Development of discriminatory multiple-locus variable number tandem repeat analysis for *Bartonella henselae*

Martine Monteil,1 Benoît Durand,2 Rim Bouchouicha,1 Elisabeth Petit,1 Bruno Chomel,3 Mardjan Arvand,4 Henri-Jean Boulouis1† and Nadia Haddad1

1UMR BIPAR/AFSSA/INRA/ENVA/UVPM, 94703 Maisons-Alfort, France  
2AFSSA/LERPAZ Unité d’épidémiologie, 94703 Maisons-Alfort, France  
3Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA 95616, USA  
4Institut für Medizinische Mikrobiologie, Virologie und Hygiene, Universität Rostock, Rostock, Germany

*Bartonella henselae* is a zoonotic bacterium that infects cats and humans. Several attempts have been made to develop typing techniques for epidemiological purposes; however, most of the techniques developed do not appear to be sufficiently discriminatory or easy to use. In order to develop multilocus variable number tandem repeat (VNTR) analysis (MLVA) for *B. henselae*, 30 VNTR candidates were selected from the genome sequence of the reference strain Houston 1 (H1). The VNTR candidates were initially tested by PCR on six *B. henselae* isolates from different geographical areas. Five VNTRs were selected from those that showed two or more alleles. These five *B. henselae* VNTRs (BHVs) were tested on 42 feline *B. henselae* isolates and strains from France (23 isolates), Denmark (17 isolates), the Philippines (one isolate) and the USA (F1 strain), on one human isolate from Germany, and on the H1 reference strain. These BHVs were sufficiently discriminatory to obtain 31 different profiles (corresponding to two different groups) among the 44 isolates and strains of *B. henselae* tested. Thirty-five profiles were obtained using these BHVs and two variant alleles. The combination of the five markers led to a diversity index of 0.98. The stability of the five BHVs was demonstrated on the feline F1 strain, with no change in stability observed after 2, 21 and 41 passages. This is believed to be the first study conducted on *B. henselae* typing using MLVA, and it demonstrates the high quality of this technique for discriminating between *B. henselae* isolates.

**INTRODUCTION**

*Bartonella* are fastidious haemotropic Gram-negative bacteria that are isolated from a wide range of mammals, including humans (Boulouis et al., 2005). The species *Bartonella henselae* is well recognized as a zoonotic agent, and several diseases have been associated with this species in humans and carnivores. Cat-scratch disease is the most benign clinical form of human infection (Chomel et al., 2004). Less-frequently occurring forms of infection have been described in immunocompromised patients; these infections include bacillary angiomatosis and parenchymal bacillary peliosis (Koehler, 1995), endocarditis (Bookman et al., 2004) and neuroretinitis (Depeyre et al., 2005). Several cases of *B. henselae* encephalopathy have been described in immunocompetent patients (Boulouis et al., 2005). Some of these human diseases have also been described in dogs (Kitchell et al., 2000; Pesavento et al., 2005) and cats (Chomel et al., 2003). On a yearly basis, about 22 000 cases of human infection occur in the USA (Jackson et al., 1993), and 2000 cases per year in the Netherlands (Bergmans et al., 1997); these prevalences are considered to be underestimated. Cats are the main reservoir of *B. henselae*, as they can remain bacteraemic for months. Depending on the cat category and the country of origin of the cat, prevalence of feline *B. henselae* infection varies from 1 % (Arvand et al., 2001) to 56 % (Chomel et al., 1999).
To date, there have been several attempts to develop molecular typing techniques for *B. henselae*. The first investigations led to the identification of two so-called ‘genotypes’ within the species *B. henselae* genotype I (*B. henselae* Houston-1 strain belongs to this genotype), and genotype II (Marseilles-like isolates) (Bergmans *et al.*, 1996; Drancourt *et al.*, 1996; La Scola *et al.*, 2002). Even when the difference between genotypes I and II is based solely on one nucleotide difference in the 16S rRNA sequence, studies have suggested that *B. henselae* genotypes may induce different pathological features in HIV-infected patients (Chang *et al.*, 2002). The two genotypes are not equally distributed in feline and human populations (Bergmans *et al.*, 1996; Chang *et al.*, 2002). Other attempts to develop more discriminatory typing techniques have been undertaken. Until recently, PFGE appeared to be the most discriminatory technique when compared with those developed earlier (Matar *et al.*, 1993; Rodriguez-Barradas *et al.*, 1995; Koehler *et al.*, 1997; Sander *et al.*, 1998; Handley & Regnery, 2000; Maruyama *et al.*, 2001; Arvand *et al.*, 2001; Dillon *et al.*, 2002).

More recently, two techniques based on the sequencing of selected genomic regions have been developed. The multilocus sequence technique (MLST) (Iredell *et al.*, 2003) has a discriminatory power equivalent to that of PFGE. The multispacer typing technique (MST) (Li *et al.*, 2006) appears to be a highly efficient method for genotyping *B. henselae* (39 profiles for 126 isolates). As these two techniques are based on sequencing, they are easily transferable; nevertheless, the necessity to sequence up to nine DNA regions (sometimes in both senses) for each isolate to be tested makes them costly, and not easily accessible for routine typing, for many potential users.

Variable number tandem repeats (VNTRs) represent good candidates for the development of a discriminatory, more user-friendly and less costly typing technique. VNTRs are microsatellite- or minisatellite-like structures initially described in eukaryotic cells. The combination of different VNTRs can lead to a good level of polymorphism (Vergnaud & Denoeud, 2000). This principle has been the basis for the development of multilocus VNTR analysis (MLVA). In a growing number of prokaryotes, including those displaying low genetic heterogeneity when using other techniques, VNTR typing has proved to be discriminatory, simple and transferable, with an excellent level of stability for the markers (Lindstedt, 2005). This type of epidemiological marker has been used for many pathogenic bacteria, such as *Bacillus anthracis*, *Yersinia pestis*, *Leptospira interrogans* and *Mycobacterium tuberculosis* (Keim *et al.*, 2000; Le Fleche *et al.*, 2001; Mazars *et al.*, 2001; Majed *et al.*, 2005; Slack *et al.*, 2005).

Despite the availability of *B. henselae*-typing techniques, several epidemiological issues have not yet been fully investigated. For example, it is important to link a human infection to a specific infection of a given potential animal carrier. One such study has been initiated by Chang *et al.* (2002), using the PFGE technique; however, a more user-friendly technique could help to investigate this link more easily. Comparison of isolates from different countries would help to establish their biodiversity and to determine whether some geographic markers can be identified, as suggested by the recent results of Li *et al.* (2006). Also, we would look for the presence of dominant isolates or clusters with potential epidemic significance, which has not been described to date (Li *et al.*, 2006). Furthermore, the role of fleas as selective carriers of *B. henselae* isolates remains unknown. In order to investigate these issues, new epidemiological markers are needed. The aim of this study was to establish the discriminatory power of 30 VNTR candidates. These potential VNTRs were selected from the sequence genome of the reference strain *B. henselae* H1, and they were tested on 44 isolates or strains of *B. henselae*, including *B. henselae* H1.

### METHODS

#### Isolates and DNA isolation.

Forty-four *B. henselae* isolates and strains were used: 23 isolates from French cats (belonging to the series studied by Gurfield *et al.* (2001)), 17 isolates from Danish cats (Chomel *et al.*, 2002), one isolate from a Philippino cat (Chomel *et al.*, 1999), one human isolate from Germany (Arvand *et al.*, 2001), the *B. henselae* reference strain Houston 1 (ATCC 49882), and the American feline type I strain F1 297172 (Yamamoto *et al.*, 2002). The characteristics of these isolates or strains according to their geographic origin and genotype are summarized in Table 1. All of the French isolates originated from the same region (Ile de France), whereas the Danish isolates were chosen because they had been collected from various areas in Denmark. The two single isolates, one from Germany, and one from the Philippinnes, were included to enlarge the geographical diversity of the tested isolates. The genotypes indicated in Table 1 are those that have been identified by the laboratories that provided the isolates, either using RFLP techniques (Chomel *et al.*, 2002), or 16S rRNA gene typing by PCR (Maruyama *et al.*, 2000). For the stability assay, we used the American feline type I strain F1 297172, which was passed 2, 21 and 41 times on rabbit blood agar medium (Gurfield *et al.*, 2001). All isolates and strains were cultivated on sheep blood agar medium (BioMérieux) for 4–5 days at 35 °C, with 5% CO2. All samples were prepared as DNA extracts, except those from the German isolate.

<table>
<thead>
<tr>
<th>Geographical origin</th>
<th>Host*</th>
<th>No. of isolates/strains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Genotype I</td>
</tr>
<tr>
<td>Denmark</td>
<td>F</td>
<td>1</td>
</tr>
<tr>
<td>France</td>
<td>F</td>
<td>7</td>
</tr>
<tr>
<td>Germany</td>
<td>H</td>
<td>1</td>
</tr>
<tr>
<td>The Philippines</td>
<td>F</td>
<td>1</td>
</tr>
<tr>
<td>USA (reference strain)</td>
<td>H</td>
<td>1</td>
</tr>
<tr>
<td>USA (passed strain)</td>
<td>F</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>–</td>
<td>12</td>
</tr>
</tbody>
</table>

*F, feline isolate; H, human isolate or strain.*
For DNA extraction, bacteria were scraped from the agar, and suspended in 500 μl sterile distilled water. These suspensions were boiled for 10 min, and centrifuged at 3000 g for 15 min. Nucleo Spin Tissue kit (Macherey-Nagel) was used for the purified DNA preparation, according to the manufacturer’s instructions. The German isolate was used as a cell lysate supernatant, after the lysate had been boiled and resuspended in distilled water. The DNA solutions and the cell lysate supernatant were stored at −20°C before testing.

**Computer analysis of repetitive DNA sequences for use as VNTR candidates.** The genomic DNA sequence of the *B. henselae* reference strain ATCC 49882 (available from NCBI) was screened for repetitive DNA sequences with the tandem repeats database developed by Le Flèche et al. (2001); the database is freely available at http://minisatellites.u-psud.fr. The following criteria were considered for application to the selection of VNTR candidates: (i) the absence of homology of a candidate sequence and of its flanking regions with sequences belonging to other genera; (ii) the length of the absence of homology of a candidate sequence and of its flanking sequences at http://minisatellites.u-psud.fr. The following criteria were consid-

**Amplification of VNTRs.** Amplification of VNTRs was conducted in a volume of 25 μl containing 1 μl purified DNA or 5 μl cell lysate, 1 × Pfx amplification buffer, 0.4 mM of each dNTP,1 mM MgSO4, 0.8 μM of each primer, 1 × PCR enhancer solution, and 1 unit Platinum Pfx DNA polymerase (Invitrogen). An initial denaturation step at 94°C for 5 min was followed by the following amplification programme: DNA was denatured for 30 s at 94°C, and primers were annealed for 30 s at the optimal temperature (50°C for BHV-A, BHV-B, BHV-C and BHV-D, and 53°C for BHV-E), and extended at 72°C for 1 min. After 40 cycles, there was a final extension programme at 72°C for 7 min.

PCR products were separated by gel electrophoresis in 1–2% agarose gels, and stained with ethidium bromide. Long gels (30 cm), long migration times (up to 27 h), and a combination of different molecular markers (from 100 bp to 1 kb), including FX174 HaeIII fragments (Invitrogen), were used for BHV-A and -B. These procedures allowed us to distinguish between two alleles, which could differ by as little as 22 bp (corresponding to 0.5 repeated unit for these two BHVs) even when these alleles were more than 1000 bp in length. For a given BHV, the expected PCR product length for the *B. henselae* H1 strain took into account the unit length, the number of units in the *B. henselae* H1 strain locus, and the length of the flanking sequences separating primer binding sites from BHV regions. The estimated size range for the different alleles was deduced from the allele size range, and from the basic unit length. For incomplete units, the calculated values were rounded up or down to the closest whole number; for instance, 13.5 or 13.8 units were considered to be 14 units, and 5.3 units was considered to be 5 units. In a second step, variant alleles (i.e. containing half-length units) that were initially rounded up (e.g. 14.5 units, which were rounded up to 15) were individualized when they provided additional information.

The final selection of BHVs was based on the following criteria: (i) candidate loci with less than two different alleles were excluded; (ii) when two BHV candidates were isomorphic, one was excluded; and (iii) the band patterns were fully readable.

**Data analysis.** For the evaluation of the discriminatory power of the selected BHVs, the Hunter and Gaston discrimination index (DI) was used (Hunter & Gaston, 1988), as recommended by the European Society of Clinical Microbiology and Infectious Diseases Study Group on Epidemiological Markers (Struelens, 1996). This index measures the probability that two isolates or strains, randomly chosen among the 44 isolates or strains selected, will have different types. It is defined by:

\[
DI = \frac{1}{N(N-1)} \sum_{j=1}^{S} nj(nj-1)
\]

where *N* is the number of isolates or strains, *S* is the total number of alleles, and *nj* is the number of isolates or strains with the allele *j* (Hunter & Gaston, 1988). Polymorphism is considered high when this index is higher than 95% (Struelens, 1996).

Clustering analysis was done using a phenetic approach, since the comparison of isolates and strains was based on small genomic sequences. The distance matrix was constructed by counting the number of different loci between isolates. With this method, the character states are considered to be unordered and, for a given BHV, the same weight is given to a small or a large difference of the number of repeats. Neighbour-joining (NJ) cluster analysis was then performed using PHYLIP (Felsenstein, 1989), with the isolate from the Philippines used as an outgroup.

**RESULTS**

**Selection of BHV candidates.** From the sequence of the reference strain *B. henselae* H1, 30 BHV candidates that fitted the selection criteria were identified. The percentage matches of tandem repeats varied from 57 to 97%. These 30 BHV candidates were further tested for evaluation of their polymorphism.

Using a first series of six isolates [three genotype I isolates (one French isolate, the Philippino isolate, and *B. henselae* H1) and three genotype II isolates (two French isolates, and one Danish isolate)], 11 out of 30 BHVs were found to be at least dimorphic and were considered for further testing.

All isolates or strains of *B. henselae*, with the exception of the American passed strain F1 297172 (Table 1), were tested with these 11 BHVs. Five BHVs (A–E) were finally selected based on the three criteria described above. BHV-D was located inside a gene, whereas the other four interfered with the beginning of coding sequences. The characteristics of the five selected BHVs are described in Table 2. For all isolates that were tested twice (more than 50% of our isolates), we obtained results that were 100% identical. In addition, in order to test the stability of our markers, the American feline strain F1 297172 was submitted to MLVA after 2, 21 and 41
passages, and the marker profiles were found to be the same as before for the five selected BHVs (A–E).

**BHV analysis**

**Genetic diversity.** A minimum of three, and a maximum of nine alleles per locus were observed with the five selected BHVs, generating a high level of polymorphism (Table 2). The discriminatory power of each BHV locus was estimated from the genetic DI values based on the number of alleles and their frequency. The high values for these individual DI s reflects their good potential as genetic diversity markers: 0.83 for BHV-A, 0.78 for BHV-B, 0.67 for BHV-C, 0.81 for BHV-D, and 0.72 for BHV-E. Moreover, the global DI value, when combining the five BHVs for the French and Danish isolates, was 0.98.

With the five BHVs, 29 different profiles were obtained for the 40 French and Danish isolates of *B. henselae* tested, corresponding to two different groups: group 1 and group 2 (Fig. 1). These two groups were distinguished at the level of the first bifurcation of the tree, which was just beyond the outgroup.

In addition, for BHV-A and BHV-B, we observed variant alleles (Fig. 2). Such variant alleles were not taken into account at the first screening level, which is reflected in the phenogram (Fig. 1).

**Comparison of isolates from different origins.** Eighteen BHV profiles were observed among the 23 French isolates, and 15 BHV profiles were observed among the 17 Danish isolates, of which four shared identical profiles. By comparison with the European isolates, the Philippino isolate was the only one that presented an original profile characterized by the presence of alleles at two different loci which were not observed in any of the other tested isolates (Fig. 1).

The Berlin isolate could not be distinguished from *B. henselae* H1 (reference strain) with the five BHV markers used (Fig. 1), or when using six additional BHVs (data not shown).

**Additional markers.** Within group 2 (Fig. 1), some isolates originating from France and Denmark looked identical when they were compared using the five BHVs, with rounded-up or -down values. However, when using the complementary tool represented by the variant alleles defined for BHV-A and -B, the group of four isolates was resolved into two groups of two isolates. In the same way, the two groups of three isolates were resolved into two groups of two isolates and two single isolates. Eventually, one pair was resolved into two different isolates (FR97 and FR123). Only seven identical pairs remained. Except for one pair (FR963 and DK229), all remaining pairs comprised isolates from the same country (French/French or Danish/Danish). In one of these four pairs the isolates (FR 74 and FR 76; Fig. 1) were from the same cat.

---

**Table 2. Characteristics of the repeat motifs of the selected BHV and the corresponding forward and reverse primers**

<table>
<thead>
<tr>
<th>Potential VNTR name*</th>
<th>Unit length (bp)</th>
<th>GC content (% of motif)</th>
<th>Conservation (% of motif)</th>
<th>Primers sequences for VNTR (5'→3'†)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHV-A (1698439)</td>
<td>14</td>
<td>52</td>
<td>95</td>
<td>F: AAATCAACACTCTCAAAAACACAAG</td>
</tr>
<tr>
<td>BHV-B (1651759)</td>
<td>20</td>
<td>84</td>
<td>97</td>
<td>R: TGCTCTGCTTC TTTGCTCA</td>
</tr>
<tr>
<td>BHV-C (1644993)</td>
<td>10</td>
<td>141</td>
<td>97</td>
<td>F: ATTCCTGAGACCTTAGTGATT</td>
</tr>
<tr>
<td>BHV-D (1482978)</td>
<td>10</td>
<td>141</td>
<td>97</td>
<td>R: GGTGATAAAGCATTCCATAA</td>
</tr>
<tr>
<td>BHV-E (1678645)</td>
<td>5</td>
<td>863</td>
<td>96</td>
<td>R: TGCTGATAGTGCGGTTTTC</td>
</tr>
</tbody>
</table>

*The position of the BHV first base in the *B. henselae* reference strain Houston 1 (ATCC 49882) is given in parentheses.
†F, forward; R, reverse.
‡Values in square brackets correspond to the allele sizes when variant alleles for BHV-A and BHV-B are not rounded up to the next whole number.
Comparison of genotypes I and II. In the phenogram, the smallest group (5 out of 44) was completely composed of strains and isolates belonging to genotype I, i.e. 45% of the isolates and strains harbouring this genotype. The five other genotype I French isolates did not appear to be randomly scattered, but formed two small subgroups within the second group, which included all the isolates belonging to genotype II. The American type I feline strain F1 (297172) was totally out of the tree.

DISCUSSION

To the best of our knowledge, this is the first study of *B. henselae* VNTR typing. Thirty BHV candidates were tested. As few as five BHVs were sufficient to obtain a very high level of polymorphism. Of the 44 isolates and strains tested, including the ATCC 49882 human strain Houston 1, a human isolate from Germany, a feline Philippino isolate, and an American feline strain, 31 different profiles were obtained (35 profiles were obtained with variant alleles). The Hunter and Gaston DI was always higher than 98%, i.e. clearly above the cut-off of 95%, either when considering all isolates together, or when considering isolates from one country only (France or Denmark, data not shown). The limited number of BHVs needed, and the standardized PCR conditions, in association with the very good reproducibility, made this technique accurate, very user friendly, and...
easily transferable. In terms of allelic stability, limited indication was provided by the observation that two isolates from a same cat (FR74 and FR76) displayed the same MLVA type with the five markers. Stability was further demonstrated by the fact that for the strain F1 297172 all the alleles remained unchanged for the five BHVs after 21 and 41 passages.

The discriminatory power of this MLVA technique, using only five BHVs, appears to be high in comparison with the results obtained with the majority of the typing techniques proposed for \( B. \) \textit{henselae}. The performance of MLST has been convincingly demonstrated to be comparable with that of PFGE by Iredell \textit{et al.} (2003), using the same isolates for testing both techniques. In their study, Li \textit{et al.} (2006) conclude that MST is more discriminatory than MLST. Some findings of our study suggest that MLVA is even more discriminatory than MST, at least when considering French isolates: Li \textit{et al.} (2006) tested 28 French isolates, and obtained five different profiles using MST; in our study, 23 French isolates from the same area (Ile de France) were tested using MLVA, and we obtained 18 different profiles. Since Li \textit{et al.} (2006) reported more heterogeneity among their Asian isolates than among their European isolates, it would be interesting to check if the apparent superiority of MLVA is confirmed when testing isolates from locations other than France and Denmark.

Furthermore, the present technique is even more discriminatory when using variant alleles for BHV-A and -B as complementary tools. Other authors, such as Farlow \textit{et al.} (2002), Pourcel \textit{et al.} (2003), and Johansson \textit{et al.} (2004), have described the presence of variant alleles for \textit{Borrelia}, \textit{Legionella} and \textit{Francisella}, respectively. These complementary tools have been useful in a few cases, when more accurate discrimination between isolates, especially French versus Danish isolates, was required.

Using an isolate from the Philippines – which had MLVA patterns very different from the European isolates – as an outgroup, two main groups were detected in the NJ tree. Isolates from France and Denmark grouped together in the majority of the cases; this is not surprising, as they belong to the same continent. A similar trend has already been observed with other bacterial species, e.g. \textit{Mycobacterium bovis}, when using MLVA (Skuce \textit{et al.}, 2002) and spoligotyping (Haddad \textit{et al.}, 2001).

When compared with the European isolates, the isolate from the Philippines, and the American feline strain F1 297172, presented original profiles. This suggests that new groups could emerge from BHV typing of non-European isolates, as observed by Li \textit{et al.} (2006) using MST. In contrast, the human isolate from Berlin was identical to \textit{B. henselae} H1 strain (human ATCC 49882 strain), despite their very different geographic origins, suggesting a possible common origin. Such a similarity between these two human isolates/strains has been described by Arvand \textit{et al.} (2001) using the PFGE technique.

There was no clear cut-off between the two \textit{B. henselae} genotypes in terms of BHV patterns. Similar observations have been made with techniques used by others (Chang \textit{et al.}, 2002; Dillon \textit{et al.}, 2002). The absence of a correlation between genotypes I and II, and the distribution of the isolates according to their MLVA profiles, is not surprising. First, discrimination between genotypes I and II is based on a unique difference of one nucleotide in the sequence of the 16S rRNA gene; thus, the phylogenic meaning of the discrimination is questionable. Second, phenograms are based on only a few informative characters, and, therefore, a difference at a single locus can change the positioning of isolates in the phenogram. Third, for a given allele, even if it is stable in a given period, a reduction or an increase in the numbers of repeated units can occur with time. This makes it possible that two isolates harbouring different I/II genotypes can evolve to closer or even identical BHV profiles, and become included in the same group (homoplasy); this could be the case for the genotype I isolates located within the group that almost otherwise consists of genotype II isolates (Fig. 1). Eventually, some modifications of BHV profiles may be linked to horizontal transfer, the occurrence of which in \textit{B. henselae} has been suspected by other authors (Dillon \textit{et al.}, 2002; Iredell \textit{et al.}, 2003). Nevertheless, 5 of the 12 isolates belonging to genotype I form a clearly individualized group (group 1). Such a distribution needs to be further confirmed using a larger collection of isolates belonging to genotype I.

In conclusion, despite the small size of the genome of \textit{B. henselae}, 11 BHV candidates revealed some degree of polymorphism. Among them, five BHVs were selected because their combination provided a high level of polymorphism, both at the level of all the 44 \textit{B. henselae} isolates tested (DI 0.98), and at the level of French or Danish isolates measured separately. Compared with the majority of the other techniques developed for \textit{B. henselae} typing, MLVA appears to be very discriminatory. However, this needs to be confirmed by comparing the levels of heterogeneity obtained for the same isolates using MLVA and other techniques, especially PFGE, MLST and/or MST.

\section*{ACKNOWLEDGEMENTS}

The authors thank Hans Petersen (Hinnerup Animal Hospital, Dyr- lægegruppen Frijsenborg, Vestergade 3, 8382 Hinnerup, Denmark) for providing the isolates from Denmark, Rick Kasten (Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA 95616, USA) for providing the American passed strain F1 297172, and Corinne Bouillin and Christelle Gandoin for their technical assistance.

\section*{REFERENCES}


Edited by: W. B. van Leeuwen