Brucella and Bartonella nucleotide pangenome comparative analysis

A.A.Simonova, A.V. Alexeevski

A.N.Belozersky Institute Of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow 119991, aba@belozersky.msu.ru

Introduction. Brucella and Bartonella species are closely related facultative pathogens belonging to Rhizobiales. According to their lifestyle they are supposed to combine genome reduction and genome divergence. Both genera are mostly host-specific zoonotic pathogens. Brucella genomes contain 2 chromosomes, and Bartonella genomes consist of a single circular chromosome, which appear to be reduced version of chromosome I of *Brucella* with chromosome II segment. [1],[2] The main objectives of this study were to compare *Brucella* and *Bartonella* nucleotide pangenomes and to look for relatively recent horizontal gene transfer events.

Matherial and methods:

Matherial. Genomes of 55 strains of 9 Brucella species and 15 strains of 9 Bartonella species assembled completely were used.

Pangenome building. For pangenome building and analyzing NPG-explorer was used. NPG-explorer (http://mouse.belozersky.msu.ru/tools/npge.html) is a package developed by B.E. Nagaev and A.V. Alexeevski that makes a nucleotide pangenome as a set of aligned blocks, each block composed of orthologues fragments; presents detail reports on evolutionary events and genes, provides detailed information on statistics and phylogeny.

HGT search. We assume that blocks recognized as unique for pangenome or common only for a clade or a set of clades might be laterally transferred from external genera. We performed BLAST search of these sequences in NR database.

Brucella and Bartonella pangenome comparison.

Table 1: Nucleotide pangenome statistics for *Brucella* and *Bartonella*.

	Brucella	*Bartonella*	*Bartonella henselae*
Genomes amount	55	15	4
Sequences amount	110[1]	18[2]	4
Stable blocks (represented once in each genome)	425	650	203
Average Identity	97%	71%	98%
Percentage of input length	61.2%	6.0%	83.4%
Partial blocks (represented once in subset of genomes)	498	25099	102
Average Identity	97%	80%	98%
percentage of input length:	13%	65%	2.4%
Unique (blocks of 1 fragment)	107	8151	67
Percentage of input length	0.1%	6.4%	0.3%
Global blocks (synteny regions)	76	68	13
Percentage of input length	66.8%	53.9%	97.7%

[1] Each genome contains 2 chromosomes

[2] 3 genomes contain plasmids

We consider stable blocks as nucleotide core of a pangenome. Table 1 demonstrates comparatively low identity and percentage of input length of stable blocks in *Bartonella* pangenome. This information lets us assume that in terms of genomics *Bartonella* seems to be different taxonomic rank than *Brucella.* For our work average identity of 71% and percentage of input length of 6% for *Bartonella* stable blocks means that orthologous nucleotide fragments are not revealed by identity threshold 70% used for this pangenome construction. Inadequate size of nucleotide core of *Bartonella* is also supported by huge ammount of partial blocks in *Bartonella*. So comparative analysis of *Brucella* and *Brucella* nucleotide pangenomes becomes ineffective. Therefore for correct comparison a nucleotide pangenome for *Bartonella henselae* was constructed.

Horizontal gene transfer (HGT) to 55 *Brucella* genomes. After analyzing *Brucella* pangenome 9 unique and 22 partial blocks were detected as potential laterally transferred loci. BLAST search confirmed transfers to 8 unique blocks (Table 2) and 14 partial blocks. Transfers to unique blocks are from other genera of Rhizobiales and Firmicutes and transfers to partial blocks are mostly from Rhizobiales and Rhodobacterales.

Table 2: Lateral transfers to unique blocks of *Brucella*.

Donor organism	Brucella strain	Length	Identity	Taxonomy
Microvirga ossetica strain V5/3M, complete genome	*B. pinnipedialis* PI-B2/94	15917	72%	Rhizobiales
Shinella sp. HZN7 plasmid pShin-01, complete sequence	*B. pinnipedialis* PI-B2/94	4485	84%	Rhizobiales
Shinella sp. HZN7 plasmid pShin-01, complete sequence	*B. pinnipedialis* PI-B2/94	2631	84%	Rhizobiales
Shinella sp. HZN7 plasmid pShin-01, complete sequence	*B. pinnipedialis* PI-B2/94	1002	95%	Rhizobiales
Pelosinus fermentans JBW45, complete genome	*B. abortus* ZW0533	1884	94%	Selenomonadales
Bacillus thuringiensis YBT-1518, complete genome	*B. abortus* ZW053	4141	99%	Bacillales
Bacillus thuringiensis YBT-1518, complete genome	*B. abortus* ZW053	4999	99%	Bacillales
Clostridium difficile complete genome, strain M120	*B. abortus* ZW053	2412	79%	Clostridiales

Brucella pinnipedialis B2/94, AC embl:CP002078.1 has several consecutive loci, separated by repeat and minor blocks, supposed to be laterally transferred from Rhizobiales.

Brucella abortus strain ZW053, AC embl:CP009098.1 has two sites supposed to be laterally transferred from *Bacillus thuringiensis* YBT-1518, Firmicutes (Table 3).

Table 3: Selected best BLASTN hits for unique block 4141 bp from *Brucella abortus* strain ZW053, AC embl:CP009098.1

Description	Max score	Query cover	E value	Ident	Accession
Bacillus thuringiensis YBT-1518, complete genome	2830	100%	0.0	99%	CP005935.1
Bacillus cereus strain CMCC P0011, complete genome	2800	100%	0.0	99%	CP011153.1
Bacillus thuringiensis strain HS18-1, complete genome	2691	100%	0.0	98%	CP012099.1
Bacillus cereus strain NJ-W, complete genome	2646	100%	0.0	97%	CP012483.1

HGT to 4 *Bartonella* genomes. For *Bartonella* transfers analysis 4 genomes of *B. henselae* were chosen. 2 blocks of 473 bp and 1269 bp turned out to be unique, without any significant similarity discovered. No strong arguments for recent HGT events from external genera to *B. henselae* were found among other unique blocks.

Conclusion. Despite the fact that *Brucella* and *Bartonella* are closely related genera with similar lifestyle their nucleotide pangenomes show that *Bartonella* genus do not support level of identity comparable with *Brucella.* This could probably be due to significant discordance between microbiological approach to taxonomy and comparative genomic data in this case and made us choose different taxonomic ranks to explore HGT events. *Brucella* pangenome analysis also showed that lately sequenced *Brucella vulpis* stands alone from the rest of *Brucella* species (data not shown here). HGT search revealed several patterns in recent transfer events in *Brucella* and two unique blocks in *Bartonella henselae* with no homologues. Genes in the transferred loci of *Brucella* and unique loci of *Bartonella hensalae* have been described (data not shown here).

1. A.R. Wattam et al. (2014) Comparative Phylogenomics and Evolution of the Brucellae Reveal a Path to Virulence, *Journal of Bacteriology,* **196(5):**920–30.
2. E.C. Berglund (2009) Genome Evolution and Host Adaptation in *Bartonella, Digital Comprehensive Summaries of Uppsala Dissertations From the Faculty of Science and Technology* **675**.