1-2020

Gene duplication and deletion, not horizontal transfer, drove intra-species mosaicism of Bartonella henselae

Rachana Banerjee
CSIR-Indian Institute of Chemical Biology

Oshina Shine
Amrita Vishwa Vidyapeetham

Vyshakh Rajachandran
Amrita Vishwa Vidyapeetham

Govind Krishnadas
Amrita Vishwa Vidyapeetham

Michael F. Minnick
University of Montana, Missoula

See next page for additional authors

Genomics 112 (2020) 467–471

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Original Article

Gene duplication and deletion, not horizontal transfer, drove intra-species mosaicism of *Bartonella henselae*

Rachana Banerjee[a], Oshina Shine[b], Vyshakh Rajachandran[b], Govind Krishnadas[b], Michael F. Minnick[c], Sandip Paul[a], Sujay Chattopadhyay[b,*,1]

[a] Structural Biology and Bioinformatics Division, CSIR-Indian Institute of Chemical Biology, Kolkata, India
[b] School of Biotechnology, Amrita Vishwa Vidyapeetham, Kollam 690 525, Kerala, India
[c] Division of Biological Sciences, University of Montana, Missoula, MT, USA

ARTICLE INFO

Keywords:
Pan-genome
Bartonella henselae
Facultative intracellular human pathogen
Gene duplication
Gene deletion
Gene mosaicism in bacterial genomes

ABSTRACT

Bartonella henselae is a facultative intracellular pathogen that occurs worldwide and is responsible primarily for cat-scratch disease in young people and bacillary angiomatosis in immunocompromised patients. The principal source of genome-level diversity that contributes to *B. henselae*'s host-adaptive features is thought to be horizontal gene transfer events. However, our analyses did not reveal the acquisition of horizontally-transferred islands in *B. henselae* after its divergence from other *Bartonella*. Rather, diversity in gene content and genome size was apparently acquired through two alternative mechanisms, including deletion and, more predominantly, duplication of genes. Interestingly, a majority of these events occurred in regions that were horizontally transferred long before *B. henselae*'s divergence from other *Bartonella* species. Our study indicates the possibility that gene duplication, in response to positive selection pressures in specific clones of *B. henselae*, might be linked to the pathogen's adaptation to arthropod vectors, the cat reservoir, or humans as incidental host-species.

1. Introduction

The *Bartonella* genus consists of roughly forty-five species, 13 of which are implicated in a diverse array of emerging zoonoses that occur worldwide [1,2]. Pathogenic *Bartonella* are typically transmitted from a mammalian reservoir species (e.g., cats, dogs, rodents) to humans through bites or contaminated feces of hematophagous arthropods, such as ticks, lice, fleas and sand flies. *Bartonella henselae* is arguably the most frequent cause of *Bartonella* infections in humans and is transmitted by animal bites or scratches. *B. henselae* is the etiologic agent of cat-scratch disease (CSD) in children and bacillary angiomatosis in immune-suppressed individuals [3,4]. It is also a causative agent of a febrile and paucisymptomatic bacteremia and endocarditis in humans [3]. *B. henselae* is a facultative intracellular pathogen and has been reported to show several characteristic features of host adaptation [5,6]. Genetic diversity is the crucial factor governing genome evolution, and it is largely directed in pathogens by the corresponding host niche [5,7]. It has been reported that gene variation in *B. henselae* is associated with extensive rearrangements and DNA amplifications in certain strains [8]. The evolution of *Bartonella quintana* and *Bartonella koehlerae*, close relatives of *B. henselae*, is associated with extensive gene deletions and partial retention of genomic islands, respectively, suggesting similar scenarios occurred in the intra-species evolution of *B. henselae* [8–10].

Additional mechanisms such as gene recombination and horizontal gene transfer (HGT) are important players in bacterial evolution [5]. Previous analyses have highlighted the relatively minor role of non-homologous recombination events compared to mutations in the evolution of *B. henselae* [7]. HGTs were also studied in *B. henselae* with different host preferences [5,7]. However, the role of HGT as a primary player in the genome evolution of *B. henselae* has not been clearly established. In addition, previous studies reported minimal novel gene acquisitions, indicative of a closed pan-genome in *B. henselae* [7]. On the other hand, genes coding for surface proteins were reported to evolve rapidly by duplication, deletion, nucleotide substitution and recombination processes [5,7]. In consideration of the diverse array of mammalian reservoirs, arthropod vectors and the human host, pathogenic bartonellae like *B. henselae* provide exceptional models for examining genomic evolution in response to various selective pressures. In this study, we performed a comparative pan-genomic profiling of the four available *B. henselae* genomes to define the relative genome-wide contribution of HGT vs. gene deletions and duplications after

* Corresponding author.
E-mail address: chatsujay@gmail.com (S. Chattopadhyay).
[1] Present address: Sujay Chattopadhyay, JIS Institute of Advanced Studies & Research, Kolkata, India.

https://doi.org/10.1016/j.ygeno.2019.03.009
Received 19 December 2018; Received in revised form 24 February 2019; Accepted 18 March 2019
Available online 19 March 2019

Table 1
List of analyzed *B. henselae* isolates along with their 16S ribosomal RNA and MLST profiles.

B. henselae strains	Country of isolation	Genome size (bp)	16S rRNA genotype	*batR*	*ftsZ*	*gltA*	*groEL*	*nlpD*	*ribC*	*rpoB*	Sequence Type (ST)
Houston – 1	USA	1,931,047	1	1	1	1	1	1	1	1	1
BM1374163 (MVT01)	France	1,905,383	2	1	1	1	1	1	1	1	9
MVT02	France	1,905,383	2	1	1	1	1	1	1	1	9
BM1374165 (MVT03)	France	1,975,503	2	3	2	2	2	1	1	2	6

The isolates BM1374163 and BM1374165 are synonymous with MVT01 and MVT03, respectively, as described earlier [3].

divergence of this species from other *Bartonella*.

2. Methods

2.1. Profiling of pan-genome

The PanCoreGen tool [11] was applied to determine core and accessory (i.e., mosaic and genome-specific) genes across the completed genomes from four *B. henselae* isolates – Houston-1 from USA [9], and BM1374163, MVT02 and BM1374165 from France [3]. Threshold values for both nucleotide sequence identity and gene-length coverage were set to 75%, since the isolates are phylogenetically close to each other as depicted from the relatedness of 16S rRNA and multi-locus sequence typing (MLST) loci (see Results and Discussion below for details) [12]. Thus, anything below that cut-off was not considered as an ortholog for a given reference gene sequence in other strains.

2.2. Detection of prophage regions

We used the PHASTER (PHAge Search Tool Enhanced Release) Web Server (http://phaster.ca/) [13] in order to identify prophage regions present in each genome (last access date: December 06, 2017). The GenBank-formatted file was uploaded to the PHASTER web-server and the output results were downloaded for each genome. All prophage regions designated as “intact”, “incomplete” or “questionable” by PHASTER were considered in the analysis.

2.3. Duplicated gene identification

All the duplicated genes (paralogs) within a genome were identified by an intra-genome gene homologous identification method. We performed BLASTn for every gene within a genome with all other genes of that genome by an in-house developed Perl program with a sequence identity and length coverage threshold of 75%. Given the BLASTn result, the in-house Perl script parses specific BLAST hits, based on user-defined nucleotide sequence identity and length coverage thresholds, and constructs a tab-delimited matrix as output.

2.4. Detection of genomic islands

Earlier work [9] detected five genomic islands in *B. henselae*. Using that report, along with information about tRNA genes (retrieved from “*_feature_table.txt” files of NCBI GenBank for the respective genomes) and considering the conserved regions present beside tRNA genes, we finally identified the GI regions for each of the genomes. Then, from the genomic position information of ST1/9-specific genes, ST6-specific genes and Houston-1-specific genes, we determined the gene-sets, which were present within the GI-region identified.

2.5. Identification of orthologs

Using the entire set of genes specific to ST1/9, ST6 and Houston-1, we carried out a BLASTn search against ‘nr database’ (using 75% as both identity and alignment length cut-offs, while using defaults for the remaining parameters in the BLASTn run), to identify orthologs. If a hit satisfied our required BLASTn criteria from a bacterial genome other than *B. henselae*, then that gene was referred to as an ‘inserted gene’. However, a gene with sequence identity ≥75%, but with an alignment length value < 50%, then the corresponding region was inferred to be a ‘truncated gene’.

All these methodological approaches are outlined in Fig. S1 as a schematic pipeline, sequentially as used.

3. Results and discussion

3.1. Completed genomes represent multiple clones of B. henselae

We determined the genotypes of all available completed genomes from *B. henselae*, including strains BM1374163, BM1374165, Houston-1 and MVT02, based on 16S rRNA and internal fragments of seven housekeeping genes that constitute the MLST loci for *B. henselae* (i.e., *batR*, *ftsZ*, *gltA*, *groEL*, *nlpD*, *ribC* and *rpoB*; https://pubmlst.org/bhenselae/). The Houston-1, BM1374163 and MVT02 strains were found to have identical MLST profiles but different types of 16S rRNA. Specifically, the American strain Houston-1 possessed 16S rRNA genotype 1, whereas the BM1374163 and MVT02 strains were found to contain 16S rRNA genotype 2. Based on these results, the strains were designated as ST1 (with genotype 1) and ST9 (with genotype 2), even though all three isolates had identical housekeeping genes as part of their MLST loci. In contrast, a second French strain, BM1374165, represented ST6 (Table 1). Previous work on *B. henselae* isolates from Australia, Europe and North America showed that ST1 and ST6 were abundant STs on all three continents [14]. Although ST9 was found at much lower frequencies in isolates from Europe and USA, and none were from Australia [14], this ST is basically identical to ST1 in terms of the MLST housekeeping loci background. These two STs, as stated above, were given different names because of their distinct 16S rRNA types. Therefore, based on MLST background, our four analyzed genomes represent two sequence types – ST1 (or ST9), and ST6. From this point forward, we group the lone isolate of ST1 and the two isolates of ST9 as representatives of ST1/9 (Table 1).

Differences between ST1/9 and ST6 were observed in the sequences of *batR* (4 changes), *ftsZ* (3 changes), *gltA* (1 change), *groEL* (1 change) and *rpoB* (1 change). Polymorphisms in *batR* and *ftsZ* were distributed across the entire length of each locus, suggesting that the changes in each gene accumulated through independent mutations, not via a single recombination event. For the concatenated sequence of these seven protein-coding genes, the pairwise nucleotide diversity (π) and the rates of synonymous (dS) and non-synonymous (dN) changes were only 0.3%, 1.2% and 0.1% respectively, as calculated by using mutation-fraction method [15]. Our recent analysis with *B. bacilliformis* showed a within-subspecies MLST diversity of 0.5% [12]. Although an even lower diversity was detected in *B. henselae*, we were dealing with only two STs (ST1/9 and ST6) here. Interestingly, the diversity values between any two STs in *B. bacilliformis* ranged from 0.03% to 1.02% [12]. This similar low diversity values suggest that clonal differentiation in these two human pathogens might be following a similar pattern.

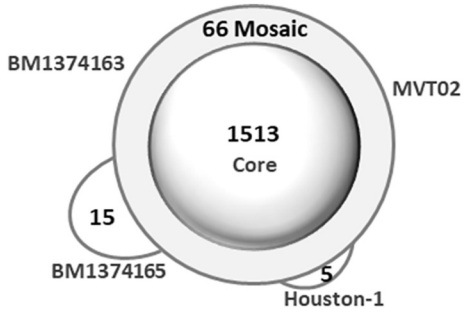

Fig. 1. Pan-genomic profile of analyzed *B. henselae* isolates.

3.2. Limited horizontal gene transfer precedes species divergence

We next performed a pan-genomic profiling based on the four *B. henselae* genomes. Each genome was considered as a reference to sequentially blast against the others for orthologs using 75% threshold values for both nucleotide sequence identity and gene-length coverage. We found 1513 (95%) core genes out of a total of 1599 genes in the pan-genome (Fig. 1 and Table S1). The average pairwise nucleotide diversity (π), dS and dN of core genes were 0.9%, 2.1% and 0.7%, respectively, where 1258 genes (83%) were found to be polymorphic.

Of the four strains, BM1374163, MVT02 (both ST1/9 with 16S rRNA genotype 2) and BM1374165 (ST6 with 16S rRNA genotype 2) were isolated from the blood of patients with tick bites in France. In contrast, Houston-1 (ST1/9 with 16S rRNA genotype 1) was isolated from the blood of an HIV-positive male in Houston, Texas, USA. Despite the isolation of ST1/9 strains from different continents (although being represented by different 16S rRNA genotypes in two continents), the within-clonal diversity of core genes of three genomes was significantly lower ($\pi = 0.3\%$, z-test $P < .0001$) than the overall values. This indicates the robustness of a stable clonal structure of ST1/9 in *B. henselae* across continents. An alternative possibility could be a relatively recent emergence and/or spread of this clonal group.

Among the 5% accessory genome fraction, 66 genes were mosaic, i.e., present in multiple but not all genomes (Fig. 1 and Table S1). Interestingly, all of these genes were present in the three ST1/9 genomes but were missing from the ST6 (BM1374165) genome. There were only five genes specific to the Houston-1 genome, while 15 genes were found only in the BM1374165 (ST6) genome (Fig. 1 and Table S1).

Previous studies showed the presence of one prophage cluster (denoted here as PI_1) and three genomic islands (denoted here as GI_1, GI_2 and GI_3) primarily in ST1 isolates but also in other STs to different degrees, along with a *Bartonella*-specific genomic island (GI_4 in this study) [8,9]. As we searched for these horizontally-transferred regions in the analyzed genomes, all of the strains were found to harbor all five genomic islands (Fig. 2 and Table S2).

Importantly, two-thirds of the apparently ST1/9-specific genes (44 genes) were part of either the prophage cluster or one of the genomic islands (Table S1). We found the closest homologs of these genes in *Bartonella grahamii*, *Bartonella tribocorum* and *Bartonella vinsonii* (not shown). Although we defined the set of 66 genes as mosaic based on 75% cut-off limits for finding orthologs (see Methods), we found that many of these genes were truncated (and mentioned as pseudogenes in the annotated genome file) or deleted in BM1374165 (Table S1). Out of five Houston-1-specific genes, one had a truncated copy in BM1374163, while the remaining genes were part of GI_3, with closest homologs in *B. grahamii* and *B. tribocorum* (Table S1). Of the BM1374165-specific gene set, eight were distributed in PI_1, GI_2 and GI_4 (again showing closest homologs in the same three *Bartonella* species), while truncated copies of the remaining seven genes were found in BM1374163 and Houston-1. However, all eight genes were missing in MVT02 (Table S1). Overall, our analyses could not detect any horizontally-transferred island unique to any genome or ST examined here. Taken together, these results strongly suggest that the relevant horizontal transfer events preceded the divergence of *B. henselae*, and they did not contribute to any intra-species mosaicism. This leads to the hypothesis that the accessory gene fraction of *B. henselae* originated by gene deletion, gene duplication or by a combination of both mechanisms. Moreover, the differential presence/absence of genes in the islands was mostly specific to ST, e.g., not separately in any single strain of ST1/9 (although ST6 in our sample set had a single representative strain), suggesting that such deletion/duplication occurred after clonal divergence at the MLST level, but prior to the diversification of 16S rRNA genotypes (as evidenced for ST1/9 with genotypes 1 and 2).

3.3. Gene content mosaicism is primarily contributed by gene duplication

The genome length of ST6 strain BM1374165 is ~1.98 Mb, compared to the ~1.91–1.93 Mb genomes of the ST1/9 strains (Table 1). Almost 80% of the difference in genome lengths between ST1/9 and ST6 could be attributed to differences in total lengths of five islands in ST1/9 (as averaged from data in Table S2) vs. the ST6 isolate. Although there was not much difference in gene content for the GI_1 and GI_4 islands between ST1/9 and ST6, the PI_1 and GI_3 sequences in the ST6 isolate were much reduced. The PI_1 was found to be about 20 Kb smaller in the ST6 genome versus ST1/9 genomes, while the length difference in the GI_3 region was 14 Kb between BM1374165 (ST6) and BM1374163/MVT02, and as high as 23 Kb between BM1374165 (ST6) and Houston-1 (Table S2). In contrast, the GI_2 island in ST6 was more than six times longer than that of ST1/9 (Table S2).

The complete or partial presence of all the horizontally-transferred islands in the analyzed ST1/9 and ST6 genomes (Fig. 2 and Table S2) rules out the possibility that new genes were acquired via any unique horizontal transfer event that led to mosaicism. Therefore, an alternative explanation is gene deletion, as evidenced by several genes in the

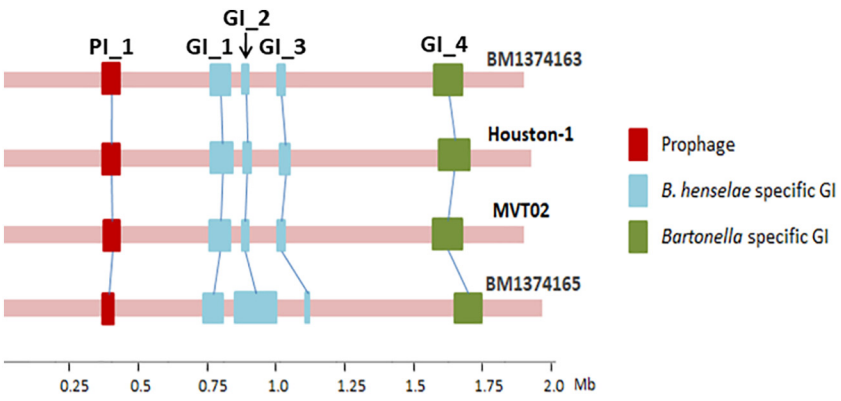

Fig. 2. Map of horizontally-transferred prophage (PI) and genomic (GI) islands detected in the analyzed *B. henselae* genomes.

Table 2
(A) Copies of duplicated genes inside and outside of horizontally-transferred islands in the analyzed *B. henselae* genomes. (B) Copies of duplicated genes exclusively present either in: BM1374163/MVT02 (these two genomes show 100% overlap), Houston-1 or BM1374165.

	Duplicated gene-copies	Within the islands					Outside the islands
		PI_1	GI_1	GI_2	GI_3	GI_4	
(A)							
BM1374165	274	26	43	109	11	15	70
Houston-1	154	26	43	6	17	26	36
BM1374163	141	27	32	6	13	22	41
MVT02	141	27	32	6	13	22	41
(B)							
BM1374165	143	6	12	81	4	0	40
Houston-1	16	1	5	0	3	7	0
BM1374163/ MVT02	15	1	1	0	3	6	4

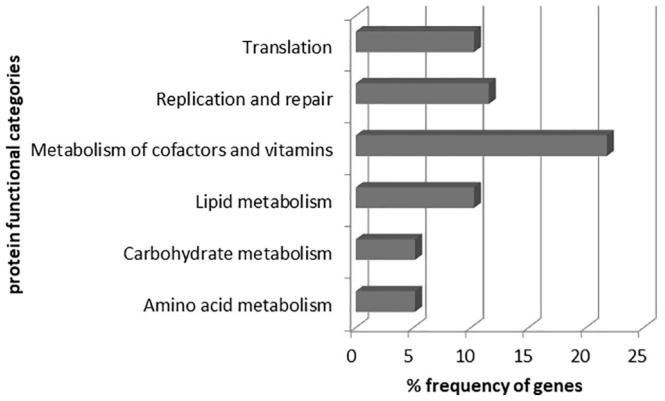

Fig. 3. KEGG categories represented by genes duplicated exclusively in the GI_2 region of BM1374165 (ST6) genome.

mosaic and genome-specific sets being either truncated or deleted from one or more genomes (Table S1). However, this does not explain the larger genome of BM1374165 relative to other strains. Therefore, to assess the contribution of a third possibility, i.e., the duplication of existing genes, we computed the number of duplicated genes in all four analyzed genomes. These results showed that BM1374165 (ST6) had almost twice the copies of duplicated genes relative to those in ST1/9 genomes (Table 2A), and > 70% of these duplications happened inside the horizontally-transferred regions in each of the analyzed strains. Importantly, the number of duplicated gene copies both inside and outside of all the islands except GI_2 in any of the ST1/9 strains (ranging from 135 to 148) was not much different from the number in the ST6 strain BM1374165 (165). In contrast, the increase in frequency of duplications in GI_2 of BM1374165 was > 18 times higher (Table 2A).

These results show that while the length difference in GI_2 could be responsible for the increased length of the BM1374165 genome compared to the other genomes, the major underlying mechanism was gene duplication, primarily in GI_2 and outside of the horizontally-transferred regions (Table 2). There could always be a possibility that the errors or differences in sequencing and/or assembly methods might have led to the increased length and an excess of gene duplications in the BM1374165 genome. However, such a possibility can be ruled out given the fact that this genome (synonymous with MVT03) was sequenced along with BM1374163 (synonymous with MVT01) and MVT02 with the same Nextera XT DNA sample prep kit and 2 × 250 paired-end protocol using Illumina MiSeq pyrosequencer as part of the same study [3] and then annotated by the same NCBI Prokaryotic Genome Annotation Pipeline (release version 2013).

Out of a total of 109 duplicated copies of GI_2 in BM1374165 (versus six in other genomes), 81 duplications represented genes that were duplicated exclusively in the ST6 genome (Table 2B). Functional classification showed that the majority of these genes included clusters of replication and repair, translation and metabolic pathways of co-factors/vitamins, lipids, amino acids and carbohydrates (Fig. 3). While the reason for duplicating these gene families is unclear, the importance of gene duplication in the adaptive evolution of microbial species is well documented, especially when populations must adapt to changing environments [16–18]. Moreover, gene duplications in the context of overall genome reduction have previously been observed in other pathogenic alphaproteobacteria, such as rickettsiae [19]. Under selection pressures, gene duplication events can help bacteria by increasing expression of genes encoding proteins for specific needs such as nutrient transport, stress management under varying environmental conditions, antibiotic resistance, etc. [20–22]. Extra gene copies may also be mutagenized and become nonfunctional or adopt new functions that help bacteria deal with selective pressures [23]. Since *B. henselae* is vectored by arthropods, reservoired by cats and infects the human host incidentally, it is expected that such drastic changes and disparate habitats would provide strong selection pressures and require adaptation by the pathogen [24].

While both horizontal gene transfer and gene deletion essentially alter the number of non-redundant gene-sets in the genome, gene duplication leads to an increase in the number of pre-existing gene-copies without the incorporation of new genes. This study reveals gene duplication, in association with some deletion events, as the principal contributor to a considerable difference in the genome-size of *B. henselae* strains from two different clonal groups. It would be important to study whether gene duplications arose as adaptive responses to selection pressures. It is noteworthy that although the results are intriguing, we realize that our conclusions are based on the four available genomes of *B. henselae*. As such, future comparative genomic studies involving supplementary *B. henselae* genomes, as they become available, are warranted. By the same token, our findings highlight the need for functional assessment of gene duplication events in bacterial genome evolution, especially in relation to host-(patho)adaptation and intracellularity as exemplified by *B. henselae*.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2019.03.009.

Acknowledgments

Financial support for RB from the DBT-RA Program in Biotechnology and Life Sciences is gratefully acknowledged. SP was supported by Ramanujan Fellowship of Science and Engineering Research Board (SERB), Government of India.

References

[1] A. Harms, C. Dehio, Intruders below the radar: molecular pathogenesis of *Bartonella* spp, Clin. Microbiol. Rev. 25 (2012) 42–78 (pmid: 22232371).
[2] U. Okaro, A. Addisu, B. Casanas, B. Anderson, *Bartonella* species, an emerging cause of blood-culture-negative endocarditis, Clin. Microbiol. Rev. 30 (2017) 709–746 pmid: 28490579.
[3] M. Vayssier-Taussat, S. Moutailler, F. Féménia, P. Raymond, O. Croce, et al., Identification of novel zoonotic activity of *Bartonella* spp., France, Emerg. Infect. Dis. 22 (2016) 457–462 (pmid: 26885624).
[4] Z. Zeaiter, P.E. Fournier, D. Raoult, Genomic variation of *Bartonella henselae* strains detected in lymph nodes of patients with cat scratch disease, J. Clin. Microbiol. 40 (2002) 1023–1030 (pmid: 11880432).
[5] C. Toft, S.G.E. Andersson, Evolutionary microbial genomics: insights into bacterial host adaptation, Nat. Rev. Genet. 11 (2010) 465–475 (pmid: 20517341).
[6] S.G.E. Andersson, C. Dehio, *Rickettsia prowazekii* and *Bartonella henselae*: differences in the intracellular life styles revisited, Int. J. Med. Microbiol. 290 (2000) 135–141 (pmid: 11045918).
[7] L. Guy, B. Nystedt, Y. Sun, K. Näslund, E.C. Berglund, et al., A genome-wide study of recombination rate variation in *Bartonella henselae*, BMC Evol. Biol. 12 (2012) 65 (pmid: 22577862).
[8] H. Lindroos, O. Vinnere, A. Mira, D. Repsilber, K. Näslund, et al., Genome

rearrangements, deletions, and amplifications in the natural population of *Bartonella henselae*, J. Bacteriol. 188 (2006) 7426–7439 (pmid: 16936024).
[9] C.M. Alsmark, A.C. Frank, E.O. Karlberg, B.A. Legault, D.H. Ardell, et al., The louse-borne human pathogen *Bartonella quintana* is a genomic derivative of the zoonotic agent *Bartonella henselae*, Proc. Natl. Acad. Sci. 101 (2004) 9716–9721 (pmid: 15210978).
[10] H.L. Lindroos, A. Mira, D. Repsilber, O. Vinnere, K. Näslund, et al., Characterization of the genome composition of *Bartonella koehlerae* by microarray comparative genomic hybridization profiling, J. Bacteriol. 187 (2005) 6155–6165 (pmid: 16109957).
[11] S. Paul, A. Bhardwaj, S.K. Bag, E.V. Sokurenko, S. Chattopadhyay, PanCoreGen - profiling, detecting, annotating protein-coding genes in microbial genomes, Genomics 106 (2015) 367–372 (pmid: 26456591).
[12] S. Paul, M.F. Minnick, S. Chattopadhyay, Mutation-driven divergence and convergence indicate adaptive evolution of the intracellular human-restricted pathogen, *Bartonella bacilliformis*, PLoS Negl. Trop. Dis. 10 (2016) e0004712(pmid: 27167125).
[13] D. Arndt, J. Grant, A. Marcu, T. Sajed, A. Pon, et al., PHASTER: a better, faster version of the PHAST phage search tool, Nucleic Acids Res. 44 (W1) (2016) W16–W21 (pmid: 27141966).
[14] M. Arvand, E.J. Feil, M. Giladi, H.J. Boulouis, J. Viezens, Multi-locus sequence typing of *Bartonella henselae* isolates from three continents reveals hypervirulent and feline-associated clones, PLoS One 2 (2007) e1346(pmid: 18094753).
[15] M. Nei, T. Gojobori, Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions, Mol. Biol. Evol. 3 (1986) 418–426 (pmid: 3444411).
[16] M.H. Serres, A.R.W. Kerr, T.J. McCormick, M. Riley, Evolution by leaps: gene duplication in bacteria, Biol. Direct 4 (2009) 46 (pmid: 19930658).
[17] A. Levasseur, P. Pontarotti, The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics, Biol. Direct 6 (2011) 11 (pmid: 21333002).
[18] N. Chia, N. Guttenberg, Dynamics of gene duplication and transposons in microbial genomes following a sudden environmental change, Mob. Genet. Elements 1 (2011) 221–224 (pmid: 22479691).
[19] A. Diop, D. Raoult, P.E. Fournier, Paradoxical evolution of rickettsial genomes, Ticks Tick Borne Dis. 10 (2019) 462–469.
[20] L. Sandegren, D.I. Andersson, Bacterial gene amplification: implications for the evolution of antibiotic resistance, Nat. Rev. Microbiol. 7 (2009) 578–588 (pmid: 19609259).
[21] F.A. Kondrashov, Gene duplication as a mechanism of genomic adaptation to a changing environment, Proc. Biol. Sci. 279 (2012) 5048–5057 (pmid: 22977152).
[22] V. Katju, U. Bergthorsson, Copy-number changes in evolution: rates, fitness effects and adaptive significance, Front. Genet. 4 (2013) 273 (pmid: 24368910).
[23] J.B. Walsh, How often do duplicated genes evolve new functions? Genetics 139 (1995) 421–428 (pmid: 7705642).
[24] S.D. Hooper, O.G. Berg, On the nature of gene innovation: duplication patterns in microbial genomes, Mol. Biol. Evol. 20 (2003) 945–954 (pmid: 12716994).