanthomonas phaseoli* pv. *manihotis*

Leidy Rache, Laurence Blondin, Paula Diaz Tatis, Carolina Flores, Andrea Camargo, Moussa Kante, Issa Wonni, Camilo López, Boris Szurek, Stephane Dupas, et al.

► To cite this version:

Leidy Rache, Laurence Blondin, Paula Diaz Tatis, Carolina Flores, Andrea Camargo, et al.. A minisatellite-based MLVA for deciphering the global epidemiology of the bacterial cassava pathogen *Xanthomonas phaseoli* pv. *manihotis*. PLoS ONE, 2023, 18 (5), pp.e0285491. 10.1371/journal.pone.0285491 . hal-04292499

HAL Id: hal-04292499

<https://hal.science/hal-04292499>

Submitted on 17 Nov 2023

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

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

RESEARCH ARTICLE

A minisatellite-based MLVA for deciphering the global epidemiology of the bacterial cassava pathogen *Xanthomonas phaseoli* pv. *manihotis*

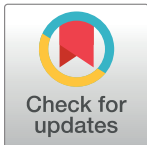
Leidy Rache^{1‡}, Laurence Blondin^{2,3}, Paula Diaz Tatis⁴, Carolina Flores³, Andrea Camargo⁴, Moussa Kante³, Issa Wonni⁵, Camilo López⁶, Boris Szurek³, Stephane Dupas¹, Olivier Pruvost⁷, Ralf Koebnik^{3*}, Silvia Restrepo¹, Adriana Bernal^{1*}, Christian Vernière^{3,7†}

1 Department of Biological Sciences, Universidad de los Andes, Bogotá D.C., Colombia, **2** CIRAD, UMR PHIM, Montpellier, France, **3** Plant Health Institute of Montpellier (PHIM), Université Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France, **4** Facultad de Ciencias, Universidad Antonio Nariño, Bogotá, Colombia, **5** INERA, LMI Pathobios, Bobo Dioulasso, Burkina Faso, **6** Departamento de Biología, Universidad Nacional de Colombia, Bogotá, Colombia, **7** CIRAD, UMR PVBMT, Saint-Pierre, La Réunion, France

† Deceased.

* Current address: Escuela de Ciencias Biológicas, Universidad Pedagógica y Tecnológica de Colombia, Tunja, Colombia

* ralf.koebnik@ird.fr (RK); abernal@uniandes.edu.co (AB)



OPEN ACCESS

Citation: Rache L, Blondin L, Diaz Tatis P, Flores C, Camargo A, Kante M, et al. (2023) A minisatellite-based MLVA for deciphering the global epidemiology of the bacterial cassava pathogen *Xanthomonas phaseoli* pv. *manihotis*. PLoS ONE 18(5): e0285491. <https://doi.org/10.1371/journal.pone.0285491>

Editor: Tushar Shaw, Faculty of Life and Allied Health Sciences, Ramaiah University of Applied Sciences, INDIA

Received: December 11, 2022

Accepted: April 25, 2023

Published: May 11, 2023

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0285491>

Copyright: © 2023 Rache et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Cassava Bacterial Blight (CBB) is a destructive disease widely distributed in the different areas where this crop is grown. Populations studies have been performed at local and national scales revealing a geographical genetic structure with temporal variations. A global epidemiology analysis of its causal agent *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) is needed to better understand the expansion of the disease for improving the monitoring of CBB. We targeted new tandem repeat (TR) loci with large repeat units, i.e. minisatellites, that we multiplexed in a scheme of Multi-Locus Variable number of TR Analysis (MLVA-8). This genotyping scheme separated 31 multilocus haplotypes in three clusters of single-locus variants and a singleton within a worldwide collection of 93 *Xpm* strains isolated over a period of fifty years. The major MLVA-8 cluster 1 grouped strains originating from all countries, except the unique Chinese strain. On the contrary, all the *Xpm* strains genotyped using the previously developed MLVA-14 microsatellite scheme were separated as unique haplotypes. We further propose an MLVA-12 scheme which takes advantage of combining TR loci with different mutation rates: the eight minisatellites and four faster evolving microsatellite markers, for global epidemiological surveillance. This MLVA-12 scheme identified 78 haplotypes and separated most of the strains in groups of double-locus variants (DLV) supporting some phylogenetic relationships. DLV groups were subdivided into closely related clusters of strains most often sharing the same geographical origin and isolated over a short period, supporting epidemiological relationships. The main MLVA-12 DLV group#1 was composed by strains from South America and all the African strains. The MLVA-12

Data Availability Statement: All relevant data are within the paper and its [supporting information files](#).

Funding: We thank the Faculty of Sciences from Universidad de los Andes-Bogotá, Colombia (INV-2021-128-2283), the UMR Interactions Plantes Microorganismes Environnement, and the Agropolis Foundation (project PAIX, grant no. 1403-073), Montpellier, France, for financial support. The Ecos Nord programme supported inter-laboratory mobility (grant no. C15A01). Leidy Rache was supported by Colciencias with a doctoral fellowship (call No. 528, 2011). Olivier Pruvost acknowledges the European Union (ERDF contract GURDT I2016-1731-0006632) and Réunion regional council for support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript

Competing interests: The authors have declared that no competing interests exist.

scheme combining both minisatellite and microsatellite loci with different discriminatory power is expected to increase the accuracy of the phylogenetic signal and to minimize the homoplasy effects. Further investigation of the global epidemiology of *Xpm* will be helpful for a better control of CBB worldwide.

Introduction

Molecular genotyping of bacterial pathogens allows a comprehensive view of their epidemiology and evolution. The ability to differentiate genotypes at different spatial and temporal scales can be informative of the dispersion and micro or macroevolution of bacterial pathogens [1–3]. At a local scale, this approach can shed light on the microevolutionary processes within populations, and track back the spread of bacterial isolates and the associated pathways. At a long-term or global scale, it can for instance unravel the invasion routes and characterize some bacterial lineages associated to pathological or adaptative traits for epidemiological surveillance. Genotyping methods based on several molecular markers have been developed to analyze the genetic diversity of pathogenic bacteria. The type of marker should be selected according to the research question to be answered because each marker has different mutation rates i.e. different speed of evolution and stability [2–6]. Markers with a desired molecular clock can be selected according to the evolutionary scale under study.

Global epidemiology studies or international surveillance of pathogens require tools that can produce unambiguous data and are easily transferrable among different laboratories. Multilocus sequence typing (MLST) is a nucleotide sequence-based approach using markers that are conserved, stable, with relatively low evolution rates. It enables international comparison of isolates and global epidemiology [7]. However, some bacterial pathogens, the so-called genetically monomorphic bacteria [8], do not exhibit enough variability in their housekeeping genes making this approach insufficiently resolute [8]. Various subtyping methods emerged with the advent of whole genome sequencing. In situations where large numbers of high-quality draft genomes can be obtained, single nucleotide polymorphism (SNP) typing is a powerful approach for epidemiological surveillance and phylogenetic analyses [8–11]. Technically simple and still robust systems are also necessary for routine epidemiological investigations in situations where an easy access to large sets of compete genomes is impossible to achieve. In such situations, Tandem Repeats (TRs) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) have been mainly applied as typing tools with regard to their reliability, discriminatory power, transferability, ease of application and cost [12,13].

TRs are short motifs of DNA fragments repeated in tandem that can vary in number due to addition or deletion of motifs. Multiple loci variable number tandem repeat analysis (MLVA) targets multiple TR loci which allows a high discriminatory power if a sufficiently large number of TRs is targeted. Two classes have been proposed to separate these informative molecular markers: microsatellites with a repeat unit size < 10 bp and minisatellites with repeat units ≥ 10 bp [14,15]. The length of TR arrays is typically determined by capillary electrophoresis from multiplex involving fluorochrome-labeled primers but it can be scored unambiguously and from agarose gels especially for large repeat motifs making them applicable at a lower throughput in laboratories where a capillary electrophoresis genotyper is not available. Several microsatellite schemes have been developed for epidemiological analyses of plant pathogenic xanthomonads [16]. Minisatellite-based typing schemes targeting large repeat motifs

were also proposed as a portable approach for the global analysis of the molecular epidemiology of bacterial strains pathogenic to human or animals [15,17,18] or to plant crops [19].

The wide range of mutation rates offered by the different TR loci classes make them useful for examining genetic patterns at several evolutionary scales. While microsatellites are most suited for local epidemiology studies, minisatellites have been mostly used for global epidemiology, *i.e.* analysis of non-epidemiologically related isolates [17,19].

Mutational processes observed in VNTRs are mainly driven by DNA replication slippage (and to a lesser extent recombination) and the length or the number of the tandem repeat and the sequence of the repeat unit are factors that drive the mutation rate of TRs [14,20,21].

Cassava Bacteria Blight (CBB) is a very important disease caused by *Xanthomonas phaseoli* pv. *manihotis* (*Xpm*) [22,23]. This disease was first reported in Brazil in 1912 [24], and later noticed in Colombia in 1970 [25], Venezuela in 1971 and African and Asian countries since 1972 [26,27]. CBB might have been present in these countries for many years, but few reports described a bacterial leaf spot on cassava before the seventies. More recently, the disease was reported in South Pacific [28]. Population studies of the pathogen have been conducted in several Latin American countries, including Brazil, Venezuela [29,30] and a continuous monitoring of the populations has been performed in Colombia [31–36]. *Xpm* is transmitted at small scale by rain and wind within production areas. At larger scales between production areas and between different countries, contaminated plant material, *i.e.* stem cuttings, are responsible for bacterial dispersal and the carry-over to the following growing season [24].

Studies of the genetic diversity in *Xpm* populations have been mainly performed at a local level using Restriction Fragment Length Polymorphisms (RFLP) or Amplified Fragment Length Polymorphisms (AFLP), as well as whole genome sequencing [37]. In Colombia *Xpm* populations were genetically diverse and geographically structured [34,38]. Genetic variation was observed within fields with high levels of genotypic diversity and haplotype frequencies significantly differing over the years [33,38]. More recent studies in different regions of Colombia confirmed the high genetic diversity and the local and geographic structure of *Xpm* populations in these regions [35,36]. A geographic structure of *Xpm* populations was also described in Venezuela at a regional scale [29]. Similarly, a high population diversity has been observed in African countries [39].

A few large-scale analyses of the genetic diversity of *Xpm* populations have been performed. Analysis of a worldwide collection using RFLP led to the identification of clusters grouping strains that originated from different countries [26]. This study suggested that African strains belong to a single subclade of South American strains, whereas Asian strains belong to different groups. A recent study based on single nucleotide polymorphism analyses supports gene flow among *Xpm* strains representative of the worldwide diversity especially between Brazil and Colombia and between the South-American and the African continents [37]. A common ancestor among Brazil, Colombia and African strains was suggested as previously suspected [26].

Herein, we built on the versatility of TR markers in order to develop improved MLVA schemes with desirable characteristics for outbreak investigation or epidemiological surveillance of *Xpm*. An MLVA-scheme targeting microsatellites was recently proposed and evaluated on local populations [40]. Our aims in this study were (i) to develop new TR markers with larger repeat size, (ii) to evaluate the discriminatory power of these minisatellites on a worldwide collection of *Xpm* strains in comparison with the MLVA scheme based on microsatellites and (iii) to propose an MLVA scheme adapted for regional or international surveillance of *Xpm*.

Materials and methods

Xpm strain collection and DNA extraction

A collection of 93 *Xpm* strains isolated over several years from different cassava producing countries from South America (Colombia, Venezuela, Brazil, Argentina), Africa (Burkina Faso, Mali, Togo), Asia (Vietnam, and China) and New Zealand, (S1 Table) was used for genotyping. The genomic DNA was extracted from 24 hour-old cultures on an LPGA culture medium (Yeast extract 5 g L⁻¹, Peptone 5 g L⁻¹, glucose 5 g L⁻¹, agar 15 g L⁻¹, pH 7.2). DNA extraction was performed using the Wizard® Genomic kit (Promega, Madison, WI, USA). The amount and purity of DNA was checked using NanoDrop technology (ThermoFisher Scientific, Illkirch, France).

Genotyping using a new MLVA scheme targeting minisatellites

Twelve *Xpm* genomes, previously sequenced with Illumina [37], were selected, based on degree of fragmentation (least fragmented were included) and number of contigs (lower number of contigs were included). Eight strains from Brazil, two from Colombia and two from Uganda [40], were screened for VNTR minisatellite loci using the *Xanthomonas* utility website (<http://bioinfo-web.mpl.ird.fr/xantho/utills/>), setting the following parameters: 30 and 1000 bp for the total length, the unit length of the repeat in a range of 10 to 100 bp, similarity of repeats within the arrays between 80 and 100%.

The primers were designed over flanking regions up- and downstream of the repeat region as previously described [40]. Primers were tested by individual PCR and multiplex PCR using DNA from a sub-collection of *Xpm* strains: CIO1, CIO151, COL303, CFBP1851 (Colombia), VEN130 (Venezuela), SCI (Ivory Coast), D2-3 (Burkina Faso), Mali45, Mali30 (Mali). These strains were only used to test the quality and design of the primers in agarose gels.

Multiplex PCR was performed to amplify three loci in each multiplex using the QIAGEN® Multiplex PCR kit (Qiagen, Courtaboeuf, France) according to the manufacturers' instructions with 40 to 50 ng of target in a total volume of 10 µl. The multiplex PCR conditions were as follows: pre-denaturation at 95°C for 5 minutes, 30 cycles of denaturation at 95°C for 30 seconds, annealing with a specific temperature per pool for 40 seconds (Table 1), extension at 72°C for 2 minutes; and a final extension at 72°C for 5 minutes. The amplicons were visualized by agarose gel electrophoresis (2%).

After the standardization of the multiplex PCR, the reverse primer in each primer pair was labelled using different fluorescent dyes to separate the PCR products by capillary electrophoresis (Table 1). PCR products were diluted 20 times (estimated from preliminary tests), and 1 µl of the dilution was mixed with HiDi formamide (9 µl) (Applied Biosystems, CA, USA), and GeneScan™ 600 LIZ® dye size standard v2.0 (0.5 µl) (Applied Biosystems, CA, USA). The detection was performed in an ABI 3500 Genetic Analyzer, and amplicon sizes were determined using GeneMapper® (Applied Biosystems) software version 4.1. The strain CIO151 [41] was used as a reference strain in each test.

Xpm genotyping using MLVA-14 targeting microsatellites

The strain collection was also genotyped using a MLVA scheme targeting 14 microsatellites, referred to as MLVA-14. This scheme is derived from MLVA-15 [40], but a single microsatellite, *Xpm* 2–20 (VNTR *Xpm* 1–20 nomenclature from [40]), was taken out because it did not amplify on a few strains from Venezuela or Africa (unpublished data). One primer from each pair in the PCR mix was labelled with one of the four different fluorescent dyes (6-FAM, NED,

Table 1. PCR primers sequences and fluorophore dyes of the minisatellite loci used in the Multiplex PCR with the corresponding annealing temperature (AT).

Pool	VNTR locus	Forward sequence	Reverse sequence	AT
I	Xpm 2-22F	CAATGGACCGCCTCGCTGGC	VIC-GTGGGTGCCTGGTCGCAACG	63
	Xpm 2-18F	TTTCGGGGTGGCGGCATTGT	6-FAM-GCGATTGATCGCCGGGTCGT	63
	Xpm 2-35F	TGACCTGAAGGGCGAGGGC	PET-CGCCGGCAGCAACTGTCCAC	63
II	Xpm 2-5F	AACGCAGCGGGTGTGGTTGC	PET-GCAAGCAGCGCAGAAGGCGA	64
	Xpm 2-33F	GCGGCTTCGTTCAGTACCCTC	VIC-CGCAATGCTCAAATCGCCCT	64
	Xpm 2-29F	GTCGGCTGTGGTGGCGGAG	6-FAM-TTCCGGCAACTGGCAGGCG	64
III	Xpm 2-3F	ACCGTGCCCATTCGGCACC	PET-CCATTACCACCGCTGCGGGC	62
	Xpm 2-20F	CGGTATGGGGCCCCGAAAGC	VIC-CTCCGTGCACACCGGCGACT	62
	Xpm 2-23F	CGACGTGCGTGCGTAGGCGA	6-FAM-GCGTGAGGCAAACATCGGCG	62

<https://doi.org/10.1371/journal.pone.0285491.t001>

PET, and VIC, Applied Biosystems) (S2 Table) and detected by capillary electrophoresis in an ABI 3500 Genetic Analyzer as indicated above.

Data scoring

Fragment sizes for each VNTR locus were estimated using GENEMAPER v 4.0 (Applied Biosystems). When necessary, a truncated number of repeats was rounded up to the nearest integer [12]. The discriminatory power of each VNTR locus was calculated based on the Hunter and Gaston discriminatory index (HGDI) [42]. Numbers of alleles and Nei's unbiased estimates of diversity were estimated using ARLEQUIN software [43]. Haplotype networks were represented by minimum spanning trees using the algorithm combining global optimal eBURST (goeBURST) and Euclidian distances in the software PHYLOViZ 2.0 [44]. The minimum spanning trees display the relationships between clonal complexes defined as groups of single locus variants (SLVs). The population structure of *Xpm* was assessed by discriminant analysis of principal components (DAPC) of MLVA-12 data using the adegenet V.2.1.1. R package [45,46]. DAPC is free of any assumption linked to a population genetic model (e.g., Hardy-Weinberg equilibrium or linkage equilibrium), and, thus, is suited for analysis of datasets produced from predominantly clonal bacteria.

Pathogenicity tests

Pathogenicity tests were performed with strains representative of different haplotypes and clonal complexes selected from the minimum spanning tree based on the MLVA-8 minisatellites scheme. The strains selected were CIAT1205, CIAT1135, UA1591, CIAT1241, UA556, UA2146, CIAT1202 and CIO151. Two month-old plants of the cassava varieties CM523-7, COL1505, NGA11, CM6438-14 and cv.60444 were inoculated. Ten μ l of *Xpm* cell suspension (optical density at 600nm (OD_{600nm}) = 0,02, $\sim 10^7$ colony forming units (CFU ml⁻¹) were inoculated in the stem. A scale of symptoms from 0 to 5 (0 = no symptoms, 1 = necrosis at the inoculation point, 2 = stem exudates, 3 = one or two wilted leaves, 4 = more than three wilted leaves, and 5 = plant death) was used. Symptom development was monitored at 7, 14, 21, and 28 days after inoculation. Disease progression AUDPC (area under disease progress curve) [47] was calculated for each of the five replicates. The determination of the resistance level of the cassava varieties was based on four categories, as defined previously [36,48]. A variety was considered resistant when Σ AUDPC \leq 39, moderately resistant: $39 < \Sigma$ AUDPC \leq 44), moderately susceptible: $44 < \Sigma$ AUDPC $<$ 49 and susceptible when Σ AUDPC \geq 49.

Results

Selection of eight new minisatellite loci

After the screening of 12 *Xpm* genomes, [49] VNTR loci met the criteria of the search. Of these candidates, only ten loci were selected because their flanking regions were almost fully conserved, and their sequences were identical for each locus in all the strains analyzed. Five TR loci did not show any polymorphism for these 12 genomes. For the polymorphic loci the HGDI was relatively high varying between 0.167 and 0.849 with an allele number ranging from two to six, respectively (Table 2).

The total length of most of the loci including the primer regions ranged between 197 to 490 bp. The unit length of the tandem repeat ranged between 10 and 26 bp. Individual bands were observed for each primer and in PCR multiplex (Fig 1). However, for loci Xpm2-29 and Xpm2-2, the size of amplicons exceeded the range size covered by capillary electrophoresis. The sequencing analysis of these two loci showed large insertions so they were removed from the typing scheme. The final minisatellite scheme was formed by eight VNTR loci and called hereafter MLVA-8 (Table 1). This scheme had three perfect loci (Xpm 2–20, Xpm 2–23, and Xpm 2–33) and five imperfect loci (Table 2).

MLVA-8 genotyping of a world collection of *Xpm* strains

Thirty-one multilocus haplotypes were obtained from the worldwide collection of 93 *Xpm* strains using the MLVA-8 scheme. All the TR loci were polymorphic except Xpm2-5, which we hypothesize was fixated by natural selection. Polymorphic loci produced between two (loci Xpm2-33 and 2–35) and eight alleles (Xpm2-18) with Nei's gene diversity varying from 0.108 to 0.831 (S3 Table). TR markers along the evolutionary path of the minimum spanning tree were analyzed by assessing the number of repeats involved in the polymorphism of recently diverging haplotypes, i.e. SLVs. Multiple-repeat variations occurred only for loci Xpm 2–3 and

Table 2. Characteristics of the minisatellite loci selected in this study from 12 genomes of *Xanthomonas phaseoli* pv. *manihotis*.

VNTR Locus	Motif sequence	FS ^a	Unit length (bp)	# Alleles	Allelic range	HGDI ^b	Other, previously used nomenclature ^c of VNTR locus
Xpm 2–2	GGTTCTAGCTTTTA, GGTTCTGCTTCTA, GGTTCTTGCTTCTC	9	14	3	4–6	0.639	
Xpm 2–3	CGGCCCCGACGT(1), CGGCCCCGCTCC(1)	10	12	1	4	0	
Xpm 2–5	GTTGGCCGGTTCGGTTA(1) GTTGGCCGGTTCGGTAG(1)	12	17	1	3	0	
Xpm 2–18	CACCGCCACTACG(1) CACCACCACAACG(1) CACCGCCACAACG (9)	12	13	6	5–11	0.849	XaG2_52
Xpm 2–20	CAACATCCACAG(2)	12	12	2	3–5	0.167	Xpm 1–20
Xpm 2–22	GCAAGCGCGGTGCAACCACGTA(1) GCAAGCGCAGTGCAACCACGCG(1) GCAAGCGCAGTGCAAGCGCA(1)	12	22	2	3–4	0.303	
Xpm 2–23	AGATCGAGACACGC(2)	12	14	4	4–7	0.636	
Xpm 2–29	CGCCGGCACCTG(2)	10	12	1	6	0	
Xpm 2–33	GAGGCTTGCGTGTCTTGGTGTGAG (2)	9	26	1	3	0	XaG2_109
Xpm 2–35	GCTGCCGGCA(1) GCCGCTGCA(1) GGTCCGCAC(1) CTGCAGGCGG(1) CTGCCGGCAG (1)	10	10	1	4	0	Xpm 1–35

^a FS: Flanking sequences (number of genomes out of the 12 for whose flanking sequences were detected).

^b HGDI: Hunter-Gaston discriminatory index.

^c Nomenclature from Trujillo et al. [35] and/or from Arrieta-Ortiz et al. [41] and/or Rache et al. [40].

<https://doi.org/10.1371/journal.pone.0285491.t002>

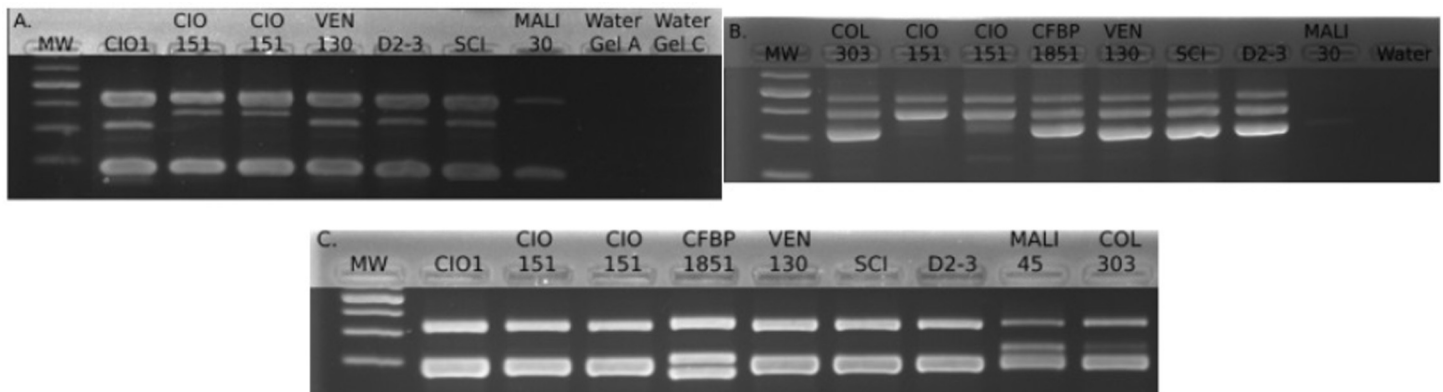


Fig 1. Visualization of multiplex PCR amplicons on agarose gels. MW: Molecular marker. A. Pool I, loci Xpm 2–22, Xpm 2–18 and Xpm 2–35. B. Pool II, loci Xpm 2–33, Xpm 2–5 and Xpm 2–29. C. Pool III loci Xpm 2–3, Xpm 2–20 and Xpm 2–23.

<https://doi.org/10.1371/journal.pone.0285491.g001>

Xpm 2–20 with 50% of double- or triple-repeat variations and 33% of double repeat variations, respectively. Changes consisting of a single-repeat variation represented 85.2% of the SLVs. Our data suggest that these TR minisatellites mainly evolve following a generalized stepwise mutation model (or two-phase model) [48].

Eighteen haplotypes out of 31 originated from a unique country. Thirteen haplotypes were shared by strains originating from different countries, six of which were from South American countries and seven from different continents. All the African strains shared haplotypes with South American strains except MLVA-8 haplotype #5 composed only of Burkinabe strains. The minimum spanning tree separated the haplotypes in three clonal complexes (CC, *i.e.* clusters of single locus variants), and one singleton, *i.e.* haplotypes with no SLVs (Fig 2). The main MLVA-8 CC1 grouped 24 haplotypes including all the African strains and strains from the other countries included in the analysis, except a single strain from China. They were isolated from 1966 to 2016. Three MLVA-8 haplotypes were recovered over a period of more than 30 years (# 1, 7 and 16) (S1 Table and Fig 2). All haplotypes originating from Mali or from Burkina-Faso were SLVs (MLVA haplotypes #2 and 4 differ only by Xpm 2–18). The MLVA-8 CC2 ($n = 4$ haplotypes), a triple locus variant of CC1, was composed by the Chinese strain and strains from Brazil and New Zealand that were all isolated between 1974 and 1996. The third CC (2 haplotypes) and the singleton, haplotype #21, grouped strains from Colombia and New Zealand isolated between 1966 and 1974. Strains from China and Vietnam did not share any haplotypes with other strains.

MLVA-14 genotyping of a world collection of *Xpm* strains

We compared the discriminatory power of MLVA-8 with the previously reported MLVA-14 microsatellite scheme. MLVA-14 scheme was able to distinguish 89 haplotypes out of the 93 *Xpm* strains. All loci were highly polymorphic with a mean allelic richness of $A = 13.5$ leading to high discriminatory index (HGDI = 0.999) and Nei's gene diversity index ($H_E = 0.863$), which were much higher than those estimated from the MLVA-8 dataset (Tables 3 and S4). Strains sharing the same haplotype originated from the same country and, when the information was available, they were isolated the same year in the same locality (S1 Table).

The minimum spanning tree describing the relationships between the 89 haplotypes produced ten small clonal complexes grouping only two or three SLVs (S1 Fig). All these CCs grouped epidemiologically related strains from the same region isolated the same year except

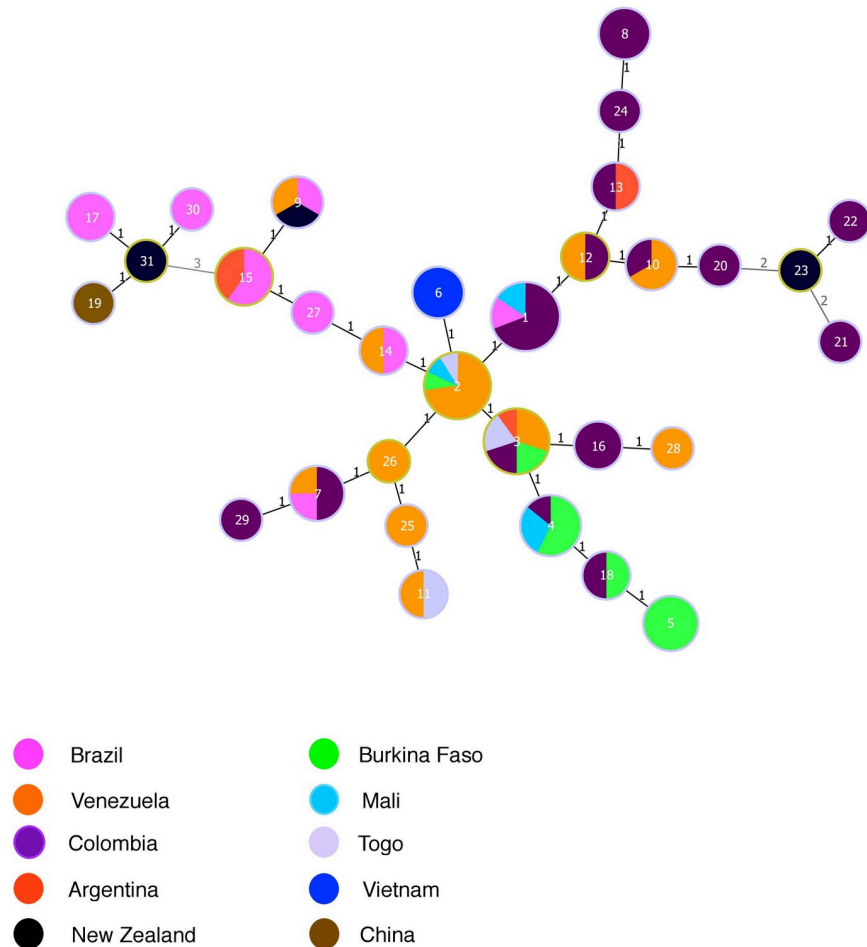


Fig 2. Minimum spanning tree displaying relationships between haplotypes using the MLVA-8 minisatellite scheme. Colors indicate the geographical origin of the haplotype, and the circle size indicates the number of strains of each haplotype. Numbers indicate the number of locus variants between haplotypes.

<https://doi.org/10.1371/journal.pone.0285491.g002>

for the two MLVA-14 haplotypes #3 and 5 (CC7) from two different Colombian regions and isolated during two consecutive years (S1 Table). Other closely related haplotypes (*i.e.* DLVs) also originated from the same country with the exception of two strains from New Zealand isolated in 1966 (haplotypes #44 and 69) which are DLVs of strains from Venezuela and Colombia isolated in the beginning of the seventies, respectively (S2 Fig and S1 Table). Almost all the other haplotypes differed from each other by at least four TR loci but mostly gathered by

Table 3. Estimators of diversity for each MLVA scheme from the *Xanthomonas phaseoli* pv. *manihotis* world collection (n = 93).

Scheme	Polymorphic loci	MLGs ^a	Mean allelic richness A	HGDI ^b	H _T ^c	Simpson' index
MLVA-8	7/8	31	3.375	0.944	0.335	0.937
MLVA-12	11/12	78	4.750	0.995	0.502	0.985
MLVA-14	14/14	89	13.5	0.999	0.867	0.988

^a MLGs, number of multilocus genotypes.

^b HGDI Hunter and Gaston discriminatory index.

^c H_T Nei's total gene diversity.

<https://doi.org/10.1371/journal.pone.0285491.t003>

country or by continent. The closest African and South American strains, from Venezuela or Colombia, were separated by variations at nine TR loci.

A combined MLVA-12 scheme proposed for large scale epidemiology

Four VNTRs (2–31, 2–6, 2–7, 2–38) among the least discriminant of the MLVA-14 scheme were combined to the eight VNTRs (2–22, 2–18, 2–35, 2–20, 2–3, 2–23, 2–5, 2–33) of the MLVA-8 scheme to slightly increase its discriminatory power. The mean allelic richness and HGDI produced by the MLVA-12 dataset were estimated to 4.75 and 0.995, respectively (Table 3). Seventy-eight MLVA-12 haplotypes out of 93 strains were separated into 16 CCs ($n = 60$ strains) composed of two to seven haplotypes and 26 singletons ($n = 33$ strains) (S2 Fig). Nine CCs contained strains originating from the same country among which six consisted of strains isolated the same year in the same region probably corresponding to the same outbreak. Five CCs grouped strains from the same continent either South America (CCs 6, 11 and 16, each with strains from two different countries) or Africa (CCs 1 and 2, both with strains originated from Mali, Burkina Faso and Togo). Most of these CCs are tightly related as they differ by variations in two loci. We used DAPC to delineate seven clusters in our dataset (Fig 3). The assignment probability to clusters was > 0.99 for all strains in the dataset. Interestingly, the large polymorphism revealed for strains from South America split them in the seven clusters. Strains originating from Argentina ($n = 4$), Brazil ($n = 12$), Colombia ($n = 28$) and Venezuela ($n = 21$) were assigned to three (Argentina) or four (other countries) distinct clusters. Cluster #1 included strains from South America (Argentina, Brazil and Colombia) together with strains from Viet Nam. Cluster #2 mostly included strains from South America (Argentina, Brazil and Venezuela) together with a single strain from New Zealand. All African strains gathered in cluster #3 together with some strains from Colombia and Venezuela. Clusters #4 and 5 solely included strains from Colombia and Venezuela, respectively. Cluster #6 primarily included strains from South America (Brazil and Colombia) together with a few strains from China and New Zealand. Cluster #7 included strains from all countries in South America (Argentina, Brazil, Colombia and Venezuela). The structure revealed by DAPC was overall consistent with the produced MST (Figs 3 and S2).

Pathogenicity patterns are not associated to genetic groups

The AUDPC values were used to determine if cassava varieties are resistant ($\Sigma\text{AUPDC} \leq 39$), moderately resistant ($39 < \Sigma\text{AUDPC} < 44$), moderately susceptible ($44 < \Sigma\text{AUDPC} < 49$), or susceptible ($\Sigma\text{AUDPC} \geq 49$) to the eight *Xpm* strains evaluated (Table 4 and S3 Fig). One cultivar, cv. 60444, was susceptible to all strains tested and two varieties (NGA11 and CM523-7) were susceptible to nearly all strains tested. On the other hand, cultivars COL1505 and CM6438-14 were resistant or moderately resistant to at least three of the strains evaluated (Table 4).

Regarding the strain pathogenicity level, strains CIAT1202, UA556, UA2164, CIO151 and CIAT1241 caused disease in all inoculated varieties. On the opposite hand, strain CIAT1135 was poorly pathogenic. Based on this information, pathological patterns were classified as follows. The more virulent strains (CIAT1202, UA556, UA2164, CIO151 and CIAT1241) were classified in group I. Strains with intermediate virulence (UA1591 and CIAT1205) were designated in group II. The least virulent strain CIAT1135 was assigned to group III (Table 4).

Data shown is the mean of the sum of AUDPC in arbitrary units at 28 dpi from five replicates per genotype. A genotype was considered resistant (R) if $\Sigma\text{AUPDC} \leq 39$, moderately resistant (MR) if $39 < \Sigma\text{AUDPC} < 44$, moderately susceptible (MS) when $44 < \Sigma\text{AUDPC} < 49$ and susceptible (S) when $\Sigma\text{AUDPC} \geq 49$, NE: Not evaluated.

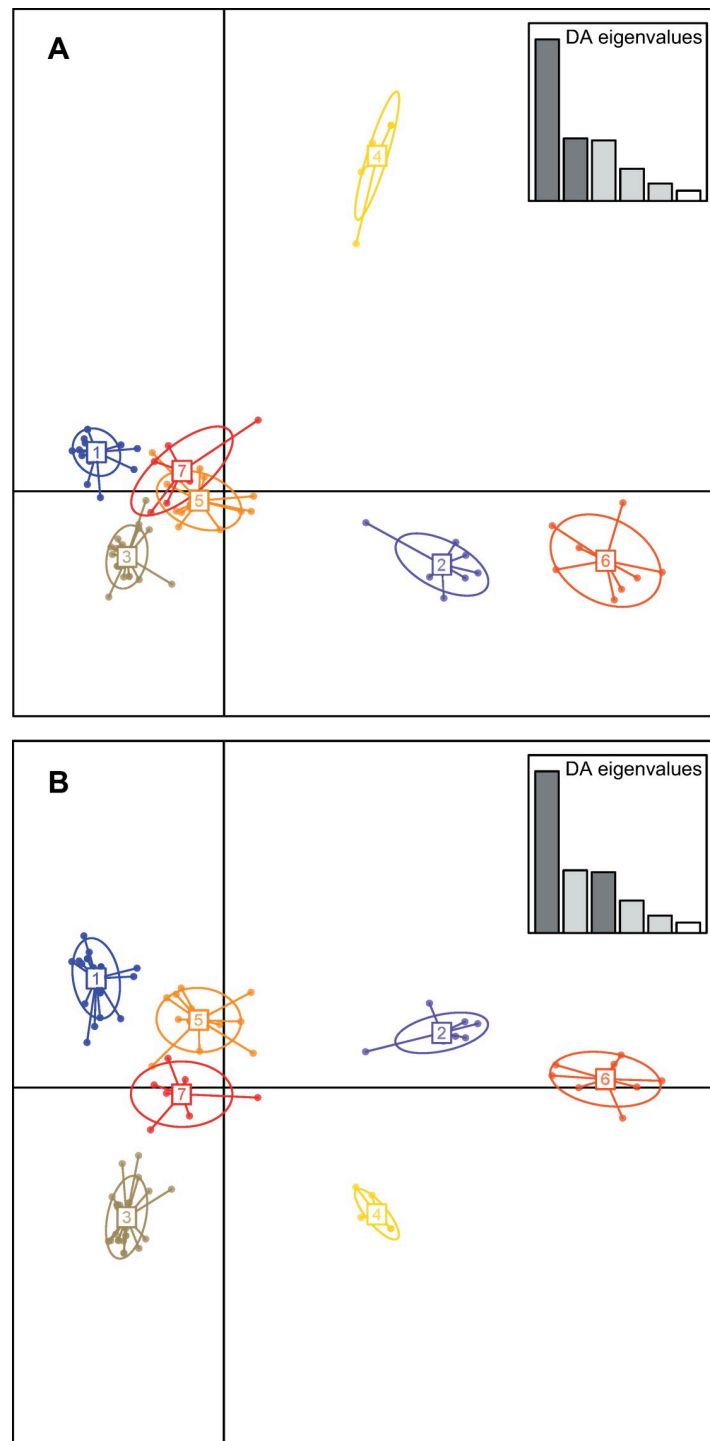


Fig 3. Genetic structure of a worldwide collection of *Xanthomonas phaseoli* pv. *manihotis* based on the discriminant analysis of principal components (DAPC) of minisatellite and microsatellite data (MLVA-12). Numbers and colors represent the seven genetic clusters retained from Bayesian information criterion (BIC) values. (A) Scatterplot representing axes 1 and 2 of the DAPC. (B) Scatterplot representing axes 1 and 3 of the DAPC.

<https://doi.org/10.1371/journal.pone.0285491.g003>

Table 4. Pathogenicity tests performed using eight representative *Xpm* strains.

Strain	Cassava Variety					Relative virulence group	Haplotype MLVA-8 (MLVA-12)	Country
	CM523-7	NGA11	COL1505	CM6438-14	cv.60444			
CIAT1202	S (58,5±3,4)	S (66,7±1,1)	S (61,8±5,0)	NE	S (52±3,1)	I	13 (30)	Colombia
UA556	S (68,2±7,6)	S (56,4±5,1)	MS (47,2±11,2)	S (69,1±7,3)	S (73,3±3,8)	I	4 (11)	Colombia
UA1591	R (29,9±8,9)	S (49,6±6,3)	MR (41,9±8,8)	MR (42,4±2,4)	NE	II	7 (59)	Colombia
UA2164	S (55,6±8,5)	S (81±2,5)	S (60,9±4,5)	NE	NE	I	1 (7)	Colombia
CIO151	S (59,3±1,2)	S (69,1±10,1)	S (53,1±5,5)	S (50,7±2,0)	MS (44,9±6,5)	I	18 (18)	Colombia
CIAT1135	R (5,2±2,0)	R (20,3±3,7)	R (0)	R (22,9±2,4)	NE.	III	19 (16)	China
CIAT1205	MS (47,1±3,0)	S (67,2±3,1)	R (26,3±6,2)	NE	MS (48,1±8,9)	II	23 (25)	New Zealand
CIAT1241	S (67±5,8)	S (68,5±2,5)	S (53,5±6,0)	S (55,4±7,4)	NE	I	15 (58)	Argentina

<https://doi.org/10.1371/journal.pone.0285491.t004>

Discussion

Epidemiology can be deciphered at different time and space scales by selecting the level of genomic resolution of the markers to monitor either the geographic spread of inoculum or the prevalence of certain epidemic or endemic clones for surveillance [2–4,6]. Cassava Bacterial Blight is a destructive disease widely distributed in the different tropical areas where cassava is grown. It has recently been reported in new areas [50]. A global molecular epidemiology analysis is needed to better understand the expansion of the disease for improving the monitoring of CBB. We recently developed a MLVA scheme targeting microsatellites with a high discriminatory power expected to be more adapted to local epidemiology and outbreak analysis [40]. Here, we evaluated new minisatellites for their usefulness in describing the relationships of non-epidemiologically related strains and large-scale epidemiology. These tandem repeat sequences, i.e. micro- and minisatellites, show motifs whose length variability is one of the main factors known to influence the mutation rate of VNTRs [14]. We propose a MLVA-12 scheme which takes advantages of combining TR loci with different mutation rates for global epidemiological surveillance.

Eight minisatellites (MLVA-8) with a unit length ranging from 10 to 26 bp (i.e. markers with an expected low mutation rate) separated a world collection of *Xpm* strains isolated over a period of fifty years (1966–2016) in three CCs and a singleton. The major CC1 grouped strains originating from all the represented countries, except the unique Chinese strain that was included in our collection. All the strains from Africa grouped in this CC while strains from South America, the continent where the disease probably originated, were distributed in the three CCs. The MLVA-8 dataset confirmed (i) a more extensive genetic diversity for the South American strains than for the African strains and (ii) the hypothesis of some genetic links between African strains and some South-American strains as previously shown using four RFLP probes and SNPs based-phylogeny [26,37].

The MLVA-8 scheme alone lacked discriminatory power and therefore did not reveal a fully meaningful epidemiological perspective at a global scale. The relatively low number of minisatellite loci identified and their low mutation rates both are factors favoring homoplasy [51,52]. In contrast, 89 out of 93 *Xpm* strains genotyped using the MLVA-14 scheme were separated as unique haplotypes clearly confirming the high mutation rate of this genotyping technique, consistent with a previous study and suitable for small scale epidemiology [40].

A MLVA-12 scheme combining the eight minisatellites and four microsatellites from the MLVA-14 and displaying the lowest evolution rates is proposed to increase discriminatory power and minimize homoplasy effects [51,52]. In consequence the MLVA-12 scheme, will increase the genetic space of possible patterns as compared to MLVA-8 and would be suitable for epidemiological surveillance. The variable mutation rates within the MLVA-12 scheme allowed a nested approach where the genetic relationships among strains is described within deeper phylogenetic clusters. Such approaches, when combining markers with different discriminatory power, progressively or simultaneously, have been found to increase the accuracy of the phylogeny and to minimize the homoplasy effects [4,52,53].

A DAPC analysis of MLVA-12 data separated the haplotypes in seven distinct clusters. Interestingly, this analysis placed strains from South America in all clusters. This further supports the wide genetic diversity already described in South America (*i.e.* center of origin of cassava and the continent where CBB was firstly reported [26,30,33,37]). In contrast, all African strains studied here grouped in a single cluster also containing strains from South America, suggesting a possible relatedness and the existence of a common ancestor. However, based on MLVA-14, the amount of genetic relatedness between African and the closest South American strains did not support recent epidemiological links. Cassava was introduced in Africa during the seventeenth and eighteenth centuries at a period where regular communications already existed between the two continents [54]. After its first description in Brazil in 1912, cassava bacterial blight was first reported in continental Africa in the early 1970's [24,54]. The pathogenic bacterium was likely introduced in this continent through transportation of infected plant material after multiple introduction events into both Eastern and Western African coasts. Three strains isolated in New-Zealand during outbreaks in 1960s or in 1980 were distributed in two different DAPC clusters, one of them sharing an MLVA-12 haplotype with a strain from Venezuela and the others were SLVs of Colombian and Brazilian strains. This result supports the occurrence of multiple introduction events of *Xpm* in this country.

No clear correlation between the MLVA profiles and virulence patterns occurred. The different haplotypes representative of the genetic variability of observed within our world *Xpm* collection from MLV8-dataset produced different pathological patterns though all the isolates were virulent on the Colombian cassava varieties. Such variability in the pathogenicity of *Xpm* from an international collection without any link between the genetic and pathological patterns was previously reported using different molecular markers [55]. Variations of strain aggressiveness and a lack of correlation between the genetic and pathological patterns were further described at a regional or local scale in different countries revealing dynamic spatio-temporal changes within the *Xpm* populations [29,30,33,36,38].

Conclusions

The MLVA-12 scheme proposed in this study combined both minisatellite and microsatellite loci that brought complementary information based on their different discriminatory power. *Xpm* is a monomorphic plant pathogen widely spread in tropical regions. This portable and resolving genotyping tool could be useful to further investigate the global epidemiology of this plant pathogenic bacterium of cassava, whose improved understanding will require additional data.

Supporting information

S1 Fig. Minimum spanning tree displaying the relationships between haplotypes using the MLVA-14 Microsatellite scheme. Colors indicate the origin of the haplotype, and the circle size indicates the number of strains of each haplotype. Blue circles indicate clonal complexes

and orange circles indicate groups of double locus variants. Numbers indicate the number of loci variants between haplotypes.

(TIF)

S2 Fig. Minimum spanning tree displaying the relationships between haplotypes using the MLVA-12 scheme. Colors indicate the origin of the haplotype, and the circle size is relative to the number of strains of each haplotype. The solid lines represent clonal complexes (CCs) grouping single locus variants and dotted lines group up to double locus variants. Numbers indicate the number of loci variants between haplotypes.

(TIF)

S3 Fig. Disease symptoms for resistant (R), moderately resistant (MR), susceptible (S), and moderately susceptible (MS) cultivars.

(TIF)

S4 Fig. Original images of agarose gel from Fig 1.

(TIF)

S1 Table. Worldwide collection of *Xanthomonas phaseoli* pv. *manihotis* strains analyzed in this study and haplotypes obtained from the different MLVA schemes.

(DOCX)

S2 Table. Multiplex scheme and primer pairs used in the MLVA-14 scheme targeting microsatellites.

(DOCX)

S3 Table. Diversity indices for MLVA-8 scheme from genotyping of a worldwide collection of *Xpm* (n = 93).

(DOCX)

S4 Table. Diversity indices for MLVA-14 scheme from genotyping of a worldwide collection of *Xpm* (n = 93).

(DOCX)

Acknowledgments

In memoriam of Christian Vernière, who was a wonderful person and researcher.

Author Contributions

Conceptualization: Ralf Koebnik, Adriana Bernal.

Data curation: Laurence Blondin, Paula Diaz Tatis, Andrea Camargo.

Formal analysis: Leidy Rache, Laurence Blondin, Paula Diaz Tatis, Andrea Camargo, Camilo López, Olivier Pruvost, Christian Vernière.

Funding acquisition: Boris Szurek, Ralf Koebnik, Silvia Restrepo, Adriana Bernal.

Investigation: Leidy Rache.

Methodology: Leidy Rache, Paula Diaz Tatis, Carolina Flores, Andrea Camargo, Moussa Kante, Issa Wonni, Camilo López, Ralf Koebnik, Adriana Bernal.

Project administration: Adriana Bernal.

Supervision: Ralf Koebnik, Silvia Restrepo, Adriana Bernal.

Validation: Ralf Koebnik.

Writing – original draft: Leidy Rache, Paula Diaz Tatis, Camilo López, Olivier Pruvost, Ralf Koebnik, Silvia Restrepo, Adriana Bernal, Christian Vernière.

Writing – review & editing: Leidy Rache, Boris Szurek, Stephane Dupas, Olivier Pruvost, Ralf Koebnik, Silvia Restrepo, Adriana Bernal, Christian Vernière.

References

1. Spratt BG, Maiden MCJ. Bacterial population genetics, evolution and epidemiology. *Philosophical Transactions of the Royal Society Series B*. 1999; 354:701–10. <https://doi.org/10.1098/rstb.1999.0423> PMID: 10365396
2. Tibayrenc M. Toward an integrated genetic epidemiology of parasitic protozoa and other pathogens. *Annu Rev Genet*. 1999; 33: 449–77. <https://doi.org/10.1146/annurev.genet.33.1.449> PMID: 10690415
3. Van Belkum A, Struelens M, De Visser A, Verbrugh H, Tibayrenc M. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. *Clin Microbiol Rev*. 2001; 14(3): 547–60. <https://doi.org/10.1128/CMR.14.3.547-560.2001> PMID: 11432813
4. Keim P, Van Ert MN, Pearson T, Vogler AJ, Huynh LY, Wagner DM. Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. *Infect Genet Evol*. 2004; 4(3): 205–13. <https://doi.org/10.1016/j.meegid.2004.02.005> PMID: 15450200
5. Kremer K, Arnold C, Cataldi A, Gutierrez MC, Haas WH, Panaiotov S, et al. Discriminatory power and reproducibility of novel DNA typing methods for *Mycobacterium tuberculosis* complex strains. *J Clin Microbiol*. 2005; 43(11): 5628–38.
6. Schlotterer C. The evolution of molecular markers—just a matter of fashion? *Nat Rev Genet*. 2004; 5(1):63–9. <https://doi.org/10.1038/nrg1249> PMID: 14666112
7. Urwin R, Maiden MCJ. Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol*. 2003; 11(10): 479–87. <https://doi.org/10.1016/j.tim.2003.08.006> PMID: 14557031
8. Achtman M. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. *Annu Rev Microbiol*. 2008; 62:53–70. <https://doi.org/10.1146/annurev.micro.62.081307.162832> PMID: 18785837
9. Vogler AJ, Birdsell D, Price LB, Bowers JR, Beckstrom-Sternberg SM, Auerbach RK, et al. Phylogeography of *Francisella tularensis*: global expansion of a highly fit clone. *J Bacteriol*. 2009; 191(8): 2474–84.
10. Wilson DJ. Insights from genomics into bacterial pathogen populations. *Plos Pathogens*. 2012; 8(9). doi: ARTN e1002874 <https://doi.org/10.1371/journal.ppat.1002874> PubMed PMID: WOS:000309816500003. PMID: 22969423
11. Cui Y, Yu C, Yan Y, Li D, Li Y, Jombart T, et al. Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proceedings of the National Academy of Sciences, USA*. 2013; 110(2):577–82. doi: 10.1073/pnas.1205750110. PubMed PMID: 23271803; PubMed Central PMCID: PMC3545753.
12. Pourcel C, Vergnaud G. Strain typing using Multiple “Variable Number of Tandem Repeat” Analysis and genetic element CRISPR. In: Persing DH, Tenover FC, Tang YW, Nolte FS, Hayden RT, Van Belkum A, editors. *Molecular microbiology: Diagnostic principles and practice*. 2nd ed. Washington, DC: ASM Press; 2011. p. 179–97.
13. Van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infect*. 2007; 13(Suppl. 3): 1–46. <https://doi.org/10.1111/j.1469-0691.2007.01786.x> PMID: 17716294
14. Van Belkum A, Scherer S, Van Alphen L, Verbrugh H. Short-sequence DNA repeats in prokaryotic genomes. *Microbiol Mol Biol Rev*. 1998; 62(2): 275–93. <https://doi.org/10.1128/MMBR.62.2.275-293.1998> PMID: 9618442
15. Vergnaud G, Pourcel C. Multiple Locus Variable Number of Tandem Repeats analysis. In: Caugant DA, editor. *Molecular epidemiology of microorganisms: methods and protocols*. 551. Dordrecht Heidelberg London New York: Springer Publs.; 2009. p. 141–58.
16. Catara V, Cubero J, Pothier JF, Bosis E, Bragard C, ermić E, et al. Trends in molecular diagnosis and diversity studies for phytosanitary regulated *Xanthomonas*. *Microorganisms*. 2021; 9(4): 862. <https://doi.org/10.3390/microorganisms9040862> PMID: 33923763
17. Mazars E, Lesjean S, Banuls AL, Gilbert M, Vincent V, Gicquel B, et al. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. *Proc Natl Acad Sci USA*. 2001; 98(4): 1901–6.

18. Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Lochet C. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. *Mol Microbiol.* 2000; 36(3): 762–71.
19. Pruvost O, Magne M, Boyer K, Leduc A, Tourterel C, Drevet C, et al. A MLVA genotyping scheme for global surveillance of the citrus pathogen *Xanthomonas citri* pv. *citri* suggests a worldwide geographical expansion of a single genetic lineage. *PLoS one.* 2014; 9(6): e98129.
20. Levinson G, Gutman GA. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Mol Biol Evol.* 1987; 4(3): 203–21. <https://doi.org/10.1093/oxfordjournals.molbev.a040442> PMID: 3328815
21. Schlotterer C. Evolutionary dynamics of microsatellite DNA. *Chromosoma.* 2000; 109(6): 365–71. <https://doi.org/10.1007/s004120000089> PMID: 11072791
22. Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, et al. Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X-axonopodis* species complex. *Plant Pathol.* 2016; 65(5): 792–806. <https://doi.org/10.1111/ppa.12461> PubMed PMID: WOS:000375897400010.
23. Oren A, Garrity JT. Notification of changes in taxonomic opinion previously published outside the IJSEM. *Int J Syst Evol Microbiol.* 2017; 67(7): 2081–6. <https://doi.org/10.1099/ijsem.0.002071> PMID: 28777068
24. Lozano JC. Cassava bacterial blight: a manageable disease. *Plant Dis* 1986; 70: 1089–93.
25. Lozano JC, Sequeira L. Bacterial blight of cassava in Colombia: etiology. *Phytopathology.* 1974; 64: 74–82.
26. Verdier V, Dongo P, Boher B. Assessment of genetic diversity among strains of *Xanthomonas campestris* pv. *manihotis*. *J Gen Microbiol.* 1993; 139(11): 2591–601.
27. Boher B, Verdier V. Cassava bacterial blight in África: the state of knowledge and implications for designing control strategies. *Afr Crop Sci J.* 1994; 2: 505–9.
28. Taylor RK, Griffin RL, Jones LM, Pease B, Tsatsia F, Fanai C, et al. First record of *Xanthomonas axonopodis* pv. *manihotis* in Solomon islands. *Australasian Plant Dis Notes.* 2017; 12: 49.
29. Verdier V, Restrepo S, Mosquera G, Duque MC, Gerstl A, Laberry R. Genetic and pathogenic variation of *Xanthomonas axonopodis* pv. *manihotis* in Venezuela. *Plant Pathol.* 1998; 47(5): 601–8.
30. Restrepo S, Valle TL, Duque MC, Verdier V. Assessing genetic variability among Brazilian strains of *Xanthomonas axonopodis* pv. *manihotis* through restriction fragment length polymorphism and amplified fragment length polymorphism analyses. *Can J Microbiol.* 1999; 45(9): 754–63.
31. Restrepo S, Duque M, Tohme J, Verdier V. AFLP fingerprinting: an efficient technique for detecting genetic variation of *Xanthomonas axonopodis* pv. *manihotis*. *Microbiology—UK.* 1999; 145: 107–14.
32. Restrepo S, Duque MC, Verdier V. Characterization of pathotypes among isolates of *Xanthomonas axonopodis* pv. *manihotis* in Colombia. *Plant Pathol.* 2000; 49(6): 680–7.
33. Restrepo S, Vélez CM, Duque MC, Verdier V. Genetic structure and population dynamics of *Xanthomonas axonopodis* pv. *manihotis* in Colombia from 1995 to 1999. *Appl Environ Microbiol.* 2004; 70(1): 255–61.
34. Restrepo S, Verdier V. Geographical differentiation of the population of *Xanthomonas axonopodis* pv. *manihotis* in Colombia. *Appl Environ Microbiol.* 1997; 63(11): 4427–34.
35. Trujillo CA, Arias-Rojas N, Poulin L, Medina CA, Tapiero A, Restrepo S, et al. Population typing of the causal agent of cassava bacterial blight in the Eastern Plains of Colombia using two types of molecular markers. *BMC Microbiol.* 2014; 14: 161. <https://doi.org/10.1186/1471-2180-14-161> PMID: 24946775
36. Trujillo CA, Ochoa JC, Mideros MF, Restrepo S, Lopez C, Bernal A. A complex population structure of the cassava pathogen *Xanthomonas axonopodis* pv. *manihotis* in recent years in the Caribbean Region of Colombia. *Microb Ecol.* 2014; 68(1): 155–67.
37. Bart R, Cohn M, Kassen A, McCallum EJ, Shybut M, Petriello A, et al. High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proceedings of the National Academy of Sciences, USA.* 2012. <https://doi.org/10.1073/pnas.1208003109> PMID: 22699502
38. Restrepo S, Velez CM, Verdier V. Measuring the genetic diversity of *Xanthomonas axonopodis* pv. *manihotis* within different fields in Colombia. *Phytopathology.* 2000; 90(7): 683–90.
39. Ogunjobi AA, Fagade OE, Dixon AGO. Molecular variation in population structure of *Xanthomonas axonopodis* pv. *manihotis* in the south eastern Nigeria. *Afr J Biotechnol.* 2006; 5(20): 1868–72.
40. Rache L, Blondin L, Flores C, Trujillo C, Szurek B, Restrepo S, et al. An optimized microsatellite scheme for assessing populations of *Xanthomonas phaseoli* pv. *manihotis*. *Phytopathology.* 2019; 109(5): 859–69. Epub 2019/03/26. <https://doi.org/10.1094/PHYTO-06-18-0210-R> PMID: 30908143.

41. Arrieta-Ortiz ML, Rodriguez LM, Perez-Quintero AL, Poulin L, Diaz AC, Rojas NA, et al. Genomic survey of pathogenicity determinants and VNTR markers in the cassava bacterial pathogen *Xanthomonas axonopodis* pv. *manihotis* strain CIO151. *PLoS one*. 2013; 8(11): <https://doi.org/10.1371/journal.pone.0079704> PMID: 24278159
42. Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol*. 1988; 26(11): 2465–6. <https://doi.org/10.1128/jcm.26.11.2465-2466.1988> PMID: 3069867
43. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform*. 2005;1: 47–50.
44. Nascimento M, Sousa A, Ramirez M, Francisco AP, Carrico JA, Vaz C. PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics*. 2017; 33(1): 128–9. <https://doi.org/10.1093/bioinformatics/btw582> PMID: 27605102.
45. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 2008; 24(11): 1403–5. <https://doi.org/10.1093/bioinformatics/btn129> PMID: 18397895
46. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*. 2010; 11: 94. <https://doi.org/10.1186/1471-2156-11-94> PMID: 20950446
47. Jorge V, Verdier V. Qualitative and quantitative evaluation of cassava bacterial blight resistance in F1 progeny of a cross between elite cassava clones. *Euphytica*. 2002; 123: 41–8.
48. Vogler AJ, Keys C, Nemoto Y, Colman RE, Jay Z, Keim P. Effect of repeat copy number on Variable-Number Tandem Repeat mutations in *Escherichia coli* O157:H7. *J Bacteriol*. 2006; 188: 4253–63.
49. Jeger MJ, Viljanen-Rollinson SLH. The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. *Theor Appl Genet*. 2001; 102: 32–40.
50. Kante M, Flores C, Moufid Y, Wonne I, Hutin M, Thomas E, et al. First Report of *Xanthomonas phaseoli* pv. *manihotis*, the causal agent of cassava bacterial blight, in Mali. *Plant Dis*. 2020; 104(6): 1852. <https://doi.org/10.1094/PDIS-12-19-2611-PDN>
51. Estoup A, Jarne P, Cornuet JM. Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol Ecol*. 2002; 11(9): 1591–604. <https://doi.org/10.1046/j.1365-294x.2002.01576.x> PMID: 12207711
52. Reyes JF, Chan CHS, Tanaka MM. Impact of homoplasmy on variable numbers of tandem repeats and spoligotypes in *Mycobacterium tuberculosis*. *Infect Genet Evol*. 2012; 12(4): 811–8.
53. Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS one*. 2009; 4(11): <https://doi.org/10.1371/journal.pone.0007815> PMID: 19915672
54. Persley GJ. Distribution and importance of cassava bacterial blight in Africa. In: G P, Terry ER, M R, editors. *Cassava bacterial blight*. Ibadan, Nigeria: International Development Research Centre; 1976. p. 9–14.
55. Verdier V, Boher B, Maraite H, Geiger JP. Pathological and molecular characterization of *Xanthomonas campestris* strains causing diseases of cassava (*Manihot esculenta*). *Appl Environ Microbiol*. 1994; 60(12): 4478–86.