Curr Microbiol (2008) 56:447–452
DOI 10.1007/s00284-008-9104-1

Variable-Number-of-Tandem-Repeats Analysis of Genetic Diversity in *Pasteuria ramosa*

L. Mouton · D. Ebert

Received: 10 September 2007 / Accepted: 30 November 2007 / Published online: 24 January 2008

Abstract Variable-number-of-tandem-repeats (VNTR) markers are increasingly being used in population genetic studies of bacteria. They were recently developed for *Pasteuria ramosa*, an endobacterium that infects *Daphnia* species. In the present study, we genotyped *P. ramosa* in 18 infected hosts from the United Kingdom, Belgium, and two lakes in the United States using seven VNTR markers. Two *Daphnia* species were collected: *D. magna* and *D. dentifera*. Six loci showed length polymorphism, with as many as five alleles identified for a single locus. Similarity coefficient calculations showed that the extent of genetic variation between pairs of isolates within populations differed according to the population, but it was always less than the genetic distances among populations. Analysis of the genetic distances performed using principal component analysis revealed strong clustering by location of origin, but not by host *Daphnia* species. Our study demonstrated that the VNTR markers available for *P. ramosa* are informative in revealing genetic differences within and among populations and may therefore become an important tool for providing detailed analysis of population genetics and epidemiology.

L. Mouton (✉) · D. Ebert
Evolutionsbiologie, Zoologisches Institut der Universitat Basel, Basel, Switzerland
e-mail: mouton@biomserv.univ-lyon1.fr

L. Mouton
Laboratoire de Biometrie et Biologie Evolutive, Universite de Lyon, Villeurbanne, France

Introduction

Bacteria of the genus *Pasteuria*, endospore-forming, Gram-positive bacteria members of the *Bacillus–Clostridium* clade, are common parasites of nematodes and crustaceans. They can be highly virulent, such as *P. ramosa*, a pathogen of the freshwater crustacean *Daphnia* [12]. These bacteria most frequently infect *Daphnia magna* but are also found in *D. pulex*, *D. longispina*, and *D. dentifera* [10]. They have been recorded in Europe and North America, with a prevalence $\leq 80\%$ [9]. *P. ramosa* are obligate extracellular parasites that grow in the body cavity of their hosts and produce several million endospores. Infected hosts suffer greatly decreased fecundity [11], which influences their population dynamics [6]. Transmission is strictly horizontal (waterborne) through spores released from the remains of dead hosts. These spores can remain dormant in pond sediment for decades [7].

Multiple-locus variable-number-of-tandem-repeats (VNTR) analysis is the method of choice for variability studies of many bacterial species, especially pathogens [16, 17, 19, 20 (review), 21, 26]. Indeed, VNTRs often provide a high level of discriminatory power for strain differentiation because of their high mutability [27]. Recently, VNTR markers were developed in *P. ramosa* [22]. In that study, genetic diversity was estimated based on sediment samples from different locations. A large degree of polymorphism was revealed within the material surveyed, indicating that these markers may constitute good tools for epidemiologic studies of natural populations. However, a better understanding of their effectiveness in determining the level of genetic diversity among *P. ramosa* isolates is a prerequisite for more effective survey.

In the present article, seven polymorphic VNTR markers were applied to study the genetic relationships among *P.*

ramosa isolates found in field populations of *Daphnia*. Two species of *Daphnia* were collected, *D. dentifera* and *D. magna*, in three different geographic locations: the United Kingdom, Belgium, and the United States. Genetic diversity and relationships among populations were estimated using three similarity coefficient calculations: the Jaccard, Dice, and simple match coefficients [18]. Principal component analysis (PCA) showed that the VNTRs are able to discriminate between populations of *P. ramosa*. The *Pasteuria–Daphnia* system has been a model for many experimental studies of host–parasite interactions [10]. It is therefore of great interest to develop valuable molecular markers that provide information on the epidemiology, diversity, and evolution of *P. ramosa* in wild populations.

Materials and Methods

Strains

DNA of *P. ramosa* was extracted from 18 infected *Daphnia* collected in the field. Two *Daphnia* species were collected: *Daphnia magna* and *D. dentifera*. Infected *D. dentifera* came from two lakes in Michigan: Sherman Lake (Barry county; latitude 42.35°N, 85.39°W) and Pine Lake (Kalamazoo county; latitude 42.57°N, 85.39°W). Infected *D. magna* were collected in one pond in each in Belgium and the United Kingdom (Table 1).

DNA Extraction and PCR Amplification

Bacterial DNA was extracted using the EZNA Tissue DNA kit (Peqlab, Erlangen, Germany). *Daphnia* were individually crushed in microtubes containing 200 µl TL buffer and 25 µl proteinase K and maintained for 2 to 3 hours at 55°C. Final elution volume was 100 µl. Polymerase chain reaction (PCR) amplification of seven VNTR loci [22] was carried out using seven primer sets (Table 2) in 25–µl volumes containing 1× PCR buffer [TrisCl, KCl, $(NH_4)_2SO_4$, 15 mM $MgCl_2$; pH 8,7], 200 µM deoxyribonucleoside triphosphate, 200 nM primers, 0.5 U HotStarTaq DNA polymerase (Qiagen GmbH, Germany), and 2 µl template DNA. Thermal cycling conditions were 15 minutes at 94°C followed by 40 cycles of 30 seconds at 94°C, 30 seconds at 50°C, and 1 minute at 72°C. Final elongation was 10 minutes at 72°C. Negative controls were done on DNA extracted from uninfected *Daphnia magna* to ensure that the VNTR amplicons were derived from the *P. ramosa* DNA.

Automated Genotype Analysis

The PCR products were electrophoretically analysed with an ABI 310 DNA sequencer. Genetic Analyser and Genemapper software packages (Applied Biosystem, Warrington, UK) were used to analyse sample sizes as described in Mouton et al. [22].

Data Analysis

Genetic similarities and relationships among the samples were estimated by multivariate analysis based on the presence or absence of each detected allele at each locus. A similarity matrix was calculated from these data based on the proportion of shared alleles across all loci. The genetic dissimilarity between pairs of individuals was estimated for all possible pair-wise comparisons within and between populations using three different coefficients: the Jaccard, Dice, [23], and simple match coefficients [25]. We denoted as "A" the number of alleles common to both individuals and populations, as "B" the number of alleles unique to the first individual and population (i_1), as "C" the number of alleles unique to the second individual and population (i_2), and as "n" the total number of alleles that appeared in at least one of the two individuals and populations (i_1 and i_2). Coefficients of dissimilarity between individuals and populations i_1 and i_2 could then be calculated as follows [18 (review)]:

Jaccard's coefficient: $j(i1,\ i2) = (B + C)/(A + B + C)$

Dice's coefficient: $d(i1,\ i2) = (B + C)/(2A + B + C)$

Simple match coefficient $m(i1,\ i2) = (B + C)/n$.

In addition, the VNTR marker data for the 18 daphnia samples were scored for the two first principal components,

Table 1 *P. ramosa* samples

Population	Host species	Geographic region	Sampling Year	No. of infected hosts tested
BE OM1	*Daphnia magna*	Heverlee, Belgium, Pond OM1	2006	3
UK K1	*Daphnia magna*	Kames, UK, Pond 1	2006	5
AM Sh	*Daphnia dentifera*	USA, Michigan, Sherman Lake	2005	5
AM Pi	*Daphnia dentifera*	USA, Michigan, Pine Lake	2005	5

Table 2 Repetitive DNA PCR primer sequences and attributes

Primer names	Primer sequences	Locus name	Repeat motif	Size range (nt)[a]	Smallest-largest no. of repetitions[b]	No. of alleles[c]
Pr1 fwd	ACCTAAAGAACAGGAATATCTGGA	Pr SSR1	AAACTAACA	195–276	3–11	4
Pr1 rev	GCATGGAATGATTTTTGCTG					
Pr2 fwd	CTGCTGGATGGATGGACTACGTGA	**Pr SSR2.1**	CCTGGTAAA	259–286	3–4	2
Pr2 rev	ACCGGTCCCGTAGGTATAGG	**Pr SSR2.2**	CATCCTGGTGGTCCTTGG		2–3	
Pr3 fwd	GGACCAATCGAACCAGGTAT	**Pr SSR3**	TATCTCCTTTAGGACCAG	365–392	6–9	3
Pr3 rev	AACGGTTTCTTCGCTTGTTG					
Pr4 fwd	GGTAACCCTGGATGTCCTGA	Pr SSR4	TT(A/G)CTTTA*	321–369	9–15	3
Pr4 rev	ATCCCGTTACAAATGGGACA					
Pr7 fwd	AACGTACTGACAAACCAAACCA	Pr SSR7	AACAACC(T/C)C*	109–172	4–11	5
Pr7 rev	AATTTTTCTTAGATTGCTAGGTTGA					
Pr8 fwd	GCATCAAATACAAAAACAAATGAAG	Pr SSR8	AGAATATGAAGAAGATGC	404–440	5–7	3
Pr8 rev	TGTTTCTCTCGCGTTTCCTT					
Pr9 fwd	ATACGACGAACGGAACAAGA	Pr SSR9	AGCAACAAC	178	5	1
Pr9 rev	AACCAAAGAATTAACGCCATT					

VNTRs located within putative coding region are indicated in bold. No significant similarities were found when compared the corresponding amino acid sequences with protein sequence databases at the NCBI database [22]

* Imperfect repetitions. The two nucleotides in the parenthesis can be found

[a] Size range of amplicons

[b] The smallest and the largest number of repetitions found in the samples

[c] Number of alleles in the samples

and PCA was plotted using Multi-Variate Statistical Package Version 3.1 (Kovach Computing Services).

Results

VNTR Amplification

Table 3 lists the alleles present at the seven primer sets for each of the 18 samples screened. A summary of the alleles' size-range and the corresponding numbers of repeats are listed in Table 2. For all these primers, PCR allowed amplification of the field samples irrespective of the host, *D. magna* or *D. dentifera*, indicating that the VNTR regions selected for typing were conserved among the *P. ramosa* that infected different *Daphnia* species. However, we failed to amplify some samples from the United States with the primer set Pr3, which may be explained by mutations in the binding regions of the primers, resulting in the loss of the PCR product (null alleles). Except for Pr SSR9, all markers proved to be effective in discriminating the isolates analysed, especially Pr SSR1, for which each population had a specific allele. However, the extent of polymorphism was variable, with as many as 5 alleles/locus observed in Pr SSR7. For this locus, some individuals harbour two different alleles, which may reflect the simultaneous presence of > 1 *P. ramosa* clones within a host individual. However, we cannot fully exclude the existence of a technical artefact caused by PCR stutter or the instability of this marker because repetitive DNA sequences has been described in other repetitive DNA in eukaryotes and prokaryotes [1, 2]. Further investigations are necessary to test these hypotheses.

Genetic Similarities and Relationships Among Isolates

Genetic distances of *P. ramosa* between and within the four infected host populations sampled were estimated using dissimilarity coefficients based on the proportion of shared alleles of *P. ramosa* among *Daphnia* individuals across the seven VNTR loci. According to Kosman and Leonard [18], Dice's coefficient is the more appropriate measure for studying similarity between haploid individuals with VNTR markers. However, two other coefficients of dissimilarity commonly used for comparison between individuals, the Jaccard and simple match coefficients, were also calculated to test the robustness of the comparisons (Table 4). These three indices gave similar results (Spearman order rank correlation coefficient, rho = 1),

Table 3 Amplicons size (nt) of the 18 infected host individuals from four populations

Population[a]	Individual number	Primer						
		Pr1	Pr2	Pr3	Pr4	Pr7	Pr8	Pr9
BE OM1	1	276	286	383	321	163	404	178
BE OM1	2	276	286	383	345	163	404	178
BE OM1	3	276	286	383	345	163	404	178
UK K1	1	240	259	365	369	163	404	178
UK K1	3	240	259	365	369	163	404	178
						172		
UK K1	4	240	259	365	369	163	404	178
UK K1	5	240	259	365	369	163	404	178
UK K1	7	240	259	365	369	163	404	178
AM Sh	1	249	259	365	321	145	422	178
						154		
AM Sh	2	249	259	365	321	145	422	178
						154		
AM Sh	3	249	259	365	321	145	422	178
						154		
AM Sh	4	249	259	365	321	145	422	178
AM Sh	5	249	259	365	321	154	422	178
AM Pi	1	195	286	x	345	163	440	178
AM Pi	2	195	259	x	345	163	440	178
AM Pi	3	195	259	365	345	163	440	178
AM Pi	4	195	286	392	345	109	440	178
						163		
AM Pi	5	195	259	365	345	109	440	178

x: Amplification failed

[a] for abbreviations see Table 1

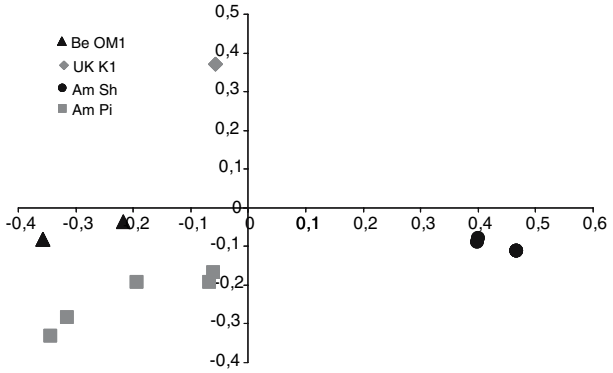

Fig. 1 Principal Component Analysis (PCA) plots from SSR data The PCA plots are components 1 (horizontal axis) vs 2 (vertical axis). Samples with identical VNTRs pattern overlap. Therefore less points than sampled individuals are visible

indicating the reliability of the conclusions. The mean distances between isolate pairs within the same population were smaller than the distances between isolates from different populations. The extent of genetic diversity varied among populations (Kruskal-Wallis, $H = 16.1$, $P = 0.0018$ for Dice's coefficient), with the United Kingdom population having the lowest and the American population from Pine Lake the highest level of dissimilarity, i.e., individual infections were the least similar within the AM Pi population.

PCA performed on the 18 samples reflected a strong clustering by location of origin but did not show clustering by host *Daphnia* species (Fig. 1). Data are shown for the first two components axes, representing 40.2% and 27.2% of the total variation, respectively. Belgium and United Kingdom bacteria were collected from the same *Daphnia* species, *D. magna*, whereas bacteria from the two American populations were isolated from *D. dentifera*. Data also showed that the populations collected in two different lakes in United States were distinct.

Discussion

Recently, the first molecular markers discriminating *P. ramosa* isolates were described [22]. To date, the only way to classify *Pasteuria* strains was through their infection

Table 4 Genetic dissimilarities between *P. ramosa* samples

Comparison		n	Jaccard 's coefficient	Dice's coefficient	Simple match coefficient
Within populations	BE OM1	3	0.17 ± 0.08	0.10 ± 0.05	0.17 ± 0.08
	UK K1	10	0.05 ± 0.02	0.03 ± 0.01	0.05 ± 0.02
	AM Sh	10	0.10 ± 0.03	0.05 ± 0.01	0.10 ± 0.03
	AM Pi	10	0.31 ± 0.05	0.20 ± 0.04	0.31 ± 0.05
Between populations	BE OM1 / UK K1		0.77	0.63	0.77
	BE OM1 / AM Sh		0.86	0.75	0.86
	BE OM1 / AM Pi		0.71	0.56	0.71
	UK K1 / AM Sh		0.77	0.63	0.77
	UK K1 / AM Pi		0.71	0.56	0.71
	AM Sh / AM Pi		0.80	0.67	0.80

The genetic distances within and between *P. ramosa* populations were estimated on the basis of the Jaccard, Dice and simple match dissimilarity coefficients as described in the material and methods section. For within comparisons, number of pairwise comparisons (n), means and standard errors (±) are indicated

profile [3, 5, 24]. In this article, we used available VNTR markers to evaluate genetic diversity and relationships of *P. ramosa* in infected *Daphnia* individuals and populations from the United Kingdom, Belgium, and the United States. This study of *P. ramosa* epidemiology showed that these markers are informative in revealing genetic differences among populations.

The genetic distance between pairs of *P. ramosa* isolates from individual *Daphnia* at the four geographic locations was estimated on the basis of three similarity coefficients (the Jaccard, Dice, and simple match coefficients). The values obtained with these coefficients showed that the extent of genetic diversity among populations varies, but that in all cases, genetic similarity is higher within than between populations. Allozyme analysis of *D. magna* populations revealed strong genetic divergences in relation to geographic distance [13]. Therefore, these results could reflect a specialization of parasites to the local host genotypes, which is consistent with previous data showing specific *Daphnia* clone–*P. ramosa* isolate interactions [3, 8]. A geographic pattern of population differentiation was revealed by PCA. However, the influence of geographic distance between populations is difficult to test here because the *P. ramosa* isolates came from two host species: *D. magna* and *D. dentifera*. Surprisingly, the *P. ramosa* infecting these two different *Daphnia* species showed no split, suggesting that the *Daphnia* species does not affect the genetic diversity of *P. ramosa;* however, more investigations are needed to clarify this question. Although the four populations showed clear genetic differences, there was also some within-population variation among bacteria from different host individuals. Large within-population variation has been reported before for *P. ramosa*, but it was based on phenotypic infection profiles [3, 13] and thus attributed to local coevolution between hosts and parasites.

We demonstrated that VNTRs are a useful resource for studying genetic diversity in *P. ramosa.* Therefore, this method can be used in epidemiologic and phylogenetic studies. It has been suggested that parasite genetic variability is a driving force of host and parasite evolution and host–parasite arms races [14, 15]. Moreover, the effect of parasite infection on host reproduction and survival, combined with genetic variation in the host–parasite association, might be responsible for coevolution based on frequency-dependent selection [3]. Investigating the genetic variability of hosts (microsatellite markers have been developed in *D. pulex* [4] and *D. magna* (I. Colson, in preparation) and their parasites is a way to test the hypothesis of coevolution between *P. ramosa* and their hosts.

Acknowledgments We thank I. Colson and L. Du Pasquier for useful discussion and comments, B. Aeschbach for technical help, and T. Little, E. Decaestecker, and S. Hall for providing samples. S. Zweizig and I. Colson improved the English language of the manuscript. This work was supported by grants from the Swiss Nationalfonds and the Freiwillige Akademische Geselschaft, Basel, Switzerland.

References

1. Bois PR (2003) Hypermutable minisatellites, a human affair? Genomics 81:349–355
2. Bzymek M, Lovett ST (2001) Instability of repetitive DNA sequences: The role of replication in multiple mechanisms. Proc Natl Acad Sci U S A 98:8319–8325
3. Carius HJ, Little T, Ebert D (2001) Genetic variation in a host-parasite association: Potential for coevolution and frequency-dependent selection. Evolution 55:1146–1152
4. Colbourne JK, Robison B, Bogart K, Lynch M (2004) Five hundred and twenty-eight microsatellite markers for ecological genomic investigations using *Daphnia*. Mol Ecol Notes 3:485–490
5. Davies KG, Redden M (1997) Diversity and partial characterization of putative virulence determinants in *Pasteuria penetrans*, the hyperparasitic bacterium of root-knot nematodes (*Meloidogyne* spp.). J Appl Microbiol 83:227–235
6. Decaestecker E, Declerck S, De Meester L, Ebert D (2005) Ecological implications of parasites in natural *Daphnia* populations. Oecologia 144:382–390
7. Decaestecker E, Lefever C, De Meester L, Ebert D (2004) Haunted by the past: Evidence for dormant stage banks of microparasites and epibionts of *Daphnia*. Limnol Oceanogr 49:1355–1364
8. Decaestecker E, Vergote A, Ebert D, De Meester L (2003) Evidence for strong host-clone-parasite interactions in the *Daphnia* microparasite system. Evolution 57:784–792
9. Duncan AB, Mitchell SE, Little TJ (2006) Parasite-mediated selection and the role of sex and diapause in *Daphnia*. J Evol Biol 19:1183–1189
10. Ebert D (2005) Ecology, epidemiology, and evolution of parasitism in *Daphnia*. Bethesda, MD: National Library of Medicine, National Center for Biotechnology Information. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=books. Accessed:
11. Ebert D, Carius H-J, Little TJ, Decaestecker E (2004) The evolution of virulence when parasites cause host castration and gigantism. Am Nat 164(Suppl.):S19–S32
12. Ebert D, Rainey P, Embley TM, Scholz D (1996) Development, life cycle, ultrastructure and phylogenetic position of *Pasteuria ramosa* Metchnikoff 1888: Rediscovery of an obligate endoparasite of *Daphnia magna* Straus. Phil Trans Royal Soc London B 351:1689–1701
13. Ebert D, Zshokke-Rohringer CD, Carius HJ (1998) Within- and between-population variation for resistance of *Daphnia magna* to the bacterial endoparasites *Pasteuria ramosa*. Proc R Soc Lond B Biol Sci 265:2127–2134
14. Haldane JBS (1949) Disease and evolution. Ric Sci 19:2–11
15. Hamilton WD (1980) Sex versus non-sex versus parasite. Oikos 35:282–290
16. Keim P, Price LB, Klevytska AM, Smith KL, Schupp JM, Okinaka R et al (2000) Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J Bacteriol 182:2928–2936
17. Klevytska AM, Brice LB, Schupp JM, Worsham PL, Wong J, Keim P (2001) Identification and characterization of variable-number of tandem repeats in the *Yersinia pestis* genome. J Clin Microbiol 39:3179–3185

18. Kosman E, Leonard KJ (2005) Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. Mol Ecol 14:415–424
19. Liao J-C, Li C-C, Chiou C-S (2006) Use of multilocus variable-number tandem repeat analysis method for molecular subtyping and phylogenetic analysis *Neisseria meningitides* isolates. BMC Microbiol 6:44
20. Lindstedt B-A (2005) Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. Electrophoresis 26:2567–2582
21. Maiden MC, Bygraves JA, Feil E, Morelli G, Russel JE, Urwin R et al (1998) Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A 95:3140–3145
22. Mouton L, Nong G, Preston J, Ebert D (2007) Identification and characterization of variable number tandem repeats in an obligate endoparasite of *Daphnia* species. Appl Environ Microbiol 73:3715–3718
23. Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleasis. Proc Natl Acad Sci U S A 76:5269–5273
24. Preston JF, Maruniak JE, Nong G, Brito JA, Schmidt LM, Giblin-Davis RM (2003) *Pasteuria spp.*: Systematics and phylogeny of these bacterial parasites of phytopathogenic nematodes. J Nematol 35:198–207
25. Sneath PA, Sokal RR (1973) Numerical taxonomy. Freeman, San Francisco, CA
26. Top J, Schouh LM, Bonten MJ, Willems RJ (2004) Multilocus variable number of tandem repeats analysis, a novel typing scheme to study the genetic relatedness and epidemiology of *Enterococcus faecium* isolates. J Clin Microbiol 42:4503–4511
27. van Belkum A, Scherer S, van Alphen L, Verbrugh H (1998) Short-sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev 62:275–293