ORIGINAL ARTICLE

New approaches in the systematics of rickettsiae

S. N. Shpynov[1], P.-E. Fournier[2], N. N. Pozdnichenko[3], A. S. Gumenuk[3] and A. A. Skiba[3]

1) *N.F. Gamaleya NRCEM, Moscow, Russia,* 2) *UMR VITROME, Aix-Marseille Université, IRD, Service de Santé des Armées, Assistance Publique-Hôpitaux de Marseille, Institut Hospitalo-Universitaire Méditerranée-Infection, Marseille, France and* 3) *OmSTU, Omsk, Russia*

Abstract

The development of a formal order analysis (FOA) allowed constructing a classification of 49 genomes of *Rickettsiaceae* family representatives. Recently FOA has been extended with new tools—'Map of genes,' 'Matrix of similarity' and 'Locality-sensitive hashing'—for a more in-depth study of the structure of rickettsial genomes. The new classification confirmed and supplemented the previously constructed one by determining the position of *Rickettsia africae* str. ESF-5, *R. heilongjiangensis* 054, *R. monacensis* str. IrR/Munich, *R. montanensis* str. OSU 85-930, *R. raoultii* str. Khabarovsk, *R. rhipicephali* str. 3-7-female6-CWPP and *Rickettsiales* bacterium str. Ac37b. The 'Map of genes' demonstrated the complete genomes and their components in a graphical form. The 'Matrix of similarity' was applied for an in-depth classification to a subtaxonomic category of the strain within the species *R. rickettsii* (11 strains) and *R. prowazekii* (ten strains). The 'Matrix of similarity' determines the degree of homology of complete genomes by pairwise comparison of their components and identification of those being identical and similar in the arrangement of nucleotides. A new genomosystematics approach is proposed for the study of complete genomes and their components through the development and application of FOA tools. Its applications include the development of principles for the classification of microorganisms, based on the analysis of complete genomes and their annotations. This approach may help in the taxonomic classification and characterization of some *Candidatus Rickettsia* spp. that are found in large numbers in arthropods worldwide.

Keywords: Arthropods, ecology, epidemiology, formal order analysis, genome, genomosystematics, rickettsiae, rickettsioses, systematics, virulence

Original Submission: 21 November 2017; **Revised Submission:** 27 January 2018; **Accepted:** 7 February 2018

Corresponding author. S.N. Shpynov, 18, N.F. Gamaleya str., Moscow, 123098 Russia.
E-mail: stanislav63@yahoo.com

Introduction

Rickettsia species are strictly intracellular vector-borne bacteria from the order *Rickettsiales* that cause mild to severe diseases in humans and other animals [1]. Currently 30 *Rickettsia* species have standing in nomenclature (http://www.bacterio.net/-allnamesmr.html). About 20 of them are pathogenic, including the causative agents of the deadly diseases epidemic typhus and Rocky Mountain spotted fever (RMSF) [2]. About 60 rickettsiae that were isolated or detected in ticks are currently considered as nonpathogenic, not validated, incompletely described and/or uncultivated species. Among uncultivated rickettsiae, 15 are potential new species and may be classified as *Candidatus* spp. (http://www.bacterio.net/-candidatus.html).

However, the systematics and nomenclature of *Rickettsia* species are based on a limited number of available phenotypic characteristics as a result of their obligate intracellular location. Members of the genus *Rickettsia* were initially classified on the basis of their morphologic, antigenic and metabolic characteristics into the following groups: (a) the spotted fever group (SFG), which includes species transmitted by hard ticks such as *Rickettsia conorii,* the causative agent of Mediterranean spotted fever [3], and *R. rickettsii,* the agent of RMSF [4]; (b) the typhus group (TG), which includes *R. typhi,* the flea-transmitted causative agent of murine typhus and *R. prowazekii,* the louse-transmitted agent of epidemic

https://doi.org/10.1016/j.nmni.2018.02.012

typhus [5,6]; and (c) the group containing *R. tsutsugamushi,* the aetiologic agent of scrub typhus [7]. Then the application of molecular and phylogenetic methods enabled defining three groups within the genus *Rickettsia:* the TG, the SFG that includes a large collection of mostly tick-borne rickettsiae and an ancestral group (AG), which includes *R. bellii* and *R. canadensis* [8]. *Rickettsia tsutsugamushi* was found to exhibit unique phenotypic and molecular characteristics and was reclassified as *Orientia tsutsugamushi* [9]. Most recently, whole-genome sequence analysis suggested the existence of another group within the *Rickettsia* genus, termed the transitional group, consisting of *R. felis* and *R. akari* [10], but this group is not widely accepted [1].

Most bioinformatic genome analysis programs such as BLAST [11], MEGA [12] and many others, are all based on mathematical and statistical methods that compare only homologous sequence fragments of genomes or their concatenation. In contrast, the formal order analysis (FOA) method transforms the nucleotide order in the sequence into a numerical sequence (value) of fixed length [13]. FOA takes into account the original arrangement of nucleotides in each genome. Previous FOA results have corroborated the separation of species in three groups within the genus *Rickettsia,* including TG, SFG and AG, and *Orientia* as a separate genus. *Rickettsia felis* and *R. akari* were not in the same group according to FOA. Therefore, this approach did not confirm the existence of a so-called transitional group [14].

In recent years, the application of new strategies of culturomics and taxonogenomics has made it possible to isolate and identify difficult-to-cultivate bacterial species, including rickettsiae, and to optimize the amount of information for their classification [15,16]. The complete genome of *Candidatus* Midichloria mitochondrii was sequenced in the female *Ixodes ricinus* tick [17]. The analysis of intergenic regions encoding noncoding small RNAs has demonstrated their role in the virulence of *R. prowazekii* [18]. Thus, in the era of high-quality genomes sequences, it is necessary to develop new approaches for obtaining the maximum amount of information from genome structure and its characteristics.

We sought to develop an approach based of FOA for improving the classification of rickettsiae, with an in-depth study of genomic characteristics for differentiating rickettsial strains and an estimation of their phenotypic characteristics, such as virulence.

Materials and methods

Genome sequences of *Rickettsia* and *Orientia* species

Forty-nine complete genome sequences of *Rickettsia* (n = 47) and *Orientia* (n = 2) spp. (Table 1) were downloaded from GenBank (https://www.ncbi.nlm.nih.gov/genome).

Formal order analysis (FOA) tools

FOA was used to analyse rickettsial genomes as previously described [13,14,19]. FOA uses a high-precision and unambiguous numerical representation of the original arrangement of nucleotides in the sequence. In order to do this, numerical characteristics (average remoteness, depth, etc.) based on intersymbol intervals (internucleotide distance [20]) were developed.

Recently FOA has been improved with the new tools 'Map of genes' (MG) [21], 'Matrix of similarity' (MS) [22] and 'Locality-sensitive hashing' (LSH) (Pozdnichenko NN et al., paper presented at 11th International Conference of Computer-Aided Technologies in Applied Mathematics) for a more in-depth study of rickettsial genome structure.

Pairs of numeric values of order characteristics from studied genomes and their components $\{<g_i, G_i>\}$ are mapped into pillars of dots on the MG. Components representing individual genomes are placed vertically, and some horizontal lines are formed with similar components in different genomes. The MG tool kit enables interactive obtaining of a detailed description of any component of the genome. Automated identification of similar components is also possible.

The MS represents the similarity values for each pair of analysed genomes. Genome similarity is determined by comparing the order characteristic values from their components. The MS tool kit enables interactive obtaining of a list of only similar components of any pair of genomes, and when necessary, a sliding window characteristics for those components can be obtained. The latter enables visualization of element-by-element similarity of genomic components.

The comparison of coding and noncoding sequences in genomes from different strains of *Rickettsia* spp. is possible using the MG tool, which can also be useful in the analysis of genomes from different microorganisms in order to find homologous genes and orthologs. The MG tool identifies interstrain genomic differences within *Rickettsia* species. The MG tool also provides a complete representation of genomes and their components in a graphical form. Genomes are placed (classified) according to the index g on the x-axis, and their components (coding and noncoding sequences) are sited according to the depth index (G) on the y-axis. Sequences that are completely similar by a characteristic are 100% homologous. Strains of the same species show a high degree of homology of their components. It is necessary to introduce a criterion (characteristic) for identification of components with a degree of homology <100% in order to compare genomes of different species. Using this approach, it is possible to conduct a selective analysis of all genomic components, individually or by grouping them by feature (coding DNA sequence, rRNA, tRNA, noncoding RNA, pseudogenes, repeat regions, etc.).

TABLE 1. Genome features of sequenced and average remoteness (*g*) of *Rickettsia* spp. and *Orientia tsutsugamushi*.

No.	No.*	Species and strains of *Rickettsia* и *Orientia*	Access Number in GeneBank	Genome size (bp)	G+C%	*g*
1	1.	*R. prowazekii* str. Katsinyian	NC_017050	1 111 454	29	1.41823242027653
2		*R. prowazekii* str. BuV67-CWPP	NC_017056	1 111 445	29	1.41824128821358
3		*R. prowazekii* str. Madrid E	NC_000963	1 111 523	29	1.41824702490549
4		*R. prowazekii* Rp22	NC_017560	1 111 612	29	1.41826432159974
5		*R. prowazekii* str. GvV257	NC_017048	1 111 969	29	1.41831400875350
6		*R. prowazekii* str. RpGvF24	NC_017057	1 112 101	29	1.41832503153582
7		*R. prowazekii* str. Chernikova	NC_017049	1 109 804	29	1.41838368933900
8		*R. prowazekii* str. Breinl	NC_020993	1 109 301	29	1.41849488956929
9		*R. prowazekii* str Naples	NZ_CP014865	1 111 769	29	1.41827647360422
10		*R. prowazekii* str NMRC Madrid E	NC_020992	1 111 520	29	1.41848807421873
11	2.	*R. typhi* str. B9991CWPP	NC_017062	1 112 957	28.9	1.41989725883639
12		*R. typhi* str. TH1527	NC_017066	1 112 372	28.9	1.41989961796300
13		*R. typhi* str. Wilmington	NC_006142	1 111 496	28.9	1.41990899257420
14	3.	*Rickettsiales bacterium* Ac37b	NZ_CP009217	1 851 238	30.8	1.42369690175481
15	4.	*R. bellii* OSU 85-389	NC_009883	1 528 980	31.6	1.42461059098798
16		*R. bellii* RML369-C	NC_007940	1 522 076	31.6	1.42478460772370
17	5.	*R. canadensis* str. CA410	NC_016929	1 150 228	31.0	1.42505825951416
18		*R. canadensis* str. McKiel	NC_009879	1 159 772	31.0	1.42576172562109
19	6.	*R. monacensis* str. IrR/Munich	NZ_LN794217	1 353 450	32.4	1.42639157617608
20	7.	*R. felis* URRWXCal2	NC_007109	1 485 148	32.6	1.42911847695814
21	8.	*R. heilongjiangensis* 054	NC_015866	1 278 468	32.3	1.4307561483309
22	9.	*R. rhipicephali* str. 3-7-female6-CWPP	NC_017042	1 290 368	32.4	1.43088037051608
23	10.	*R. japonica* YH	NC_016050	1 283 087	32.4	1.43117915307354
24	11.	*R. australis* str. Cutlack	NC_017058	1 296 670	32.3	1.43123685110919
25	12.	*R. montanensis* str. OSU 85-930	NC_017043	1 279 798	32.6	1.43172145972565
26	13.	*R. africae* ESF-5	NC_012633	1 278 530	32.4	1.43188388215182
27	14.	*R. slovaca* str. D-CWPP	NC_017065	1 275 720	32.5	1.43199033266176
28		*R. slovaca* 13-B	NC_016639	1 275 089	32.5	1.43200176043628
29	15.	*R. parkeri* str. Portsmouth	NC_017044	1 300 386	32.4	1.43209421751330
30	16.	*R. conorii* str. Malish 7	NC_003103	1 268 755	32.4	1.43213975511604
31	17.	*R. raoultii* str. Khabarovsk	NZ_CP010969	1 344 517	32.5	1.43249957974534
32	18.	*R. rickettsii* str. Arizona	NC_016909	1 267 197	32.4	1.43264659062342
33		*R. rickettsii* str. Iowa	NC_010263	1 268 188	32.4	1.43268083393081
34		*R. rickettsii* str. Brazil	NC_016913	1 255 681	32.5	1.43272749268022
35		*R. rickettsii* str. Hino	NC_016914	1 269 837	32.5	1.43279406482271
36		*R. rickettsii* str. Colombia	NC_016908	1 270 083	32.5	1.43286684945432
37		*R. rickettsii* str. Hlp#2	NC_016915	1 270 751	32.5	1.43286996863738
38		*R. rickettsii* str. "Sheila Smith"	NC_009882	1 257 710	32.5	1.43291630589912
39		*R. rickettsii* str. R	NZ_CP006009	1 257 005	32.5	1.43288001638348
40		*R. rickettsii* str. Iowa	NZ_CP000766	1 268 201	32.4	1.43268083393081
41		*R. rickettsii* str. Iowa Large Clone	NZ_CP018913	1 268 220	32.4	1.4326837678714
42		*R. rickettsii* str. Iowa Small Clone	NZ_CP018914	1 268 242	32.4	1.43268429687864
43	19.	*R. massiliae* MTU5	NC_009900	1 360 898	32.5	1.43314583756584
44	20.	*R. philipii* str. 364D	NC_016930	1 287 740	32.5	1.43325070488125
45	21.	Candidatus *R. amblyommii* str. GAT-30V	NC_017028	1 407 796	32.45	1.43334598270974
46	22.	*R. peacockii* str. Rustic	CP001227	1 288 492	32.6	1.43514665884247
47	23.	*R. akari* str. Hartford	NC_009881	1 231 060	32.3	1.43747339509228
48	24.	*O. tsutsugamushi* str. Boryong	NC_009488	2 127 051	30.5	1.44599460730303
49		*O. tsutsugamushi* str. Ikeda	NC_010793	2 008 987	30.5	1.44642319193296

All genomes were imported from GenBank NCBI (USA): http://www.ncbi.nlm.nih.gov/genome/.
*Number on Fig. 1.

Each component can be identified among all compared organisms by its name in the annotation, which allows checking its presence in each genome. The LSH of nucleotide sequences is provided by differing values of numerical characteristics. All genomes were analysed using FOA [13] software, which is available online (http://foarlab.org/).

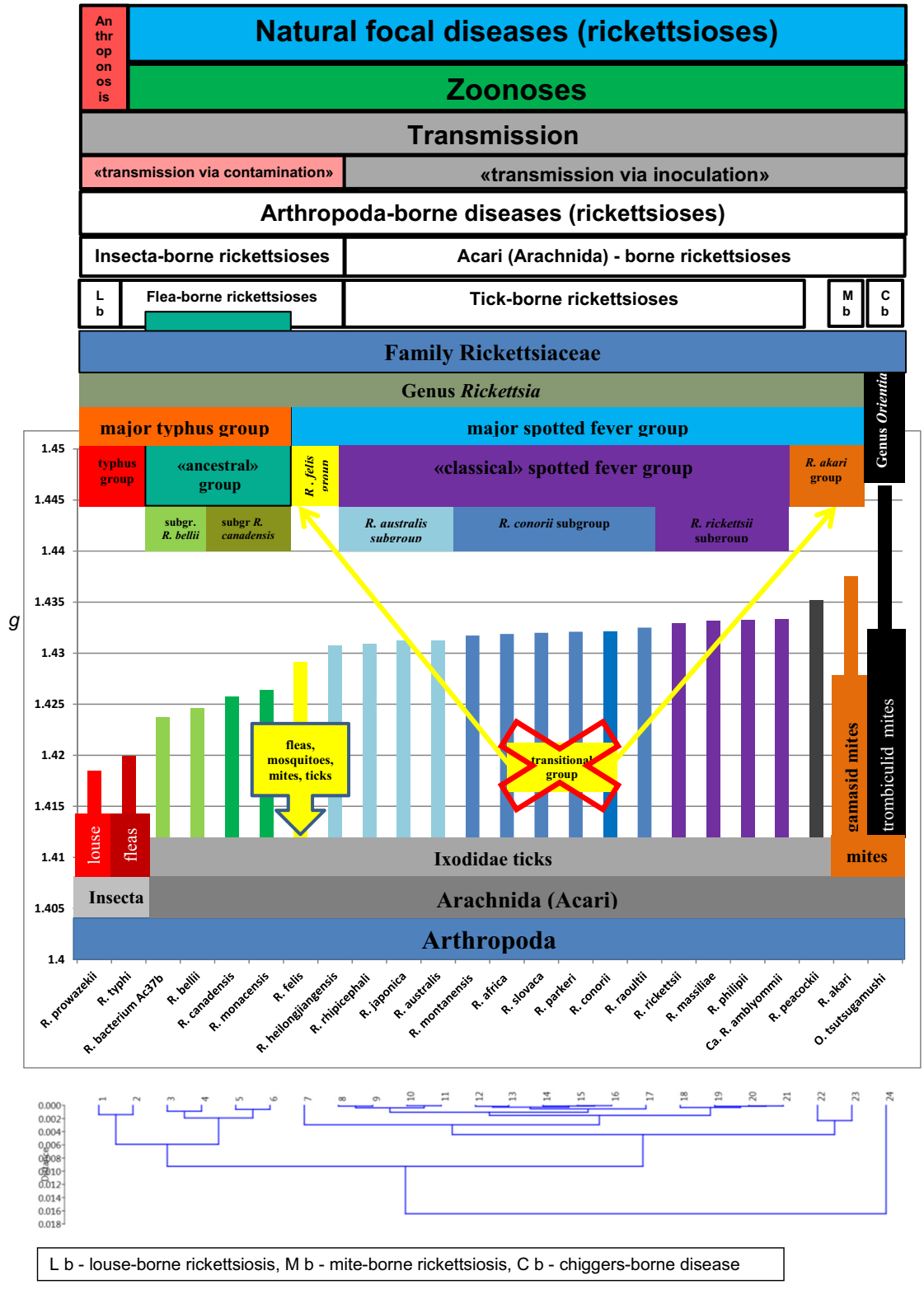
Cluster analysis

Cluster analysis was carried out using PAST software (http://folk.uio.no/ohammer/past/) for verification of the obtained classification scheme for representatives of the *Rickettsiaceae* family and criteria for the formation of taxa within the genus *Rickettsia* (Fig. 1). The UPGMA algorithm (unweighted pair-group average) was used for the analysis of average distances. Clusters were formed on the basis of the average distance between members of all groups.

Results

The classification obtained in this work confirmed and improved the previous one, using only a single characteristic of the order that is the average remoteness (*g*) [14]. The *Rickettsiales* bacterium str. Ac37b and *R. monacensis* were localized on the borders of the AG, beside *R. bellii* str. IrR/Munich OSU 85-389 and *R. canadensis* str. McKiel. *Rickettsia heilongjiangensis* str. 054, *R. rhipicephali* str. 3-7-female6-CWPP, *R. montanensis* str. OSU 85-930, *R. africae* str. ESF-5 and *R. raoultii* str. Khabarovsk were placed in the SFG (Fig. 1).

By using cluster analysis, sets of genomes from members of the family *Rickettsiaceae* were grouped into disjointed subset clusters consisting of genomes that are close by index *g* and represented as a dendrogram (Fig. 1). The family *Rickettsiaceae*



L b - louse-borne rickettsiosis, M b - mite-borne rickettsiosis, C b - chiggers-borne disease

FIG. 1. Systematics of *Rickettsia* spp. and *Orientia tsutsugamushi* using characteristics of average remoteness (*g*) of their genomes, as well as ecological, epidemiologic and nosologic (aetiologic) features (genomosystematics of rickettsiae).

was divided into the genera *Rickettsia* and *Orientia,* with a distance index ranging from 0.016 to 0.018. Two major groups are formed inside the genus *Rickettsia* in the range of 0.008 to 0.01. The major TG (MTG) includes the previous TG and AG. The TG and AG are formed by a value index of 0.006. The AG comprised two subgroups: *R. bellii* (*R. bacterium* str. Ac37b and *R. bellii* str. OSU 85-389) and *R. canadensis* (*R. canadensis* str. McKiel and *R. monacensis* str. IrR/Munich). The major SFG (MSFG) group included the *R. akari* group, the *R. felis* group and the classical SFG. The *R. akari* group (mite-borne, rickettsialpox) exhibits a distance value index of 0.004. The *R. felis* and classical SFG were discriminated by a value index of 0.003. The classical SFG group is also subdivided into three subgroups: *R. rickettsii, R. conorii* and *R. australis.* The *R. rickettsii* subgroup (*R. massiliae* str. MTU5, *R. philipii* str. 364D and *R. amblyommatis* str. GAT-30V) exhibits a distance value index of 0.0018, followed by the subgroups *R. conorii* (*R. raoultii* str. Khabarovsk, *R. montanensis* str. OSU 85-930, *R. africae* str. ESF-5, *R. slovaca* str. 13-B and *R. parkeri* str. Portsmouth) and *R. australis* (*R. heilongjiangensis* str. 054, *R. rhipicephali* str. 3-7-female6-CWPP and *R. japonica* str. YH), with distance value indexes of 0.0015.

The MG tool detected the homology of rRNA genes (5S rRNA, 16S rRNA and 23S rRNA) of rickettsiae. The analysis of rRNA demonstrated that all three *R. typhi* strains of were 100% homologous for these genes, while nine of ten *R. prowazekii* strains had complete 5S rRNA and 16S rRNA homology, with the exception of *R. prowazekii* str. Madrid E. This strain differed in 23S rRNA from all other strains, whereas the remaining nine strains were divided into two groups; Breinl, Chernikova and BuV67-CWPP formed one group, and all remaining strains formed another group. All *R. rickettsii* strains were 100% homologous with *R. philipii, R. conori* and *R. parkeri* for the 5S rRNA gene. All *R. rickettsii* strains were also 100% homologous for 16S rRNA, except *R. rickettsii* str. Hlp#2, which was identical to *R. philipi.*

The analysis of the 23S rRNA gene enabled discriminating *R. rickettsii* strains into three groups: (a) Sheila Smith, R, Brazil and Colombia, (b) Hlp#2 and (c) all remaining strains. Two *R. canadensis* strains exhibited complete homology with each other only for the 5S rRNA gene and were also 100% homologous with *R. bellii, R. rhipicephali, R. montanensis, R. monacensis, R. felis* and *R. japonica.* Two *R. amblyommatis* strains exhibited 100% absolute homology for 5S rRNA.

The MS tool was used for in-depth intraspecific analysis of the most pathogenic *Rickettsia* species, *R. prowazekii* (ten strains) and *R. rickettsii* (11 strains). We studied the correlation between the MS score (genotypic characteristic) obtained by analysing their genomes with FOA and virulence. The MS tool determines the degree of homology of complete genomes by a pairwise comparison of their components and identification of identical and similar components in the arrangement of nucleotides. The results are presented in Tables 2 and 3, respectively. Strains were ranked within these species according to the decrease of MS score with *R. prowazekii* str. Breinl and *R. rickettsii* str. Sheila Smith, which are type strains that have a well-characterized high degree of virulence.

According to their decreasing MS score with str. Breinl, *R. prowazekii* strains were classified as follows: Chernikova (85.7%), Naples (80.37%), Rp22 (80.11%), BuV67-CWPP (74.42%), Katsinyian (73.9%), NMRC Madrid E (64.55%), Madrid E (60.07%), RpGvF24 (52.31%) and GvV257 (50.88%; Table 2).

Similarly, according to their decreasing MS score with str. Sheila Smith, *R. rickettsii* strains were classified as follows; R (98.75%), Brazil (80.79%), Colombia (79.23%), Arizona (75.68%), Hino (75.05%), Morgan (75.01%), Iowa (74.31%), Iowa isolate Large Clone (63.32%), Iowa isolate Small Clone (63.32%) and Hlp#2 (26.22%; Table 3).

Numerical characteristics of order can be used for compact representation and LSH of complete nucleotidic genome sequences. Thus, the genome from *R. prowazekii* str. Madrid E is 1 111 520 bp long and requires 300 pages in FASTA format, it but can be represented in FOA characteristic by a 14-decimal number (e.g., 1.41848807421873).

TABLE 2. Study of homology degree of components for genomes of strains of *Rickettsia prowazekii* using 'Matrix of similarity'

No.	*Rickettsia prowazekii* strain	Katsinyian	BuV67-CWPP	Madrid E	Rp22	Naples-1	GvV257	RpGvF24	Chernikova	NMRC madrid E	Breinl
1	Katsinyian	100.00%	91.44%	79.40%	86.36%	85.67%	60.11%	61.65%	84.71%	77.01%	73.90%
2	BuV67-CWPP	91.44%	100.00%	74.90%	86.78%	85.87%	60.07%	62.38%	85.68%	73.04%	74.42%
3	Madrid E	79.40%	74.90%	100.00%	69.75%	70.52%	47.79%	48.91%	68.41%	66.63%	60.08%
4	Rp22	86.36%	86.78%	69.75%	100.00%	96.09%	58.65%	60.19%	92.44%	69.21%	80.11%
5	Naples-1	85.67%	85.87%	70.52%	96.09%	100.00%	58.12%	59.67%	91.24%	69.30%	80.37%
6	GvV257	60.11%	60.067%	47.79%	58.65%	58.12%	100.00%	86.56%	58.32%	49.59%	50.88%
7	RpGvF24	61.65%	62.38%	48.91%	60.19%	59.67%	86.56%	100.00%	60.30%	50.68%	52.31%
8	Chernikova	84.71%	85.68%	68.41%	92.44%	91.24%	58.32%	60.30%	100.00%	67.57%	85.70%
9	NMRC Madrid E	77.001%	73.04%	66.63%	69.21%	69.30%	49.59%	50.68%	67.58%	100.00%	64.55%
10	Breinl	73.90%	74.42%	60.08%	80.11%	80.37%	50.88%	52.30%	85.70%	64.55%	100.00%

TABLE 3. Study of homology degree of components for genomes of strains of *Rickettsia rickettsii* using 'Matrix of similarity'

No.	*Rickettsia rickettsii* strain	Arizona	Iowa	Iowa isolate, large clone	Iowa isolate, small clone	Brazil	Morgan	Hino	Colombia	Hlp#2	R	Sheila smith
1	Arizona	100.00%	92.50%	78.18%	78.11%	74.91%	93.92%	94.00%	75.85%	26.43%	75.81%	75.68%
2	Iowa	92.49%	100.00%	82.66%	82.60%	74.20%	94.73%	96.63%	74.19%	25.80%	74.44%	74.31%
3	Iowa isolate, large clone	78.18%	82.66%	100.00%	99.87%	63.44%	80.10%	81.24%	63.03%	21.52%	63.64%	63.32%
4	Iowa isolate, small clone	78.11%	82.59%	99.87%	100.00%	63.44%	80.03%	81.24%	63.03%	21.52%	63.64%	63.32%
5	Brazil	74.91%	74.20	63.44%	63.44%	100.00%	73.95%	74.20%	77.94%	26.26%	80.93%	80.80%
6	Morgan	93.92%	94.73%	80.10%	80.03%	73.95%	100.00%	96.14%	74.60%	26.26%	75.28%	75.01%
7	Hino	94.00%	96.63%	81.24%	81.24%	74.20%	96.14%	100.00%	75.07%	26.22%	75.17%	75.05%
8	Colombia	75.85%	74.19%	63.03%	63.03%	77.94%	74.60%	75.07%	100.00%	26.01%	79.29%	79.24%
9	Hlp#2	26.43%	25.80%	21.52%	21.52%	26.26%	26.26%	26.22%	26.01%	100.00%	26.31%	26.22%
10	R	75.81%	74.44%	63.64%	63.64%	80.93%	75.28%	75.17%	79.30%	26.31%	100.00%	98.75%
11	Sheila Smith	75.68%	74.31%	63.32%	63.32%	80.79%	75.01%	75.05%	79.24%	26.22%	98.75%	100.00%

Discussion

The classification obtained using the FOA method showed a significant divergence of the genera *Rickettsia* and *Orientia* within the family *Rickettsiaceae*. Initially two groups were formed in the *Rickettsia* genus: MTG and MSFG. The MTG was divided into the TG and AG. This classification was supported by the detection of antigenic cross-reactions between *R. canadensis* and TG members [23,24]. Furthermore, *R. canadensis* was also suspected to be responsible for cases of acute cerebral vasculitis [24,25]. The nearby position of *R. monacensis* and *R. canadensis* str. IrR/Munich McKiel from the AG group may be due to the detached position from the SFG [26]. The MSFG includes the *R. akari* group, *R. felis* group (*Candidatus R. senegalensis* and *Candidatus R. asemboensis*) and classical SFG members [14,19]. Representatives of the subgroup *R. rickettsii* are predominantly distributed in North and Central America, only in hard ticks (*Dermacentor, Rhipicephalus, Amblyomma* and *Haemaphysalis*); those of the subgroup *R. conorii* are present in Europe, Asia, North Africa, Sub-Saharan Africa, the Pacific Islands and North and Central America (*Amblyomma, Dermacentor* and *Rhipicephalus*); and members of the subgroup *R. australis* are predominantly distributed in the Asian–Australian region (genera *Haemaphysalis, Ixodes, Amblyomma, Dermacentor* and *Rhipicephalus*). On the basis of FOA data and their common vectors (mites), the *R. akari* group was intermediate between the SFG and the *Orientia* genus [19]. The position of *R. peacockii* close to *R. akari* and far from *R. rickettsii* may be explained by significant genomic rearrangements caused by the presence of ISRpe1 transposons [27] and other genomic reorganizations (including deletions) that provoked a loss of virulence [28].

The FOA-based classification demonstrated a broad resemblance to the classification of rickettsiae and rickettsioses based on the complex characteristics of this group of infections published by Zdrodovskii and Golinevich in 1960 [29], as follows.

I. Louse- or flea-borne typhus fever group (aka typhus fever group)
   1. Epidemic, or louse-borne, European or historic typhus fever—agent: *R. prowazekii* (Rocha Lima, 1916) or *R. prowazekii* var. *prowazekii* (Pinkerton, 1936).
   2. Endemic or murine typhus fever (Marcy, 1926)—agent: *R. mooseri* (Monteiro, 1931) or *R. prowazekii* var. *mooseri* (Pinkerton, 1936). Synonym: *R. typhi* (Wolbach and Todd, 1920).

II. Tick-borne SFG
   1. New World subgroup—(1) Rocky Mountain spotted fever (Maxey, 1899); (2) Brazilian or Sao Paolo typhus fever (Monteiro, 1935), agent: *D. rickettsii* (*R. rickettsii*) (Wolbach, 1919).
   2. Old World subgroup—(1) Marseilles or Mediterranean Fever (Fièvre boutonneuse, pimple fever) (Conor and Bruch, 1910), agent: *D. conorii* (*R. conorii*) (Brumpt, 1932); (2) South and East African tick typhus, agent: *D. rickettsii* var. *pijperi* (Alexander and Mason, 1939); (3) North Asian tick rickettsiosis or tick typhus (Velik, Savul'kin, Shmatikov, Krontovskaia et al., 1935–1938), agent: *D. sibiricus* (*R. sibirica*) (Zdrodovskii and Golinevich, 1949); (4) North Australian tick typhus (Andrew, Bonnin, and Williams, 1946), agent: *D. nov. spec.* (Plotz et al., 1946).
   3. Subgroup of gamasid rickettsioses—Varioliform or vesicular rickettsiosis (Huebner, Greenberg et al., 1946–1947; Drobinskii, Zhdanov, Kulagin, 1948–1950), agent: *R. acari* (Huebner et al., 1946).

III. Mite-borne fever group (aka, in Japanese terminology, tsutsugamushi group)

Tsutsugamushi fever or Japanese river fever (Baelz and Kawakami 1879), agent: *R. orientalis* (Nagayo et al., 1930). Synonyms: *R. tsutsugamushi* (Ogata, 1931).

IV. Pneumotropic group of rickettsioses (Q fever group)

V. Paroxysmal group of rickettsioses
VI. Group of rickettsiae and rickettsial diseases of domestic animals

Groups IV, V and VI have lost their significance as a result of the reclassification of aetiologic agents.

Medical taxonomy is critically important to define diseases, based on epidemiologic characteristics, clinical manifestations and vectors involved in the transmission of the aetiologic agents [1]. Arthropods from the class *Insecta* (insects), lice and fleas, are the hosts of *Rickettsia* species from the TG and *R. felis*. Arthropods of the *Arachnida* class are the hosts of *Rickettsia* species from the AG (ixodid), SFG (ixodid), *R. akari* (mites) and *O. tsutsugamushi* (mites). A strong ecological association was established between representatives of each of the rickettsial groups with members of different taxa of the *Arthropoda*. Using the average remoteness characteristic (score), rickettsial groups were classified as follows: TG members (g = 1.418232–1.419908) (Fig. 1) ecologically associated with insects (lice, fleas) were separated from AG (g = 1.424610–1.425761) and SFG (g = 1.431179–1.435146) rickettsiae that are associated with *Arachnida* (ticks) [14]. *Rickettsia akari* (g = 1.437473) was located on the border between the SFG and *O. tsutsugamushi* (g = 1.445994–1.446423), which is also associated with mites (trombiculid, *Arachnida*). It was recommended to isolate *R. akari* as a separate group within the genus *Rickettsia* on the basis of its genomic characteristics (average remoteness) and the taxonomic position of its gamasid mites vectors [19]. *Ctenocephalides felis* and other flea species are the vectors of *R. typhi* and *R. felis*. AG rickettsiae, SFG rickettsiae, *R. akari* and *O. tsutsugamushi* are ecologically associated with representatives of the superorders *Parasitiformes* and *Acariformes*. Ixodid ticks of the family *Ixodidae* (genera *Dermacentor, Rhipicephalus, Amblyomma, Haemaphysalis* and *Ixodes*) of the superfamily *Ixodoidea* of the order *Ixodida* of the superorder *Parasitiformes* are the hosts of AG and SFG rickettsiae, and gamasid mites from the order *Mesostigmata* are the hosts of *R. akari*. Trombiculid mites (gamasid mites) in the order *Trombidiformes* of the superorder *Acariformes* [30] are the vectors and hosts of *O. tsutsugamushi*.

This classification demonstrated a close ecological association between pathogenic rickettsial species and their arthropod hosts. These rickettsiae and *O. tsutsugamushi* are aetiologic agents of distinct nosologic forms, with rickettsioses and scrub typhus exhibiting different ecoepidemiologic and clinical features (host, transmission, seasonal manifestation, etc.).

In contrast with other rickettsioses and scrub typhus, which are zoonoses, epidemic typhus is an anthroponosis. Rickettsioses are divided into two groups: insect-borne and acari-borne rickettsioses (Fig. 1). Insect-borne rickettsioses include louse-borne rickettsiosis (epidemic typhus caused by *R. prowazekii*) and flea-borne rickettsioses (murine typhus and flea-borne spotted fever caused by *R. typhi* and *R. felis*, respectively). Acari-borne rickettsioses consist of tick-borne spotted fevers (among others, RMSF, Mediterranean spotted fever, Siberian tick typhus, Queensland tick typhus caused by *R. rickettsii, R. conorii, R. sibirica* and *R. australis*), mite-borne rickettsiosis (rickettsialpox caused by *R. akari*) and chigger-borne disease (scrub typhus caused by *O. tsutsugamushi*).

Rickettsiae are associated with arthropods, which can transmit them to vertebrates via saliva or faeces. *Rickettsia prowazekii* is transmitted by the human body louse (*Pediculus hominis corporis*), and its main reservoir is humans [31]. Transmission of this bacterium does not occur directly by a bite but by contamination of scratch sites with the faeces or the crushed bodies of infected lice [5,32]. A similar infection mode occurs for *R. typhi* and *R. felis* infections, transmitted by fleas [33,34]. Transmission of *R. typhi* to humans occurs by contamination of the skin or respiratory tract by aerosols of dust containing infective material or via contamination of the conjunctivae of the host with infected flea faeces [32]. Thus, the human infection caused by *R. prowazekii, R. typhi* and *R. felis* is carried out as transmission via contamination.

Ticks are the main vectors and reservoirs of SFG rickettsiae. Rickettsiae infecting the ticks' salivary glands can be transmitted to vertebrate hosts during feeding [32]. *R. akari* is responsible for rickettsialpox, which is an urban disease involving mites of the genus *Allodermanyssus* (*Liponyssoides*), the house mouse *Mus musculus* and, accidentally, humans [32,35]. Humans get rickettsialpox after being bitten by an infected mite. *Orientia tsutsugamushi* is transmitted by bites of feeding larval trombiculid mites (chiggers), which are the reservoir of the agent and the only life stage that feeds on a vertebrate host [36–38]. Therefore, human infection by SFG rickettsiae *R. akari* and *O. tsutsugamushi* is carried out as transmission via inoculation. *Rickettsia felis* has been proposed to be one of the most ancient *Rickettsia* species [39]. It was identified worldwide in more than 20 different haematophagous species of fleas, mosquitoes, ticks and mites [40]. Although *C. felis* fleas were initially considered to be the only vector of *R. felis*, evidence supports the role of other vectors, notably *Anopheles*, in the transmission of the bacterium [41,42]. The acquisition and persistence of *R. felis* in *Anopheles* have been demonstrated, and live bacteria were detected in mosquito faeces and their salivary glands, gut and ovaries [41]. The transmission of *R. felis* to vertebrates by *A. gambiae* and *Liposcelis bostrychophila* is experimentally proven, but transmission to humans is only hypothetical. Vertebrate infection by fleas may occur by blood feeding or contamination of excoriations by faeces [40]. Probably *R. felis* is transmitted to humans both through transmission via contamination and transmission via inoculation modes.

Using the MG tool enabled building a rickettsial classification for each of the genes available online. The presence or absence of a gene is a classification feature for closely related species of rickettsia.

We attempted to rank *R. rickettsii* and *R. prowazekii* strains for virulence by analysing the number of components of the genome having complete (100%) MS homology (Tables 2 and 3). Strains from *R. rickettsii* differ significantly in virulence [43–46]. Virulence varies from the most virulent Sheila Smith strain to the avirulent Iowa. Strains Sheila Smith, Brazil and Morgan were defined as highly cytopathic isolates based on an *in vitro* model, whereas the Colombia and Hlp#2 strains caused lower reactions [43]. The phylogenetic tree constructed using the analysis of multilocus sequences showed that strains Sheila Smith and R are closely related and differ significantly from strains Iowa and Morgan, which are close to each other [46]. Strain Hlp#2 exhibited differences in the *ompA* gene compared to 12 other *R. rickettsii* strains isolated from ticks and patient's blood in the Americas [44]. The virulence loss of strain Iowa may be associated with the disruption of *ompA* and the defect in processing of *ompB,* which are demonstrated to cause protection of guinea pigs in subsequent infection by *R. rickettsii* str. Sheila Smith [47]. The Sheila Smith, R and Brazil strains exhibit a 10 kbp deletion, unlike the others. The significance of this deletion is unclear, since this region is present both in str. Iowa, Morgan, Hino, Hauke, Arizona, Colombia and Hlp#2 [46]. It is believed that the presence or absence of this region does not have a direct effect on virulence, since it is present in the virulent Morgan strain. The indel genotyping system enabled the identification of 25 genotypes of *R. rickettsii* in 4 groups [45]. Strain Hlp#2, which is often considered as nonpathogenic, showed the greatest diversity compared to other strains. Our results support the phylogenetic trees obtained by Clark and colleagues [46], and Genome Tree is publicly available online (https://www.ncbi.nlm.nih.gov/genome/tree/674?). Furthermore, the division of *R. rickettsii* strains into three groups confirms the MG analysis results of the 23S rRNA gene.

Our results support the current understanding of the virulence of rickettsial strains based on phenotypic and genotypic characteristics. *Rickettsia prowazekii* strains exhibit various degrees of virulence [48]. The Breinl strain and a more recent isolate, Rp22, are considered to be highly virulent. The Madrid I strain was isolated in 1941 during the Madrid outbreak of epidemic typhus. After passages in embryonated eggs, strain Madrid I has lost its virulence and has been used under the name of Madrid E as a vaccine in humans since 1944 [49]. A comparative genomic microarray study revealed a highly conserved genome content between str. Breinl and Madrid E (only 3% nucleotide variations) [50]. The draft genomes from the flying squirrel strain GvF12 was found to differ from those of str. GvV257 and GvF24 at 226 and 11 positions, respectively, whereas the GvF12 and Madrid E genomes were found to vary at 869 positions. In comparison, the Breinl and Madrid E genomes were found to differ at 292 positions. These preliminary data indicate that flying squirrel isolates may be more similar to each other than to human isolates [51]. Furthermore, rRNA screening with the MG tool demonstrated differences between *R. prowazekii* str. Madrid E and *R. rickettsii* str. Hlp#2 and other strains.

It has been shown that the ranking of the genomes from *R. rickettsii* and *R. prowazekii* strains is correlated to the Gapped Identity (%) on the National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov/genome/neighbors/674?&genome_assembly_id=300283 and https://www.ncbi.nlm.nih.gov/genome/neighbors/737?genome_assembly_id=168378x, respectively). This correlation requires additional research. A comparative analysis of the degree of virulence in the MG analysis showed that the percentage of coincident genome components correlated with an index of Gapped Identity (%) to a high degree. Thus, the elimination (decrease in the degree) of the virulence of strains is associated with the accumulation of the gapped effect from the highly virulent strains *R. rickettsii* str. Sheila Smith and *R. prowazekii* str. Breinl.

Use of LSH is important for reduction of the amount of information stored and is required for identification of genetic texts. The LSH is an algorithm that results in the nucleotide sequence being recoded into a numerical sequence.

Bacterial taxonomy relies on a polyphasic approach based on the combination of phenotypic and genotypic characteristics (DNA-DNA hybridization, 16S rRNA gene sequence similarity and phylogeny, DNA G+C content). Recently, the polyphasic approach has been adapted to culturomics [52] through the development of taxonogenomics [16]. Taxonogenomics is a polyphasic strategy combining phenotypic characteristics obtainable and comparable by most laboratories with matrix-assisted desorption ionization–time of flight mass spectrometry analysis, genome sequence characteristics and comparison using average genomic identity of orthologous gene sequences (AGIOS), average nucleotide identity (ANI) and/or other genome comparison software for the taxonomic description of new bacterial taxa [15,16,53].

The complete genome of *Candidatus Midichloria mitochondrii* IricVA was sequenced (GenBank accession no. NC_015722.1) from the ovarian tissue of a female *Ixodes ricinus* tick collected in nature (Varese, Italy) [17]. Bioinformatic analysis enabled studying the genotypic characteristics and modelling the phenotypic characteristics of this uncultivated microorganism from the order *Rickettsiales*. We believe that sequencing the genome of an uncultivated microorganism directly from the organs of a naturally infected female tick or using a experimentally infected tick model will develop. Then, after genome sequence annotation, the

genotypic characteristics may be studied and the phenotypic characteristics modelled using bioinformatics analysis. To classify these uncultivated bacteria taxonomically, taxonogenomics may be used. However, to determine the position at the family, genus and species ranks, FOA, which takes into account the arrangement of nucleotides in the genome, may be included in the taxonogenomic analysis. The application of this method may provide an occasion for the International Committee of Systematic Bacteriology to implement a new procedure for recognizing the status of new species among noncultivated bacteria [54].

For the first time, an attempt to compare the genomes from prokaryotes (bacteria and archaea) was adapted from the works of E. Kunin and coauthors in 1997 [55]. The term 'phylogenomics,' proposed by Eisen and Fraser in 2003 [56], covers systematics-based study of genes and genomes, as well as analysis of the evolution of gene families within genomes. Genosystematics has been successfully applied to rickettsiae using 16S rRNA analysis and four protein-coding genes [57].

Conclusion

In this study, a new genomosystematics approach is proposed for the study of complete genomes and their components through the development and application of FOA tools. Its applications include the development of principles for the classification of microorganisms based on the analysis of complete genomes and their annotations. The classification of rickettsial genomes obtained on the basis of FOA has a strong correlation with the taxonomy of arthropods, which are the hosts of rickettsia and which is confirmed by their ecological associations. The objectivity of the classification of rickettsial genomes is confirmed by the classification of rickettsioses, built on the basis of mechanisms of infection by various groups of rickettsia, with these mechanisms being of great importance in the epidemiology and aetiology of rickettsioses.

Thus, genomosystematics (systematics of genomes) underlies the classification of rickettsiae and rickettsioses based on ecological, epidemiologic and aetiologic principles.

Conflict of interest

None declared.

References

[1] Merhej V, Raoult D. Rickettsial evolution in the light of comparative genomics. Biol Rev Camb Philos Soc 2011;86:379–405.

[2] Parola P, Paddock CD, Socolovschi C, Labruna MB, Mediannikov O, Kernif T, et al. Update on tick-borne rickettsioses around the world: a geographic approach. Clin Microbiol Rev 2013;26:657–702.

[3] Raoult D, Weiller PJ, Chagnon A, Chaudet H, Gallais H, Casanova P. Mediterranean spotted fever, clinical, laboratory and epidemiological features of 199 cases. Am J Trop Med Hyg 1986;35:845–50.

[4] Dumler JS, Walker DH. Rocky Mountain spotted fever—changing ecology and persisting virulence. N Engl J Med 2005;353:551–3.

[5] Bechah Y, Capo C, Mege JL, Raoult D. Epidemic typhus. Lancet Infect Dis 2008;8:417–26.

[6] Zinsser H. Rats, lice, and history. London: Broadway House; 1935.

[7] Merhej V, Angelakis E, Socolovschi C, Raoult D. Genotyping, evolution and epidemiological findings of *Rickettsia* species. Infect Genet Evol 2014;25:122–37.

[8] Stothard DR, Clark JB, Fuerst PA. Ancestral divergence of *Rickettsia bellii* from the spotted fever and typhus groups of *Rickettsia* and antiquity of the genus *Rickettsia*. Int J Syst Bacteriol 1994;44:798–804.

[9] Tamura A, Ohashi N, Urakami H, Miyamura S. Classification of *Rickettsia* tsutsugamushi in a new genus, *Orientia* gen. nov., as *Orientia tsutsugamushi* comb. nov. Int J Syst Bacteriol 1995;45:589–91.

[10] Gillespie JJ, Beier MS, Rahman MS, Ammerman NC, Shallom JM, Purkayastha A, et al. Plasmids and rickettsial evolution: insight from *Rickettsia felis*. PLoS One 2007;2. e266.

[11] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10.

[12] Kumar S, Tamura K, Nei M. MEGA: Molecular evolutionary genetics analysis. Pennsylvania State University: University Park; 1993., version 1.01.

[13] Gumenyuk AS, Pozdnichenko N, Rodionov I, Shpynov SN. About the means of formal analysis system of nucleotide chains [in Russian]. Mat Bio Bioinform 2013;8:373–97.

[14] Shpynov S, Pozdnichenko N, Gumenuk A. Approach for classification and taxonomy within family *Rickettsiaceae* based on the formal order analysis. Microbe. Infect 2015;17:839–44.

[15] Mediannikov O, Nguyen TT, Bell-Sakyi L, Padmanabhan R, Fournier PE, Raoult D. High quality draft genome sequence and description of *Occidentia massiliensis* gen. nov., sp. nov., a new member of the family *Rickettsiaceae*. Stand Genomic Sci 2014;9:9. Available at: http://www.standardsingenomics.com/content/9/1/9.

[16] Fournier PE, Lagier JC, Dubourg G, Raoult D. From culturomics to taxonomogenomics: a need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 2015;36:73–8.

[17] Sassera D, Lo N, Epis S, D'Auria G, Montagna M, Comandatore F, et al. Phylogenomic evidence for the presence of a flagellum and cbb(3) oxidase in the free-living mitochondrial ancestor. Mol Biol Evol 2011;28:3285–96.

[18] Schroeder CL, Narra HP, Sahni A, Rojas M, Khanipov K, Patel J, et al. Identification and characterization of novel small RNAs in *Rickettsia prowazekii*. Front Microbiol 2016;8(7):859.

[19] Shpynov SN, Gumenuk AS, Pozdnichenko NN. Application of the numerical characteristic of formal order analysis of the prokaryotic genomes for reclassification within the genus *Rickettsia* [in Russian]. Mat Bio Bioinform 2016;11:336–50.

[20] Afreixo V, Bastos CA, Pinho AJ, Garcia SP, Ferreira PJ. Genome analysis with inter-nucleotide distances. Bioinformatics 2009;25:3064–70.

[21] Gumenuk AS, Pozdnichenko NN, Skiba AA, Shpynov SN. Computer program 'Map of genes' [in Russian]. Certificate of State Registration of the Computer Program in the Register of Computer Programs 2017616730, 13.06. 2017.

[22] Gumenuk AS, Pozdnichenko NN, Skiba AA, Shpynov SN. Computer program 'Matrix of similarity of nucleotide sequences by their components' [in Russian]. Certificate of State Registration of the Computer Program in the Register of Computer Programs 2017616679, 09.06. 2017.

[23] Ignatovich VF. Antigenic relations of *Rickettsia prowazekii* and *Rickettsia Canada*, established in the study of sera of patients with Brill's disease [in Russian]. J Hyg Epidemiol Microbiol Immunol 1977;21:55–60.
[24] Linnemann Jr CC, Pretzman CI, Peterson ED. Acute febrile cerebrovasculitis. A non–spotted fever group rickettsial disease. Arch Intern Med 1989;149:1682–4.
[25] Wenzel RP, Hayden FG, Gröschel DH, Salata RA, Young WS, Greenlee JE, et al. Acute febrile cerebrovasculitis: a syndrome of unknown, perhaps rickettsial, cause. Ann Intern Med 1986;104:606–15.
[26] Jado I, Oteo JA, Aldámiz M, Gil H, Escudero R, Ibarra V, et al. *Rickettsia monacensis* and human disease. Spain Emerg Infect Dis 2007;9:1405–7. V. 13. .
[27] Simser JA, Rahman MS, Dreher-Lesnick SM, Azad AF. A novel and naturally occurring transposon, ISRpe1 in the *Rickettsia peacockii* genome disrupting the *rickA* gene involved in actin-based motility. Mol Microbiol 2005:71–9. V. 58.1. .
[28] Felsheim RF, Kurtti TJ, Munderloh UG. Genome sequence of the endosymbiont *Rickettsia peacockii* and comparison with virulent *Rickettsia rickettsii*: identification of virulence factors. PLoS One 2009;4(12): e8361. V. 21. .
[29] Zdrodovskii PF, Golinevich HM. The rickettsial diseases. London: Pergamon Press; 1960.
[30] Krantz GW, Walter DE. A manual of acarology. 3rd ed. Lubbock: Texas Tech University Press; 2009.
[31] Weiss E, Moulder JW, Order I. Rickettsiales, gieszczkiewicz 1939. In: Krieg NR, Holt JG, editors. Bergey's manual of systematic bacteriology. Baltimore, MD: Williams & Wilkins; 1984. p. 687–703. 1.
[32] Raoult D, Roux V. Rickettsioses as paradigms of new or emerging infectious diseases. Clin Microbiol Rev 1997;10:694–719.
[33] Azad AF, Radulovic S, Higgins JA, Noden BH, Troyer JM. Flea-borne rickettsioses: ecologic considerations. Emerg Infect Dis 1997;3: 319–27.
[34] Civen R, Ngo V. Murine typhus: an unrecognized suburban vector-borne disease. Clin Infect Dis 2008;46:913–8.
[35] Nichols E, Rindge ME, Russell GG. The relationship of the habits of the house mouse and the mouse mite (*Allodermanyssus sanguineus*) to the spread of rickettsialpox. Ann Intern Med 1953;39:92–102.
[36] Traub R, Wisseman Jr CL. The ecology of chigger-borne rickettsiosis (scrub typhus). J Med Entomol 1974;11:237–303.
[37] Shirai A, Saunders JP, Dohany AL, Huxsoll DL, Groves MG. Transmission of scrub typhus to human volunteers by laboratory-reared chiggers. Jpn J Med Sci Biol 1982;35:9–16.
[38] Paris DH, Shelite TR, Day NP, Walker DH. Unresolved problems related to scrub typhus: a seriously neglected life-threatening disease. Am J Trop Meg Hyg 2013;2:301–7. V. 89. .
[39] Markov AV, Zakharov IA. The use of quantitative measures of similarity of gene orders for phylogenetic reconstructions build on the example of the genus *Rickettsia* bacteria [in Russian]. Genetics 2008;44: 456–66.
[40] Angelakis E, Mediannikov O, Parola P, Raoult D. *Rickettsia felis*: the complex journey of ans emergent human pathogen. Trends Parasitol 2016;32:554–64.
[41] Dieme C, Bechah Y, Socolovschi C, Audoly G, Berenger JM, Faye O, et al. Transmission potential of *Rickettsia felis* infection by *Anopheles gambiae* mosquitoes. Proc Natl Acad Sci USA 2015;112:8088–93.
[42] Mediannikov O, Socolovschi C, Edouard S, Fenollar F, Mouffok N, Bassene H, et al. Common epidemiology of *Rickettsia felis* infection and malaria, Africa. Emerg Infect Dis 2013;Nov;19:1775–83.
[43] Eremeeva ME, Dasch GA, Silverman DJ. Quantitative analyses of variations in the injury of endothelial cells elicited by 11 isolates of *Rickettsia rickettsii*. Clin Diagn Lab Immunol 2001;8:788–96.
[44] Eremeeva ME, Klemt RM, Santucci-Domotor LA, Silverman DJ, Dasch GA. Genetic analysis of isolates of *Rickettsia rickettsii* that differ in virulence. Ann N Y Acad Sci 2003;990:717–22.
[45] Eremeeva ME, Dasch GA. Closing the gaps between genotype and phenotype in *Rickettsia rickettsii*. Ann N Y Acad Sci 2009;1166:12–26.
[46] Clark TR, Noriea NF, Bublitz DC, Ellison DW, Martens C, Lutter EI, et al. Comparative genome sequencing of *Rickettsia rickettsii* strains that differ in virulence. Infect Immun 2015;83:1568–76.
[47] Ellison DW, Clark TR, Sturdevant DE, Virtaneva K, Porcella SF, Hackstadt T. Genomic comparison of virulent *Rickettsia rickettsii* Sheila Smith and avirulent *Rickettsia rickettsii* Iowa. Infect Immun 2008;76: 542–50.
[48] Turco J, Winkler HH. Differentiation between virulent and avirulent strains of *Rickettsia prowazekii* by macrophage-like cell lines. Infect Immun 1982;35:783–91.
[49] Bechah Y, El Karkouri K, Mediannikov O, Leroy Q, Pelletier N, Robert C, et al. Genomic, proteomic, and transcriptomic analysis of virulent and avirulent *Rickettsia prowazekii* reveals its adaptive mutation capabilities. Genome Res 2010;20:655–63.
[50] Ge H, Chuang YY, Zhao S, Tong M, Tsai MH, Temenak JJ, et al. Comparative genomics of *Rickettsia prowazekii* Madrid E and Breinl strains. J Bacteriol 2004;186:556–65.
[51] Bishop-Lilly KA, Ge H, Butani A, Osborne B, Verratti K, Mokashi V, et al. Genome sequencing of four strains of *Rickettsia prowazekii*, the causative agent of epidemic typhus, including one flying squirrel isolate. Genome Announc 2013;1. e00399–13.
[52] Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.
[53] Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 2014;64(Pt 2):384–91.
[54] Samoilenko IE, Kumpan LV, Shpynov SN, Obert AS, Butakov OV, et al. Methods of isolation and cultivation of new *rickettsiae* from the Nosoarea of the North Asian tick typhus in siberia. Ann N Y Acad Sci 2006;1078:613–6.
[55] Eisen JA, Fraser CM. Phylogenomics: intersection of evolution and genomics. Science 2003;300(5626):1706–7.
[56] Koonin EV, Mushegian AR, Galperin MY, Walker DR. Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the *Archaea*. Mol Microbiol 1997;25:619–37.
[57] Fournier PE, Dumler JS, Greub G, Zhang J, Wu Y, Raoult D. Gene sequence–based criteria for identification of new rickettsia isolates and description of *Rickettsia heilongjiangensis* sp. nov. J Clin Microbiol 2003;41:5456–65.