Genetic diversity of *Yersinia pestis* in Brazil

M.B.M. Oliveira¹, M.P.S. Barros²,³, V.M. Silveira-Filho²,³,
M.R. Araújo-Nepomuceno², V.Q. Balbino¹, N.C. Leal²,
A.M.P. Almeida² and T.C. Leal-Balbin²,³

¹Departamento de Bioquímica, Universidade Federal de Pernambuco, Recife, PE, Brasil
²Centro de Pesquisas Ageu Magalhães (CPqAM/FIOCRUZ-PE), Recife, PE, Brasil
³Departamento de Genética, Universidade Federal de Pernambuco, Recife, PE, Brasil

Corresponding author: T.C. Leal-Balbino
E-mail: cristina@cpqam.fiocruz.br / tcristinaleal@yahoo.com.br

Received November 8, 2011
Accepted April 11, 2012
Published September 25, 2012
DOI http://dx.doi.org/10.4238/2012.September.25.10

ABSTRACT. Plague outbreaks are occasionally reported in Brazil. Unfortunately, due to great genetic similarity, molecular subtyping of *Yersinia pestis* strains is difficult. Analysis of multiple-locus variable number of tandem repeats (VNTR), also known as MLVA, has been found to be a valuable tool to discriminate among strains. To check for genetic differences, strains obtained from two different ecological complexes in Brazil collected during two different epidemiological events, an epizootic in Sítio Alagoinha in 1967 and an outbreak in Planalto da Borborema in 1986, were subtyped through MLVA using 12 VNTR loci. Three clusters (A, B and C) were observed. Of the 20 strains from the epizootic, 18 fit into cluster A. Cluster A was divided into two subgroups: A₁ (15 strains) and A₂ (3 strains). Of the 17 strains from the outbreak, 15 fit into cluster B. Cluster B was divided into three subgroups: B₁ (4 strains), B₂ (4 strains) and B₃ (7 strains). Cluster C is a singleton with one epizootic strain. The external standards, *Y. pestis* CO92 and *Y. pseudotuberculosis* IP32953, formed two clusters of singletons. The stability of 12 VNTR loci of three unrelated cultures
included in this study was assessed. The 12 VNTR loci were stable through multiple serial subcultures in the laboratory. MLVA revealed that *Y. pestis* populations in Brazil are not monomorphic, and that there is intraspecific genetic diversity among Brazilian plague strains. We conclude that there is some correlation among genetic groups of this species, related to the temporal and geographic origin of isolates.

**Key words:** Plague; Molecular epidemiology; MLVA; *Yersinia pestis*

**INTRODUCTION**

The plague, caused by infection with *Yersinia pestis*, was introduced in Brazil by sea through infected rats and fleas carried on steam ships during the last pandemic, which originated in China in 1855. Records of the early history of the plague in Brazil are sporadic and sparse, which makes detailed tracking difficult to impossible. Systematic information became available only after the 1930s (Pollitzer, 1954; Baltazard, 1968). Nevertheless, there is a general consensus that the first human case was registered in October 1899, in Santos, São Paulo State. From then on, other coastal cities became successively infected. Government control measures were promptly taken, and the infection was eliminated from the coast; however, these measures did not stop the spread of the plague to inland cities, towns and villages. The plague reached rural areas, afflicting small districts, farms and ranches. The plague then established many natural foci among the wild rodents in several ecological complexes, from the State of Ceará to the northern Minas Gerais (collectively called the Northeast focus) and in the Serra dos Orgãos, Rio de Janeiro (WHO, 1965; Baltazard, 1968).

Tight control of the plague in the 1930s reduced human cases and mortality. Unfortunately, this program was interrupted, and the disease reemerged in the 1960s. To achieve new insights into the features of plague epidemiology and to reorganize an effective control strategy, an extensive project comprised of field and laboratory research was established and carried out from 1966 to 1974 (Baltazard, 2004). By the end of the project, more than 600 *Y. pestis* strains were isolated, which resulted in a unique collection of *Y. pestis* cultures in Brazil. Subsequent epidemiological investigations and surveillance activities in northeastern Brazilian plague areas expanded this collection (Rocha et al., 2009). The collection is now maintained by the National Service for Reference in Plague of the Centro de Pesquisas Aggeu Magalhães (SRP/CPqAM), Recife, PE, and linked to FIOCRUZ culture collections (Fiocruz-CYP).

Attempts have been made to genotype many of these isolates. In almost every case, no differences between isolates were found, even across multiple host species, different geographic origins, or years (Sobreira M and Leal NC, unpublished results). Molecular subtyping of *Y. pestis* for epidemiological tracking has been difficult due to the great genetic similarity among isolates (Vogler et al., 2011). However, analysis of multiple-locus variable number of tandem repeats (VNTR), also known as MLVA, has become a valuable tool to discriminate *Y. pestis* strains. Using MLVA, it is possible to trace the dispersion routes of bacterial populations at the local, regional and global levels (Adair et al., 2000; Klevytska et al., 2001; Achtman et al., 2004; Girard et al., 2004; Pourcel et al., 2004; Lowell et al., 2005; Vogler et al., 2011).

In the present study, we analyzed *Y. pestis* strains from two different ecological com-
plexes in Brazil during two different epidemiological events. MLVA using 12 VNTRs revealed that *Y. pestis* populations in Brazil are not monomorphic. We also demonstrated intraspecific genetic diversity among the strains, which establishes a relative correlation between genetic groups and the temporal and geographic origin of the isolates.

**MATERIAL AND METHODS**

**Bacterial strains studied**

The study involved 37 *Y. pestis* strains that were kindly provided by the Fiocruz-CYP culture collection, which is maintained by the SRP/CPqAM-FIOCRUZ/PE. Twenty strains were isolated from rodents and fleas during an epizootic in August 1967. This epizootic occurred in Sítio Alagoinha, a locality situated in Chapada do Araripe, Pernambuco State (de Almeida et al., 1985). Seventeen strains were isolated from an outbreak in 1986, which involved several municipalities in Planalto da Borborema, Paraíba State (de Almeida et al., 1989). Three unrelated strains (P. CE 882, P. Exu 340 and P. Peru 375) and 11 derived cultures obtained through multiple serial subcultures of the parental strains (Leal-Balbino et al., 2004) were included to assess *in vitro* VNTR stability. Figure 1 shows the localization of the Brazilian plague foci and origin of the strains.

Epizootic strains were identified as P (plague), Exu (laboratory city), and sequential number. Outbreak strains were identified as P (plague), PB (state of origin) and sequential number.

![Figure 1. Geographic origin of the *Yersinia pestis* strains analyzed. CE = Ceará (1: Ipu); PB = Paraíba (2: Solânea, 3: Remígio), and PE = Pernambuco (4: Ipubi, 5: Exu).](image)
Culture conditions

Cultures were maintained on peptone agar stabs at 4°C. The strains were inoculated into brain heart infusion broth at 28°C and inspected daily for up to seven days. Each culture grown during this period was plated on blood agar base, incubated at 28°C and analyzed on a daily basis for up to five days to assess the morphology of the developing colonies. Identification of the culture and confirmation of purity were carried out by the plague phage test (Karimi, 1978).

VNTR primers

Primers designed to amplify the 12 chromosomal VNTRs published by Adair et al. (2000), Le Flèche et al. (2001) and Klevytska et al. (2001) were used for PCR amplification in the Y. pestis cultures. It should be pointed out that some loci were given different names by different authors. In this paper, the standardized nomenclature from Pourcel et al. (2004) was used. A description of the primers and references is presented in Table 1.

<table>
<thead>
<tr>
<th>Locus*</th>
<th>Alias</th>
<th>Primer sequence</th>
<th>Repeat unit (bp)</th>
<th>Reference for primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>mo04 M58</td>
<td>F: 5'-GGG ATA ACC CAC ATT AAT AAT ACC AAC AC-3'</td>
<td>17</td>
<td>Klevytska et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>ms05 M59</td>
<td>F: 5'-GGT TGG CCA GGA AAG AGT TGG C-3'</td>
<td>17</td>
<td>Klevytska et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>mo06 -</td>
<td>F: 5'-AAT TGG CCA AAA AAT GCA T-3'</td>
<td>60</td>
<td>Le Flèche et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>mo07 M37</td>
<td>F: 5'-GTT TCC CCA TTA GAA TAA GTA G-3'</td>
<td>10</td>
<td>Klevytska et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>mo09 -</td>
<td>F: 5'-CTT CAT CCA CCA GCT TCA C-3'</td>
<td>18</td>
<td>Le Flèche et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>ms20 M51</td>
<td>F: 5'-GGA ACC ACC CCG TGA AAT TAA AAT CAA C-3'</td>
<td>15</td>
<td>Klevytska et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>mo30 -</td>
<td>F: 5'-ACA CAA TAC CAG GGT CTA CTA C-3'</td>
<td>54</td>
<td>Le Flèche et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>ms45 M42</td>
<td>F: 5'-GGG GGG CAA GGT GGT GGT TTA TTA TTA TTC T-3'</td>
<td>12</td>
<td>Klevytska et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>mo46 -</td>
<td>F: 5'-CAG TTA CCT TAT TGT TAT TAA ACT G-3'</td>
<td>7</td>
<td>Le Flèche et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>ms54 -</td>
<td>F: 5'-GTG TCG TGA TGA TGA TAA G-3'</td>
<td>22</td>
<td>Le Flèche et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>ms62 M34</td>
<td>F: 5'-AAA TCC CCA GCA TTT GGG TGG TGG C-3'</td>
<td>9</td>
<td>Klevytska et al. (2001)</td>
<td></td>
</tr>
<tr>
<td>1AB -</td>
<td>F: 5'-GGT TAG GTG AGA TGG TAA GGA G-3'</td>
<td>4</td>
<td>Adair et al. (2000)</td>
<td></td>
</tr>
</tbody>
</table>

* Nomenclature as outlined in Pourcel et al. (2004).

DNA extraction and PCR amplification

Genomic DNA was prepared as described by Souza et al. (2007). The PCRs were carried out in a final volume of 25 μL that contained 20 ng DNA, 20 pmol of each primer, 50 mmol KCl, 20 mmol Tris-HCl, pH 8.0, 0.16 mmol deoxynucleoside triphosphates (dNTPs, Invitrogen, Brazil), 1.5 mmol MgCl₂, and 1 U Taq DNA polymerase (Invitrogen).

The amplification of locus 1AB was carried out as described by Adair et al. (2000). The amplification of the other loci was carried out according to Le Flèche et al. (2001).
The 1-D Image Analysis imaging software (version 3.5; Kodak Digital Science) was used to determine the size of the amplicons. The number of repeat units (NRU) for each locus was determined using the following formula, which we inferred from Adair et al. (2000):

\[ NRU = A_S - FR_S / RU_S, \]

where \( A_S \) is the amplicon size, \( FR_S \) is the size of the flanking region sequence and \( RU_S \) is the size of each repeat unit. \( FR_S \) and \( RU_S \) were identified by in silico analysis of the \( Y. pestis \) CO92 genome sequence with the Tandem Repeats Finder software (Benson, 1999). The Li et al. (2009) standardized nomenclature was used for all loci, except for ms30 (Le Flèche et al., 2001) and 1AB (Adair et al., 2000).

One representative amplicon of each VNTR was sequenced to confirm the size of the repeat and the copy number. Due to their small size (4 bp), all the amplicons generated with primer 1AB were sequenced to confirm the repeat size and copy number. For clustering analyses, the categorical coefficient and unweighted pair group method using arithmetic averages (UPGMA) were used. The use of the categorical parameter assumes that the character states are unordered. The same weight is given to a large or a small number of differences in the number of repeats at each locus. The UPGMA tree was constructed using MEGA v. 5.0 (Tamura et al., 2011).

MLVA patterns were normalized using \( Y. pestis \) CO92 (GI:30407161) and \( Y. pseudotuberculosis \) IP32953 (GI:51594359) public genomic data as external standards.

The numerical discriminatory index \((DI)\) of each VNTR locus was evaluated using the Hunter and Gaston (1988) formula:

\[ DI = 1 - [1 / N (N - 1)] \sum nj(nj - 1), \]

where \( N \) is the total number of isolates in this population, and \( nj \) is the number of isolates representing each type. This approach is based on the probability that two unrelated strains from the test population will be placed into different groups.

Sequencing

For sequencing, the amplicons were purified with the Purelink PCR purification kit (Invitrogen), following manufacturer instructions. The sequencing reaction was performed with the same primers used for PCR. An ABI 3100 automated DNA sequencer (Applied Biosystems) and the BigDye kit (Applied Biosystems) were used for sequencing. The Basic Local Alignment Search Tool (BLASTN) (Altschul et al., 1997) was used to confirm the identity of the sequenced amplicons by comparing them against the NCBI databank of non-redundant nucleotide sequences. Seq Tools 8.3.022 (Rasmussen, 2002) and BioEdit v.
7.0.5.2 (Hall, 1999) were used to edit the generated sequences. The sequences were submitted to and published in the GenBank database under the following accession Nos: 1AB (GI:89112565); ms04 (GI:238867221); ms05 (GI:238867222); ms06 (GI:89112563); ms07 (GI:238867223); ms09 (GI:238867224); ms20 (GI:238867225); ms30 (GI:238867226); ms45 (GI:238867227); ms46 (GI:238867228); ms54 (GI:238867229), and ms62 (GI:238867230). The copy number variation of the VNTR loci was determined with the Tandem Repeats Finder tool (Benson, 1999) using the basic alignment parameters described on the homepage of the software (accessible at http://tandem.bu.edu/trf/trf.html).

RESULTS

MLVA based on 12 VNTR chromosomal loci was used to subtype two sets of Y. pestis strains from two different ecological complexes in Brazil during two different epidemiological events, an epizootic in Sítio Alagoinha in 1967 (20 strains) and an outbreak in Planalto da Borborema in 1986 (17 strains).

The primers directed to the 12 VNTRs yielded PCR products in all the cultures tested, which means that the loci were all present in the strains analyzed. MLVA was also performed on the 12 loci in the parental strain and their derivatives obtained by multiple serial subcultures of P. CE 882, P. Peru 375 and P. Exu 340 strains. Identical patterns were revealed in the parental and derived cultures (data not shown).

Of the 12 VNTR loci analyzed in the 37 outbreak and epizootic strains, seven (ms04, ms05, ms06, ms07, ms20, ms45, and ms62) revealed polymorphisms. These seven loci displayed differences in amplicon length and copy number, with a \( D_I \) ranging from 0.4 to 0.9. Strains from the two groups shared at least one allele in each locus. Five VNTRs revealed a monomorphic pattern: ms09 (20 repeats), ms30 (8), ms46 (4), ms54 (5), and 1AB (10) (Figures 2 and 3).

Figure 2. Analysis of PCR products by 3.0% agarose gel electrophoresis to determine repeat size polymorphisms at the ms06 and ms07 loci. The number of repeats is indicated under each amplicon.
Using a cut-off value of 2.5 differences, the 37 epizootic and outbreak strains were separated into three distinct clusters (A, B and C). The external standards, *Y. pestis* CO92 and *Y. pseudotuberculosis* IP32953, formed two distinct singletons (Figure 4). Cluster A was divided into two subgroups: A1 (15 strains) and A2 (5 strains). Cluster B was divided into three subgroups: B1 (4 strains), B2 (5 strains) and B3 (7 strains). Cluster C was a singleton (Figure 4).

**Figure 3.** Distribution of the number of alleles (bars) and the discriminatory index (markers) found among the 12 VNTRs.

**Figure 4.** UPGMA dendrogram showing clustering and correlation of the *Yersinia pestis* strains studied with their origin and distribution by groups. H = Human; Hs = *Holochilus sciureus*; Nl = *Necromys lasiurus*; Os = *Oryzomys subflavus*; Rr = *Rattus rattus*; Xc = *Xenopsylla cheopis*; Pbj/Nl = *Polygenis bolshi jordani* from *N. lasiurus*; S = Solânea; E = Exu; R = Remígio; MLVA = multiple-locus variable number of tandem repeats.
Cluster A, which was the dominant cluster, grouped all but two epizootic strains and included 2 outbreak strains (P. PB 866 and P. PB 869). Subgroup A₁ was specific for the epizootic strains and included most of them: Necromys lasiurus (9), Polygenis bolshi jordani (4), Rattus rattus (1), and Xenopsylla cheopis (1). Subgroup A₂ included 3 epizootic strains: N. lasiurus (1), P. b. jordani (1), and R. rattus (1). Subgroup A₃ also included 2 outbreak strains: Holochilus sciureus (1) and R. rattus (1) (Figure 4).

Cluster B included all but two outbreak strains plus one epizootic strain. Subgroups B₁ and B₂ were exclusively composed of outbreak strains. Subgroup B₁ included strains from R. rattus (1), Oryzomys subflavus (1) and human (1); subgroup B₂ included strains from N. lasiurus (4), O. subflavus (2) and human (1). Subgroup B₃ included outbreak strains from R. rattus (2), N. lasiurus (1) and O. subflavus (1) and one epizootic strain from N. lasiurus (Figure 4).

Some strains from the wild host, the rodent N. lasiurus, and its flea vector, P. b. jordani, e.g., P. Exu 33/P. Exu 36, P. Exu 53/P. Exu 59, P. Exu 31/P. Exu 35, and P. Exu 40/P. Exu 44, were closely related. These strains only differed at one to three loci. The two epizootic strains unaffiliated with cluster A grouped in different clusters: B₂ (P. Exu 42) and C (P. Exu 47) (Figure 4).

**DISCUSSION**

Studies on the molecular epidemiology of *Y. pestis* have been hampered due to the great similarity of isolates. However, new molecular approaches have uncovered a relatively high level of genetic diversity. The global expansion of *Y. pestis* strains from central Asian plague foci was demonstrated by Li et al. (2009) with an MLV A of 25 VNTRs. MLVA was also successfully employed to study the maintenance and spread of *Y. pestis* in Madagascar (Vogler et al., 2011).

In the present study, a phylogenetic analysis based on 12 VNTRs was performed on *Y. pestis* strains from two different epidemiological events in two northeastern Brazilian plague areas: an epizootic in Chapada do Araripe, Pernambuco State (de Almeida et al., 1985) and an outbreak in Planalto da Borborema, Paraíba State (de Almeida et al., 1989).

In Pernambuco State, plague was registered for the first time in 1902 on the docks of Recife, the capital city. Plague is suspected of spreading via transportation of cargo and reached Chapada do Araripe in 1919, establishing one of the most active plague foci in Brazil (WHO, 1965; Baltazard, 2004). MLVA performed on 20 strains from an epizootic in Sítio Alagoinhin, a small property located in Chapada do Araripe in 1967, allowed us to distinguish two subgroups (Figure 4). Most of the strains (75%) grouped in subgroup A₁, suggesting that they were all from a single “epizootic path”. This subgroup assembled strains from the wild rodent, N. lasiurus, and its fleas (P. b. jordani). This subgroup also included strains from the commensal rat (R. rattus) and the rat flea, X. cheopis, collected on the ground inside the houses.

After studies on the ecology of plague in the Chapada do Araripe focus, Baltazard (2004) concluded that *Y. pestis* progresses as individual “epizootic paths”. These “paths” wind their way randomly in short incursions through the fields, exterminating the populations of afflicted rodents. Given the extreme sedentary nature of rodents and the small radius of their excursions, such “epizootic paths” progress slowly and cover very small areas. Each “path” winds its way through the fields autonomously. This behavior could explain the presence of a dominant MLVA genotype among different species of rodents and fleas from the epizootic in a small area.
According to Baltazard (2004), the field rodent *N. lasiurus* and its fleas (*Polygenis* spp) are by far the most frequently infected species in the wild and play an important role in plague amplification. However, other species of rodents and fleas are also involved. Our results provide further evidence of *Y. pestis* interchange between wild rodent hosts and the commensal rat and their fleas in Brazilian foci.

In Paraíba State, plague was first registered in 1912. In September 1986, after a quiescent period of 10 years, an outbreak emerged. This outbreak spread quickly throughout 41 municipalities and affected humans and several species of rodents from many localities on the Planalto da Borborema (de Almeida et al., 1989). MLVA distinguished three subgroups (B₁, B₂ and B₃) among the outbreak isolates analyzed. Thus, it appears that distinct *Y. pestis* populations (the multiple “epizootic paths”) circulated in the area during the outbreak. According to Baltazard (2004), when the rodent population density is high, each “epizootic path” extends and crosses and crisscrosses the other “paths”. This pattern continues until the rodent population climaxes. The epizootic can then spread over vast areas, reaching all species, and then human cases arise. Most of the outbreak strains (41.2%) clustered in subgroup B₃ and one human strain (P. PB 862) clustered with two *N. lasiurus* strains isolated from the same locality. These different *Y. pestis* populations described in the present study may reflect the microevolution of the invading strain due to environmental pressure experienced while spreading inland from the coast. Alternatively, the strains could be inaccurate representatives of the original isolates, and these differences could be due to handling procedures during the prolonged storage of the strains in the laboratory since their isolation (Revazishvili et al., 2008; Vogler et al., 2011). Leal et al. (2000) and Leal-Balbino et al. (2004, 2006) observed spontaneous changes in the genome (both plasmid and chromosomal) of Brazilian *Y. pestis* following prolonged storage and laboratory handling. However, analysis of the 12 VNTR loci in the three parental strains and their derived cultures did not reveal changes in the studied regions, confirming their stability as previously observed (Adair et al., 2000; Keim et al., 2000). Furthermore, according to Vogler et al. (2011), the rate of VNTR evolution in the laboratory, even with passaging, should be much slower than in nature. While these isolates may not be exactly the same as when they were first isolated, they should be quite similar to the original isolates.

In Brazil, all the plague foci have been mostly dormant since the mid-1980s. The occurrence of new cases has been rare, and the last cases were in 1997 and 2005 in Ceará. However, serological surveys continuously detect plague antibodies in sentinel animals in plague areas (Leal and Almeida, 1999; Aragão et al., 2009), suggesting that these plague foci remain active and may reappear (Duplantier et al., 2005; Gage and Kosoy, 2005; Stenseth et al., 2008; WHO, 2006, 2009).

Brazilian plague foci are autonomous, i.e., totally independent of each other, and there is no connection between them. They have different epidemiological characteristics and patterns of disease manifestations in the human population. The foci have alternating periods of high incidence followed by other periods of low incidence or even the absence of cases (Baltazard, 2004; Giles et al., 2011). The presence of similar genotypes separated by long geographical and temporal distances as in the A₂ and B₂ subgroups remains to be explained.

Further studies with all 43 VNTR markers previously described for *Y. pestis* (Adair et al., 2000; Klevytska et al., 2001; Girard et al., 2004) should contribute to a better understanding of the changes that have occurred in the *Y. pestis* genome over time in the Brazilian foci. The study of a larger number of strains from all the known Brazilian foci and strains isolated in
different epidemiological periods may reveal the dispersion of the different \( Y. \) \textit{pestis} “paths”.

In conclusion, MLV A based on 12 VNTRs revealed that \( Y. \) \textit{pestis} populations in Brazil are not monomorphic and demonstrated intraspecific genetic diversity among Brazilian plague strains. This study establishes a relative correlation for the first time among genetic groups based on the temporal and geographic origin of the Brazilian isolates. The knowledge of the characteristics of local strains will be very helpful in elucidating the origin of new strains, if by emergence of natural foci, reemergence or coming from other countries.

ACKNOWLEDGMENTS

Research supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico; Edital MCT/CNPq 14/2009 Process #482694/2009-2) and SRP/CPqAM-FIOCRUZ. We thank Silvana Santos from the Department of Microbiology of CPqAM for technical assistance.

REFERENCES


