PLOS ONE

RESEARCH ARTICLE

Distinction between *Borrelia* and *Borreliella* is more robustly supported by molecular and phenotypic characteristics than all other neighbouring prokaryotic genera: Response to Margos' et al. "The genus *Borrelia* reloaded" (PLoS ONE 13(12): e0208432)

Radhey S. Gupta *

Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada

* gupta@mcmaster.ca

Check for updates

OPEN ACCESS

Citation: Gupta RS (2019) Distinction between *Borrelia* and *Borreliella* is more robustly supported by molecular and phenotypic characteristics than all other neighbouring prokaryotic genera: Response to Margos' et al. "The genus *Borrelia* reloaded" (PLoS ONE 13(12): e0208432). PLoS ONE 14(8): e0221397. https://doi.org/10.1371/journal.pone.0221397

Editor: Sven Bergström, Umeå University, SWEDEN

Received: January 9, 2019

Accepted: July 25, 2019

Published: August 27, 2019

Copyright: © 2019 Radhey S. Gupta. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This study was funded by grant no 249924 to RSG from the Natural Science and Engineering Research Council of Canada, https://www.canada.ca/en/science-engineering-research.html. The funders had no role in study design, data

Abstract

In a recent publication in PLOS ONE, Gabriele Margos and colleagues have questioned the division of the genus *Borrelia* into two genera on the basis that the differences in percentage of conserved proteins (POCP) between these two groups is >50%, which an earlier study has suggested as the threshold for differentiating prokaryotic genera. However, the POCP threshold is a poorly characterized and rarely used criterion for establishing distinction among prokaryotic genera. Detailed evaluation of the intergeneric POCP values for 37 genera from 3 different families (viz. *Enterobacteriaceae*- 24 genera, *Morganellaceae*-8 genera and *Cystobacteraceae*-5 genera) presented here shows that the POCP values for all genera within each of these families exceeded >58%. Thus, the suggested POCP threshold *is not a useful criterion for delimitation of genus boundary* and the objection by Margos *et al*. on this ground is invalid. Additionally, Margos *et al*. have questioned the specificities of ~15–20% of the conserved signature indels (CSIs) described in our work. However, as shown here, this concern is due to misunderstanding of the results and the CSIs in question are still highly-specific characteristics of the members of these genera and they provide important information regarding the evolutionary relationships of two new reptiles-echidna-related species viz. *Borrelia turcica* and Candidatus Borrelia tachyglossi to other *Borrelia* species. Results presented here show that both these species are deeper-branching members of the genus *Borrelia* and their placement within this genus is strongly supported by phylogenetic analyses and multiple uniquely shared CSIs with the other *Borrelia* species. Based on the large body of evidence derived from phylogenetic, genomic, molecular, phenotypic and clinical features, it is contended that the characteristics clearly distinguishing the *Borrelia* and *Borreliella* genera are far more numerous and of different kinds than those discerning most (all) other neighbouring genera of prokaryotes. Thus, the placement of these two groups of microorganisms into distinct genera, *Borrelia* and *Borreliella*, which clearly recognizes the differences among them, is highly appropriate and it should lead to a better understanding

collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The author has declared that no competing interests exist.

of the clinical, molecular and biological differences between these two important groups of microbes.

Introduction

The family *Borreliaceae* includes species that are causative agents of Lyme disease (LD) and others that are causative agents of tick- and louse-borne relapsing fever (RF) [1–5]. Our earlier comprehensive phylogenomics and comparative studies on protein sequences from *Borreliaceae* genomes provided compelling evidence for the existence of two genetically distinct groups of organisms within this family [6,7]. Of these two groups, one group included all species that are causative agents of the clinically distinctive disorder known as RF, whereas the second group encompassed causative agents of LD along with some other closely related species [6]. The existence of these two groups was supported by different independent lines of evidence which included: (i) Distinct branching of the LD and RF groups of species in the 16S rRNA trees and multiple genome scale trees based on protein sequences [6–8]; (ii) Clear distinction of the LD and RF groups of species based on pairwise comparison of either the average nucleotide identity (ANI) or the average amino acid identity (AAI) of different genes/proteins from the *Borreliaceae* genomes [6,7]; (iii) Identification of >70 highly-specific molecular signatures consisting of conserved signature insertions/deletions (indels) (CSIs) in protein sequences and conserved signature proteins (CSPs) that are exclusively shared by different members of either the LD or the RF group of species [6,7]; and (iv) Several phenotypic characteristics known from earlier work including the distinct pathogenicity profiles of the two groups of organisms and differences in arthropod vectors used by them [1,2,4,5]. Based on the robust evidence provided by all of these analyses, we have previously proposed a division of the family *Borreliaceae* (and the genus *Borrelia*) into two main genera, *Borrelia* and *Borreliella* [6]. In this proposal, all of the species that are part of the RF group were retained within the genus *Borrelia*, whereas all species related to the LD group were placed into a new genus called *Borreliella* [6]. This latter group of species is widely referred to by the name "*Borrelia burgdorferi sensu lato*", recognizing their distinctness from the RF group of species [1,2,5,9].

Recently, Margos *et al.* [10,11] have analyzed the genome sequences from two new *Borreliaceae* species, viz. *Borrelia turcica* and Candidatus Borrelia tachyglossi, which are associated with reptiles and echidna. In their publication, Margos *et al.* [10] acknowledge that the LD and RF groups of species "have different clinical, biological, and epidemiological characteristics, and phylogenetic data is concordant with this, demonstrating that these two groups are genetically similar yet distinct and form independent monophyletic sister clades that once shared a common ancestor". Additionally, they state that the proposal by Adeolu and Gupta [6] to divide the genus *Borrelia* into two genera "was largely based on the identification of conserved signature insertions/deletions (indels) (CSIs) and conserved signature proteins (CSPs) that are differentially present in the LD or RF *Borrelia* genogroup, as well as average nucleotide identity (ANI) values calculated between whole genomes of 18 *Borrelia* species including eight LD species and ten RF species . . . it is uncontested that these differences exist between LD and RF *Borrelia*".

Margos *et al.* [10] have questioned the division of genus *Borrelia* into two genera on three accounts. Their main argument for questioning the division is based on the consideration that the differences in the ANI or AAI values between these two groups, shown in our work as part of the evidence indicating that these two groups of species differ from each other [6,7], are not

suitable means for differentiation of prokaryotic genera. Instead, they assert that a method proposed by Qin *et al*. [12] based on percentage of conserved proteins (POCP) between genomes from different species is a more reliable means for the determination of a genus level boundary. In addition to this main objection, the authors also criticize our work on two other grounds: (i) that the methodology used in our work only identifies CSIs and CSPs that are exclusive to only one *Borrelia* genogroup and it precludes the detection of those characteristics that are shared non-exclusively between both genogroups, and (ii) that upon inclusion of sequence information for the two new *Borrelia* species, about 17–20% of the previously reported 53 CSIs are unable to differentiate between the LD and the RF groups of species. I discuss below our responses to all of these criticisms and specifically the problem of using or relying on the suggested POCP threshold as a criterion for the delimitation of prokaryotic genera, which is the main basis of Margos *et al*.'s [10] resistance to our division of the genus *Borrelia* into two genera.

Materials and methods

Construction of the phylogenetic tree based on the core genome proteins and the calculation of percentage of conserved proteins (POCP) between different genomes was carried out using an internally developed software pipeline described in earlier work [13–15]. Information regarding genome sequences for different species from the families *Enterobacteriaceae* [15], *Morganellaceae* [14] and *Cystobacteraceae*, for which the POCP values were calculated is provided in S1 and S2 Tables. Briefly, using the CD-HIT program [16], proteins sharing a minimum of 50% sequence identity and sequence length were identified in different genomes. Based on this information, the POCP between different pairs of genomes was calculated as described by Qin *et al*. [12]. Multiple sequence alignment (MSA) of the proteins which were found in at least 80% of the input genomes (a total of 703 protein families) were created using Clustal Omega [17]. For phylogenetic analysis, the sequence alignments were trimmed using TrimAl [18] before their concatenation into a single file. The combined sequence for the 703 core genome proteins, which after trimming consisted of 248452 aligned amino acids, was utilized for phylogenetic analysis. A maximum likelihood (ML) tree based on this sequence alignment was constructed and optimized in RAxML 8 as described in our earlier work [13–15].

The 16S rRNA gene sequences for different *Borreliaceae* species were downloaded from All-Species Living Tree Project [19] and aligned using ClustalX2. The tree was constructed using the Maximum-likelihood (ML) method in MEGA6 [20]. Updating of the sequence information and group specificity of different CSIs and CSPs was carried out by performing BLASTp searches on the sequences of the indicated proteins. Formatting of the sequence alignment files was carried out using SIG_CREATE and SIG_STYLE programs described in our work [21]. It should be mentioned that based on different lines of evidence, the following *Borrelia* species (viz. *B. bissettii*, *B. lanei*, *B. mayonii* and *B. yangtzensis*) consistently group with the LD group. Unlike the other LD group of species, which are now transferred to the genus *Borreliella* [6], the proposal to reclassify these four species to the genus Borreliella has not yet been made. However, in the interim, to avoid any confusion due to the grouping of these *Borrelia* species within other *Borreliella* species, the genus name of these species is abbreviated as "Bor." in the manuscript and different Figs.

Results and discussion

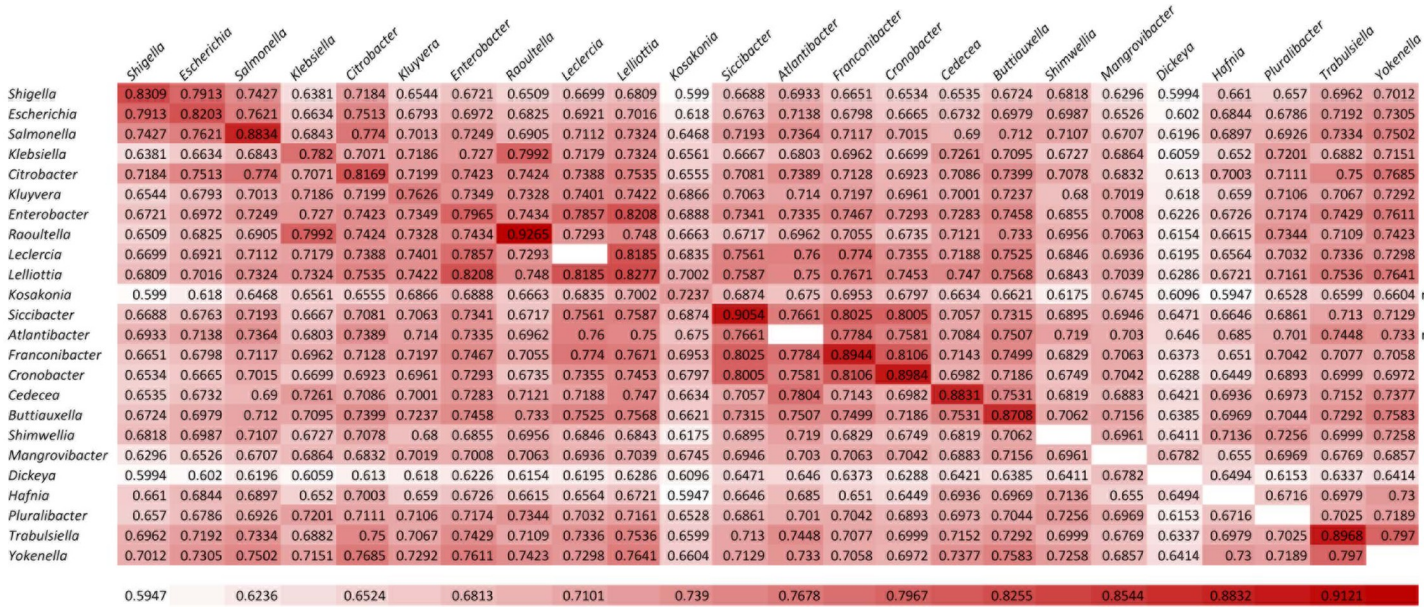
The inadequacy of using a 50% POCP threshold for genus level boundaries

Prokaryotic systematics involves assemblage of organisms into groups of different ranks from most inclusive to least inclusive (e.g. Phylum, Class, Order, Family, Genus and Species) on the

basis of their observed similarities and differences and phylogenetic/evolutionary relationships [22–25]. Species are the basic unit of any biological classification scheme. For prokaryotic organisms, although a formal definition of "the species" is lacking, for practical purposes, it is now generally accepted that strains showing >70% similarity in DNA-DNA hybridization values, or >98.65% sequence similarity in 16S rRNA, or those exhibiting >95% similarity in ANI values provide comparable means for delimiting a prokaryotic species and for identification of new species [23,25–33]. In contrast to these accepted criteria for species delimitation, *there are no commonly accepted or used criteria for identification of genus or higher level taxa* [34]. A genus is commonly defined as *"a monophyletic grouping of species with many characters in common"*[22,35]. Further, there is a general consensus that the division into higher taxonomic ranks including genus level taxon should reflect phylogenetic relationships.

While there are no accepted criteria for genus level boundaries, some authors have suggested that the 16S rRNA similarity values between 94.5% and 86.6% [26] or the POCP values <50% [12] can be used as thresholds for differentiation among genera. However, these suggestions are based on studies using a limited number of prokaryotic taxa and the general utilities of these methods (or suggested thresholds) for delimitation of prokaryotic genera remains to be properly evaluated. Let us now specifically consider the utility of using the 50% POCP threshold value as a genus level boundary, which Margos *et al*. [10] contend provides a more suitable method for demarcation of prokaryotic genera. The study by Qin *et al*. [12], which suggested the use of POCP values for genus level delimitation was based on a limited number of prokaryotic taxa and the inferences based on it suffer from a number of drawbacks: (i) Interspecies POCP comparison in this study was carried out for only 17 genera. Of these, several genera such as *Bacillus*, *Lactobacillus* and *Clostridium* are highly polyphyletic and only a selected group of closely related species were chosen from them to represent the entire genera [12]. Due to arbitrary delimitation of these genera to a small group of selected species, the closest relatives of these genera, which are other species from the same genera, were not considered in either the interspecies or intergeneric POCP comparisons. (ii) Intergeneric POCP comparisons were carried out with only 1 arbitrarily chosen species from these 17 genera to only single species from other genera, families and orders of bacteria [12]. As many of these latter comparisons were made for species that are part of different families or orders of bacteria, the POCP values obtained for them do not reflect intergeneric differences, but rather family or order level differences. The latter values are expected to be lower than intergeneric differences and the results from such comparisons should not have been included in the comparison of intergeneric POCP values as they artificially lower the observed intergeneric values. (iii) Several genera used for interspecies comparison viz. *Thermotogae*, *Clostridium*, *Mycobacterium*, for which the POCP values were indicated to be higher than 50%, have since been divided into multiple genera [13,36–38] indicating that the POCP threshold is not a useful or required criterion for genus level separation.

To further evaluate the usefulness of POCP values for genus level separation/boundary, we have independently determined interspecies and intergeneric POCP values for a number of families each containing multiple genera. Three well-studied families that we have examined in this regard include the family *Enterobacteriaceae* containing 24 genera [15], the family *Morganellaceae* containing 8 genera [14], and the family *Cystobacteraceae* containing 5 genera [39]. For all of these families, pairwise interspecies and intergeneric POCP values were determined for all species for which genome sequences were available. From the pairwise POCP matrix, average POCP values were determined for different species within each genus (interspecies POCP values) and for different genera within each of these three families (intergeneric POCP values). The results of these comparisons for the family *Enterobacteriaceae* are presented in the pairwise POCP matrix in Fig 1.



	Shigella	Escherichia	Salmonella	Klebsiella	Citrobacter	Kluyvera	Enterobacter	Raoultella	Leclercia	Lelliottia	Kosakonia	Siccibacter	Atlantibacter	Franconibacter	Cronobacter	Cedecea	Buttiauxella	Shimwellia	Mangrovibacter	Dickeya	Hafnia	Pluralibacter	Trabulsiella	Yokenella
Shigella	0.8309	0.7913	0.7427	0.6381	0.7184	0.6544	0.6721	0.6509	0.6699	0.6809	0.599	0.6688	0.6933	0.6651	0.6534	0.6535	0.6724	0.6818	0.6296	0.5994	0.661	0.657	0.6962	0.7012
Escherichia	0.7913	0.8203	0.7621	0.6634	0.7513	0.6793	0.6972	0.6825	0.6921	0.7016	0.618	0.6763	0.7138	0.6798	0.6665	0.6732	0.6979	0.6987	0.6526	0.602	0.6844	0.6786	0.7192	0.7305
Salmonella	0.7427	0.7621	0.8834	0.6843	0.774	0.7013	0.7249	0.6905	0.7112	0.7324	0.6468	0.7193	0.7364	0.7117	0.7015	0.69	0.712	0.7107	0.6707	0.6196	0.6897	0.6926	0.7334	0.7502
Klebsiella	0.6381	0.6634	0.6843	0.782	0.7071	0.7186	0.727	0.7992	0.7179	0.7324	0.6561	0.6667	0.6803	0.6962	0.6699	0.7261	0.7095	0.6727	0.6864	0.6059	0.652	0.7201	0.6882	0.7151
Citrobacter	0.7184	0.7513	0.774	0.7071	0.8169	0.7199	0.7423	0.7424	0.7388	0.7535	0.6555	0.7081	0.7389	0.7128	0.6923	0.7086	0.7399	0.7078	0.6832	0.613	0.7003	0.7111	0.75	0.7685
Kluyvera	0.6544	0.6793	0.7013	0.7186	0.7199	0.7626	0.7349	0.7328	0.7401	0.7422	0.6866	0.7063	0.714	0.7197	0.6961	0.7001	0.7237	0.68	0.7019	0.618	0.659	0.7106	0.7067	0.7292
Enterobacter	0.6721	0.6972	0.7249	0.727	0.7423	0.7349	0.7965	0.7434	0.7857	0.8208	0.6888	0.7341	0.7335	0.7467	0.7293	0.7283	0.7458	0.6855	0.7008	0.6226	0.6726	0.7174	0.7429	0.7611
Raoultella	0.6509	0.6825	0.6905	0.7992	0.7424	0.7328	0.7434	0.9265	0.7293	0.748	0.6663	0.6717	0.6962	0.7055	0.6735	0.7121	0.733	0.6956	0.7063	0.6154	0.6615	0.7344	0.7109	0.7423
Leclercia	0.6699	0.6921	0.7112	0.7179	0.7388	0.7401	0.7857	0.7293		0.8185	0.6835	0.7561	0.76	0.774	0.7355	0.7188	0.7525	0.6846	0.6936	0.6195	0.6564	0.7032	0.7336	0.7298
Lelliottia	0.6809	0.7016	0.7324	0.7324	0.7535	0.7422	0.8208	0.748	0.8185	0.8277	0.7002	0.7587	0.75	0.7671	0.7453	0.747	0.7568	0.6843	0.7039	0.6286	0.6721	0.7161	0.7536	0.7641
Kosakonia	0.599	0.618	0.6468	0.6561	0.6555	0.6866	0.6888	0.6663	0.6835	0.7002	0.7237	0.6874	0.675	0.6953	0.6797	0.6634	0.6621	0.6175	0.6745	0.6096	0.5947	0.6528	0.6599	0.6604
Siccibacter	0.6688	0.6763	0.7193	0.6667	0.7081	0.7063	0.7341	0.6717	0.7561	0.7587	0.6874	0.9054	0.7661	0.8025	0.8005	0.7057	0.7315	0.6895	0.6946	0.6471	0.6646	0.6861	0.713	0.7129
Atlantibacter	0.6933	0.7138	0.7364	0.6803	0.7389	0.714	0.7335	0.6962	0.76	0.75	0.675	0.7661		0.7784	0.7581	0.7084	0.7507	0.719	0.703	0.646	0.685	0.701	0.7448	0.733
Franconibacter	0.6651	0.6798	0.7117	0.6962	0.7128	0.7197	0.7467	0.7055	0.774	0.7671	0.6953	0.8025	0.7784	0.8944	0.8106	0.7143	0.7499	0.6829	0.7063	0.6373	0.651	0.7042	0.7077	0.7058
Cronobacter	0.6534	0.6665	0.7015	0.6699	0.6923	0.6961	0.7293	0.6735	0.7355	0.7453	0.6797	0.8005	0.7581	0.8106	0.8984	0.6982	0.7186	0.6749	0.7042	0.6288	0.6449	0.6893	0.6999	0.6972
Cedecea	0.6535	0.6732	0.69	0.7261	0.7086	0.7001	0.7283	0.7121	0.7188	0.747	0.6634	0.7057	0.7804	0.7143	0.6982	0.8831	0.7531	0.6819	0.6883	0.6421	0.6936	0.6973	0.7152	0.7377
Buttiauxella	0.6724	0.6979	0.712	0.7095	0.7399	0.7237	0.7458	0.733	0.7525	0.7568	0.6621	0.7315	0.7507	0.7499	0.7186	0.7531	0.8708	0.7062	0.7156	0.6385	0.6969	0.7044	0.7292	0.7583
Shimwellia	0.6818	0.6987	0.7107	0.6727	0.7078	0.68	0.6855	0.6956	0.6846	0.6843	0.6175	0.6895	0.719	0.6829	0.6749	0.6819	0.7062		0.6961	0.6411	0.7136	0.7256	0.6999	0.7258
Mangrovibacter	0.6296	0.6526	0.6707	0.6864	0.6832	0.7019	0.7008	0.7063	0.6936	0.7039	0.6745	0.6946	0.703	0.7063	0.7042	0.6883	0.7156	0.6961		0.6782	0.655	0.6969	0.6769	0.6857
Dickeya	0.5994	0.602	0.6196	0.6059	0.613	0.618	0.6226	0.6154	0.6195	0.6286	0.6096	0.6471	0.646	0.6373	0.6288	0.6421	0.6385	0.6411	0.6782		0.6494	0.6153	0.6337	0.6414
Hafnia	0.661	0.6844	0.6897	0.652	0.7003	0.659	0.6726	0.6615	0.6564	0.6721	0.5947	0.6646	0.685	0.651	0.6449	0.6936	0.6969	0.7136	0.655	0.6494		0.6716	0.6979	0.73
Pluralibacter	0.657	0.6786	0.6926	0.7201	0.7111	0.7106	0.7174	0.7344	0.7032	0.7161	0.6528	0.6861	0.701	0.7042	0.6893	0.6973	0.7044	0.7256	0.6969	0.6153	0.6716		0.7025	0.7189
Trabulsiella	0.6962	0.7192	0.7334	0.6882	0.75	0.7067	0.7429	0.7109	0.7336	0.7536	0.6599	0.713	0.7448	0.7077	0.6999	0.7152	0.7292	0.6999	0.6769	0.6337	0.6979	0.7025	0.8968	0.797
Yokenella	0.7012	0.7305	0.7502	0.7151	0.7685	0.7292	0.7611	0.7423	0.7298	0.7641	0.6604	0.7129	0.733	0.7058	0.6972	0.7377	0.7583	0.7258	0.6857	0.6414	0.73	0.7189	0.797	
	0.5947		0.6236		0.6524		0.6813		0.7101		0.739		0.7678		0.7967		0.8255		0.8544		0.8832		0.9121	

Fig 1. A comparison matrix showing the averages of the percentage of conserved proteins (POCP) within and between different genera of the family *Enterobacteriaceae*. POCP was determined for all genome sequenced species from the family *Enterobacteriaceae* detailed in our earlier work [26]. The values along the diagonal shows the average POCP values for different species within a given genus (i.e. interspecies values), whereas all other values represent average intergeneric POCP values for different genera within this family. The blank cells indicate that only a single species was available for these genera and hence their interspecies values could not be calculated.

https://doi.org/10.1371/journal.pone.0221397.g001

As seen from the matrix in Fig 1, the intergeneric POCP values for all 24 genera that are part of this family range from a low of 59.4% to a high of 82.0% and they are all higher than 50%. Similarly, the intergeneric POCP values for the 8 genera that are part of the family *Morganellaceae* range from 58.7% to 84.5% (S3 Table), and for the family *Cystobacteraceae*, they range from 61.9%– 82.4% (S3 Table). Thus, if a POCP cut-off value of <50% was to be used for genus level boundary, then all of the different genera present within each of these three families would be part of a single genus. These results demonstrate that the usefulness of the 50% POCP threshold value for determination of genus level boundaries is very limited, if any.

Margos *et al.* [10] have also presented a comparison of the POCP values for the *Borrelia* and *Borreliella* genera along with some other genera within the phylum Spirochaeta. However, of the four other genera for which the POCP comparisons were made, *Brachyspira* and *Leptospira* are part of two separate orders viz. *Brachyspirales* and *Leptospirales* within the phylum Spirochaeta [40,41]. Based on the 16S rRNA sequence similarity comparisons, Yarza *et al.* [26] have previously noted that the species from these two orders, which are very distantly related to each other as well as other orders within the phylum Spirochaeta, should in fact be assigned *class* level ranks within the phylum. Thus, a comparison of the POCP values for these two genera with the other genera is misleading as they provide an indication of the order or class level differences and not intergeneric differences. The other two genera included in the comparison are *Treponema* and *Spirochaeta*. Although both of these genera are part of the family *Spirochaetaceae* [40,41], in phylogenetic trees, members of these genera form different clades indicating extensive divergence (unpublished results) [19,41,42]. Based on the results shown by Margos *et al.* (S1 Table of their publication) [10], the interspecies POCP values for members of these two genera are mostly in the range of 20–40% with an average POCP value of 33.7% for the *Treponema* species and 35.5% for the *Spirochaeta* species. Based on the 50% POCP threshold value for genus level boundaries, the species from both *Treponema* and *Spirochaeta* genera

should each be divided into multiple genera. These results again point to the inadequacy of using the suggested POCP threshold value as a reliable means for the genus level boundaries.

Although a specific POCP value is not very useful for establishing a genus level boundary, a comparison matrix based on POCP, similar to the matrices based on ANI or AAI values, can still provide an overall indication of the genomic similarity and differences between two closely related groups of species. In the POCP matrix presented by Margos *et al.* [10], while the species from the genus *Borreliella* (LD group) exhibited a high degree of similarity to each other, the species from the *Borrelia* (RF) group exhibited considerable variability and this group was not clearly differentiated. However, the POCP matrix constructed by Margos *et al.* [10] was based on genome sequences that included genes present on both the linear chromosomes as well as different plasmids. The distribution of plasmids is highly variable in different *Borreliaceae* species/strains unlike the conservation of linear chromosome structure and chromosomal genes, [5,43–46] and inclusion of plasmid sequences will introduce considerable variability in genome sequence or POCP comparison. Thus, in order to reliably compare the POCP values among different species, such comparisons should be based only on the chromosomal genes not including the plasmid genes. A POCP matrix for the *Borreliaceae* species based on genes present on chromosomal sequences is presented in Fig 2. As seen, this matrix clearly distinguishes the *Borreliaceae* species into two groups corresponding to the *Borrelia* and *Borreliella* genera. Based on this matrix, the average POCP for species from the genera *Borrelia* and *Borreliella* are 93.4% and 94.7%, respectively, whereas the average POCP value between these two groups is only 82.2%. Thus, a comparison of the POCP values based on chromosomal genes actually supports the genetic distinction between the *Borrelia* and *Borreliella* genera.

Specificity of the molecular signatures for the *Borrelia* and *Borreliella* genera

In the Margos *et al.* [10] paper, concerns were also raised regarding our methodology for identifying CSIs and CSPs, which they assert only considered those molecular signatures which

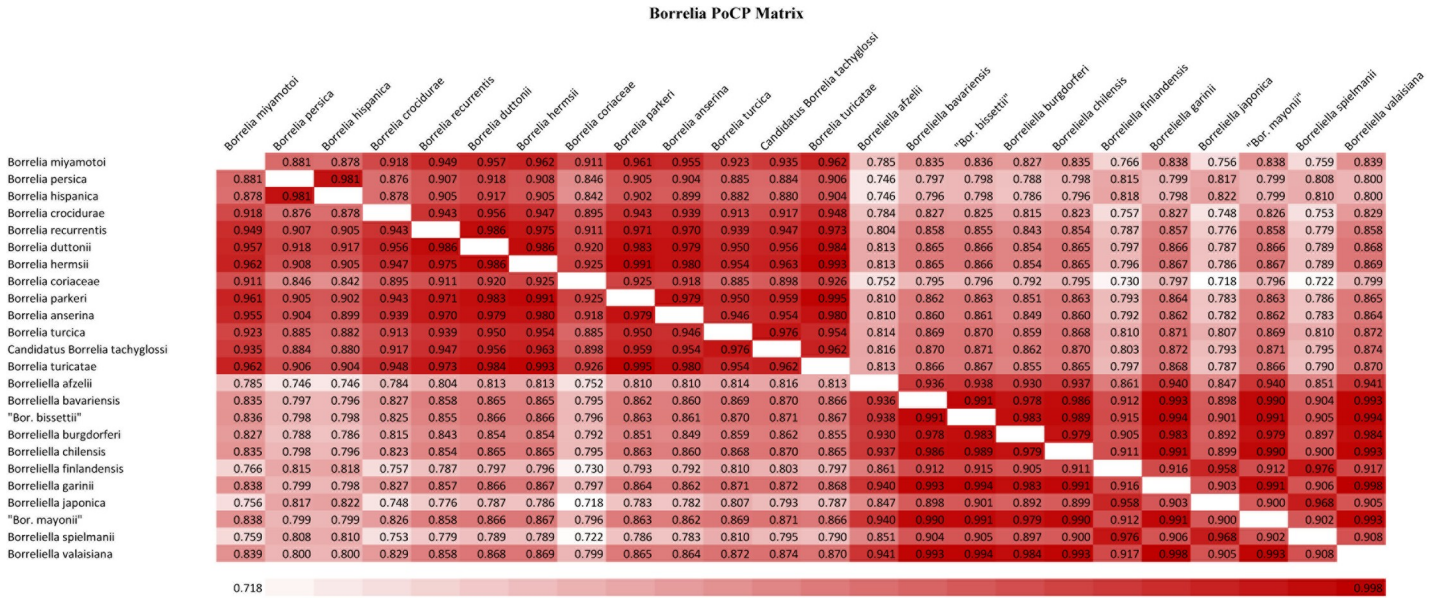


Fig 2. A pair-wise comparison matrix based on percentage of conserved proteins (POCP) in chromosomal genes from different genome sequenced *Borreliaceae* species. The matrix was constructed using an internally developed pipeline [13,14]. Genome pairs sharing higher POCP are shaded more darkly (red). Based on their POCP values, species belonging to the family *Borreliaceae* form two main groups, with one group containing all of the LD and related species (or *Borreliella*), and the other encompassing RF group of species together with the reptile-and echidna- associated species *B. turcica* and Candidatus Borrelia tachyglossi (genus *Borrelia*).

https://doi.org/10.1371/journal.pone.0221397.g002

were exclusively found in one *Borrelia* genogroup and precluded detection of such characteristics that are shared non-exclusively between both genogroups. However, in our original work, in addition to the CSIs and CSPs that are specific for the two main groups (viz. *Borrelia* and *Borreliella*), we also reported 31 CSIs and 82 CSPs that are specifically found in all *Borreliaceae* species [6,41]. This information was also provided and emphasized in our rebuttal response [7] to an earlier criticism of our work by these authors [47]. By non-exclusive, however, if Margos *et al*. [10] mean that the CSIs or CSPs are commonly shared by only some members from each of the two main clades of *Borreliaceae* species, then in our work we have not come across significant number of such characteristics showing any specific pattern. However, isolated characteristics of this kind can result from lateral gene transfers and they are not useful for understanding evolutionary relationships or for taxonomic purposes [32,48].

Margos *et al*. [10] also state that between 17–20% of the CSIs identified by us are not specific for *Borrelia* or *Borreliella* genera and do not differentiate between these two groups. However, subsequent to our earlier work describing the specificities of the CSIs for two *Borreliaceae* genera [6], genome sequences have become available for two new *Borrelia* isolates viz. *B*. *turcica* and Candidatus Borrelia tachyglossi [10,11], and they were included by Margos *et al*. [10] in their analyses. Of these two species/strains, *B*. *turcica* is associated with reptiles whereas Candidatus Borrelia tachyglossi was isolated from an echidna (*Tachyglossus aculeatus*) species [11]. In phylogenetic trees based on 16S rRNA sequences as well as multiple genome-scale phylogenetic trees and trees based on individual protein sequences (Fig 3), these two species form deeper branching lineages of the *Borrelia* (RF) clade [10,11]. Although in a number of trees, particularly those based on large datasets of protein sequences (Fig 3A) [10], these two species form a clade, such an association is often not seen in trees based on sequences for many individual proteins (see Fig 3B) or in the tree based on 16S rRNA gene sequences (Fig 3C). However, we will refer to *B*. *turcica* and Candidatus Borrelia tachyglossi as the Reptiles-related (RR) group/clade in this work.

The inclusion of these two new species in the dataset, depending upon their branching position, is expected to alter the specificity of some of the identified signatures. In our earlier rebuttal response to Margos *et al*. [7,47], we had clearly outlined the different scenarios of how the inclusion of sequence information for the RR group of species, depending upon their branching positon within the family *Borreliaceae*, will affect the group-specificity of some of the identified CSIs. It was stated that if "the RR species/strains branch either within the RF group or as an outgroup of this clade, then such a group of species is expected to contain either some or all of the signatures for the RF clade, but generally none for the LD group"[7]. This is exactly what is observed upon the inclusion of sequence information for *B*. *turcica* and Candidatus Borrelia tachyglossi sequences. Thus, the questions raised by Margos *et al*. [10], regarding the specificities of some of the CSIs indicate that they are misinterpreting the results for the species distribution of the indicated CSIs.

To go over their objections, let us consider the results for different CSIs that were reported previously and how they have been affected upon the inclusion of sequence information for *B*. *turcica* and Candidatus Borrelia tachyglossi. As noted earlier, 31 identified CSIs were specific for the family *Borreliaceae* (Table 2 in Ref. [6]). These CSIs, as expected, are also present in protein homologs from *B*. *turcica* and Candidatus Borrelia tachyglossi (results not shown). The remaining CSIs, which distinguished the two main groups within the family *Borreliaceae* were/are of two kinds. Of these, the first category of 15 CSIs are in proteins whose homologs besides the family *Borreliaceae* are also found in other bacteria (i.e. outgroup species) (Fig 4A and 4B). Based on the presence or absence of these CSIs in the outgroup species, one can infer whether these CSIs represent an insert(s) or deletion(s) and at what specific stage in the evolution of *Borreliaceae* family the genetic changes responsible for these CSIs have occurred

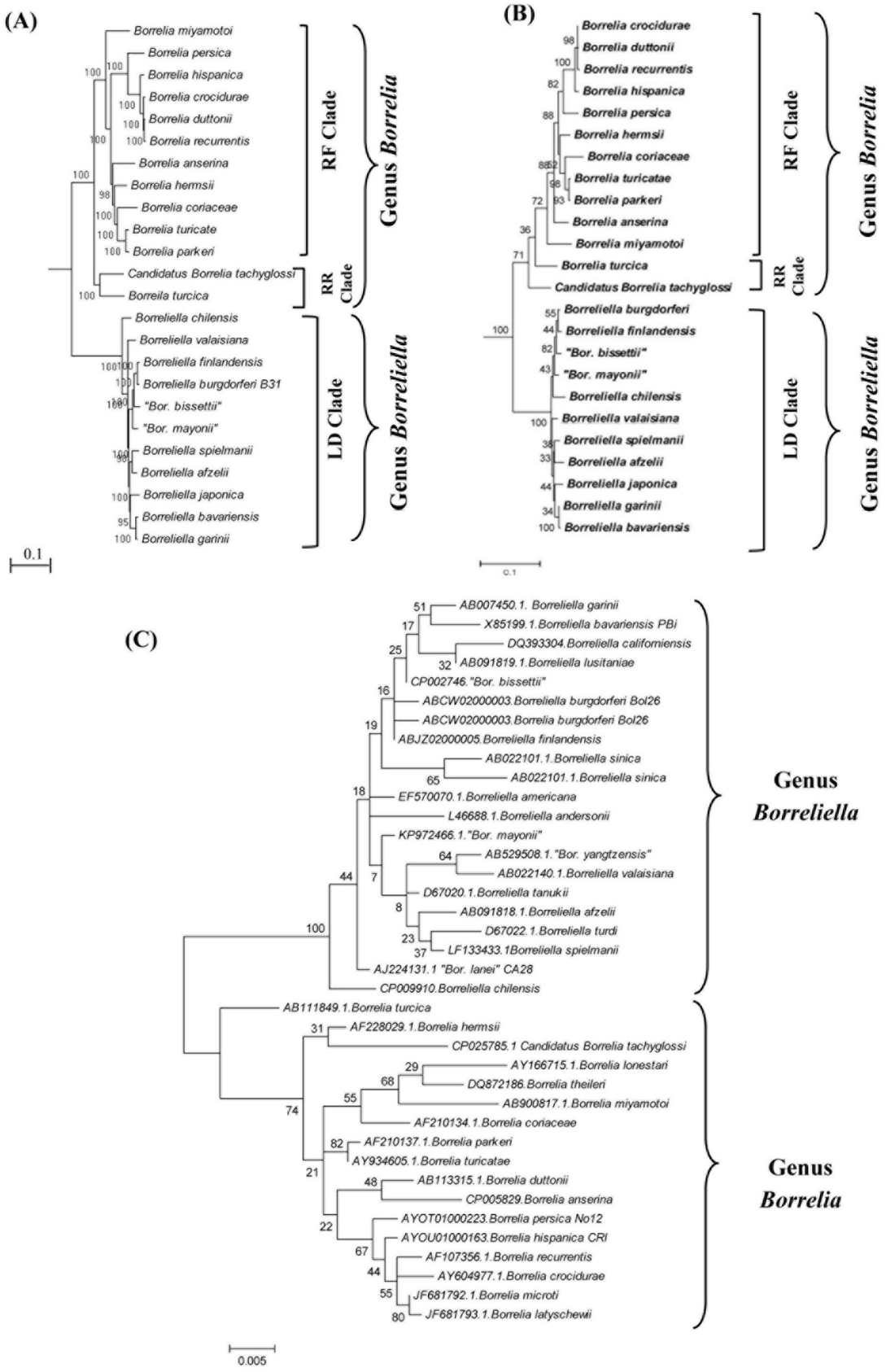

(A)
Borrelia miyamotoi
Borrelia persica
Borrelia hispanica
Borrelia crocidurae
Borrelia duttonii
Borrelia recurrentis
Borrelia anserina
Borrelia hermsii
Borrelia coriaceae
Borrelia turicate
Borrelia parkeri
Candidatus Borrelia tachyglossi
Borreila turcica
Borreliella chilensis
Borreliella valaisiana
Borreliella finlandensis
Borreliella burgdorferi B31
"Bor. bissettii"
"Bor. mayonii"
Borreliella spielmanii
Borreliella afzelii
Borreliella japonica
Borreliella bavariensis
Borreliella garinii
0.1
RF Clade
RR Clade
LD Clade
Genus Borrelia
Genus Borreliella
(B)
Borrelia crocidurae
Borrelia duttonii
Borrelia recurrentis
Borrelia hispanica
Borrelia persica
Borrelia hermsii
Borrelia coriaceae
Borrelia turicatae
Borrelia parkeri
Borrelia anserina
Borrelia miyamotoi
Borrelia turcica
Candidatus Borrelia tachyglossi
Borreliella burgdorferi
Borreliella finlandensis
"Bor. bissettii"
"Bor. mayonii"
Borreliella chilensis
Borreliella valaisiana
Borreliella spielmanii
Borreliella afzelii
Borreliella japonica
Borreliella garinii
Borreliella bavariensis
RF Clade
RR Clade
LD Clade
Genus Borrelia
Genus Borreliella
(C)
AB007450.1. Borreliella garinii
X85199.1.Borreliella bavariensis PBi
DQ393304.Borreliella californiensis
AB091819.1.Borreliella lusitaniae
CP002746."Bor. bissettii"
ABCW02000003.Borreliella burgdorferi Bol26
ABCW02000003.Borrelia burgdorferi Bol26
ABJZ02000005.Borreliella finlandensis
AB022101.1.Borreliella sinica
AB022101.1.Borreliella sinica
EF570070.1.Borreliella americana
L46688.1.Borreliella andersonii
KP972466.1."Bor. mayonii"
AB529508.1."Bor. yangtzensis"
AB022140.1.Borreliella valaisiana
D67020.1.Borreliella tanukii
AB091818.1.Borreliella afzelii
D67022.1.Borreliella turdi
LF133433.1Borreliella spielmanii
AJ224131.1 "Bor. lanei" CA28
CP009910.Borreliella chilensis
AB111849.1.Borrelia turcica
AF228029.1.Borrelia hermsii
CP025785.1 Candidatus Borrelia tachyglossi
AY166715.1.Borrelia lonestari
DQ872186.Borrelia theileri
AB900817.1.Borrelia miyamotoi
AF210134.1.Borrelia coriaceae
AF210137.1.Borrelia parkeri
AY934605.1.Borrelia turicatae
AB113315.1.Borrelia duttonii
CP005829.Borrelia anserina
AYOT01000223.Borrelia persica No12
AYOU01000163.Borrelia hispanica CRI
AF107356.1.Borrelia recurrentis
AY604977.1.Borrelia crocidurae
JF681792.1.Borrelia microti
JF681793.1.Borrelia latyschewii
0.005
Genus Borreliella
Genus Borrelia

Fig 3. Phylogenetic trees showing the branching of *Borreliaceae* species. (A) A maximum-likelihood (ML) tree based on concatenated sequences of 703 core proteins found in the genomes of *Borreliaceae* species; (B) A tree based on sequence alignment for the RNA polymerase β'- subunit (RpoC protein). (C) A ML tree for *Borreliaceae* species based on 16S rRNA gene sequences.

https://doi.org/10.1371/journal.pone.0221397.g003

[6,7,32,41]. Of these 15 CSIs, based on the available information, 7 CSIs were indicated to be specific for the LD group, whereas in the remaining 8, the genetic changes leading to the CSIs occurred in the lineage leading to the RF group of species. Upon inclusion of sequence information for *B. turcica* and Candidatus Borrelia tachyglossi, which form deeper branching lineages of the RF group, no changes were observed in the specificities of any of the CSIs specific for the LD group and the homologs of the two new species lacked these CSIs (see Table 1).

However, the CSIs which were previously indicated to be specific for the RF clade showed two patterns. Of these, 4 CSIs are commonly shared by all members of the RF group as well as *B. turcica* and Candidatus Borrelia tachyglossi (RR group), whereas the remaining 4 CSIs were only found in the RF group of species and not found in the two deeper branching RR group of species. In Fig 5, an example of CSIs showing the two types of patterns are presented. Information regarding the species specificities of all other CSIs for this group is presented in Table 2. The species distribution pattern of the CSIs for this group is exactly as we had predicted previously and the observed results, independent of the phylogenetic trees, strongly support the following inferences: (i) RR group of species, i.e. *B. turcica* and Candidatus Borrelia tachyglossi, are specifically associated with the RF group (i.e. genus *Borrelia*) as indicated by the 4 CSIs they uniquely share with the other RF group of species (Table 2A; Figs 4 and 5); (ii) *B. turcica* and Candidatus Borrelia tachyglossi are earlier branching members of the genus *Borrelia* and the genetic changes in the 4 CSIs that are absent in these two species have occurred in a common ancestor of the other *Borrelia* species, after the divergence of these two species (Table 2B; Figs 4 and 5). Thus, the species distribution patterns of the CSIs, upon inclusion of sequence information for *B. turcica* and Candidatus Borrelia tachyglossi, rather than showing any lack of specificity, provide important information clarifying and strongly supporting the observed evolutionary relationship of these species to the other *Borreliaceae* species (Fig 4). The CSIs whose specificities are questioned by Margos *et al.* [10] are marked by an asterisk (*) in Fig 4.

The remaining 38 CSIs, which constitute the second category, are present in proteins that are found only in different *Borreliaceae* species [6]. Although these CSIs differentiate members of the LD and the RF group of species, due to the absence of these proteins in outgroup species, it is difficult to determine whether the genetic changes giving rise of these CSIs represent insertion(s) in the LD (RF) group, or deletion(s) in the RF (LD) group (see Fig 4C). Thus, Margos *et al.* [10] are misinterpreting the results for these CSIs, when they indicate that a specific CSI of this kind is an insert or a deletion in the RF or the LD group of species. Nonetheless, with the inclusion of sequences for *B. turcica* and Candidatus Borrelia tachyglossi, which are deeper branching species associated with the RF group, depending upon where the genetic changes responsible for these CSIs have occurred, the species distribution pattern of some of these CSIs will be altered. The presence and absence of the indels in all 38 CSIs from this category and their correct interpretation is provided in Table 3.

If the genetic change leading to the CSI occurred in a common ancestor of either the LD group or the entire RF group (inclusive of the RR group) then the CSIs will be present in one of these groups and absent in the other, similar to that reported in the earlier work. Of the 38 CSIs in this category, 29 showed this pattern and they differentiate between the members of the two *Borreliaceae* genera. One example of a CSI of this kind is shown in Fig 6A. However, if the genetic change in a given gene/protein occurred in a common ancestor of the RF group after the divergence of the RR group of species (viz. *B. turcica* and Candidatus Borrelia

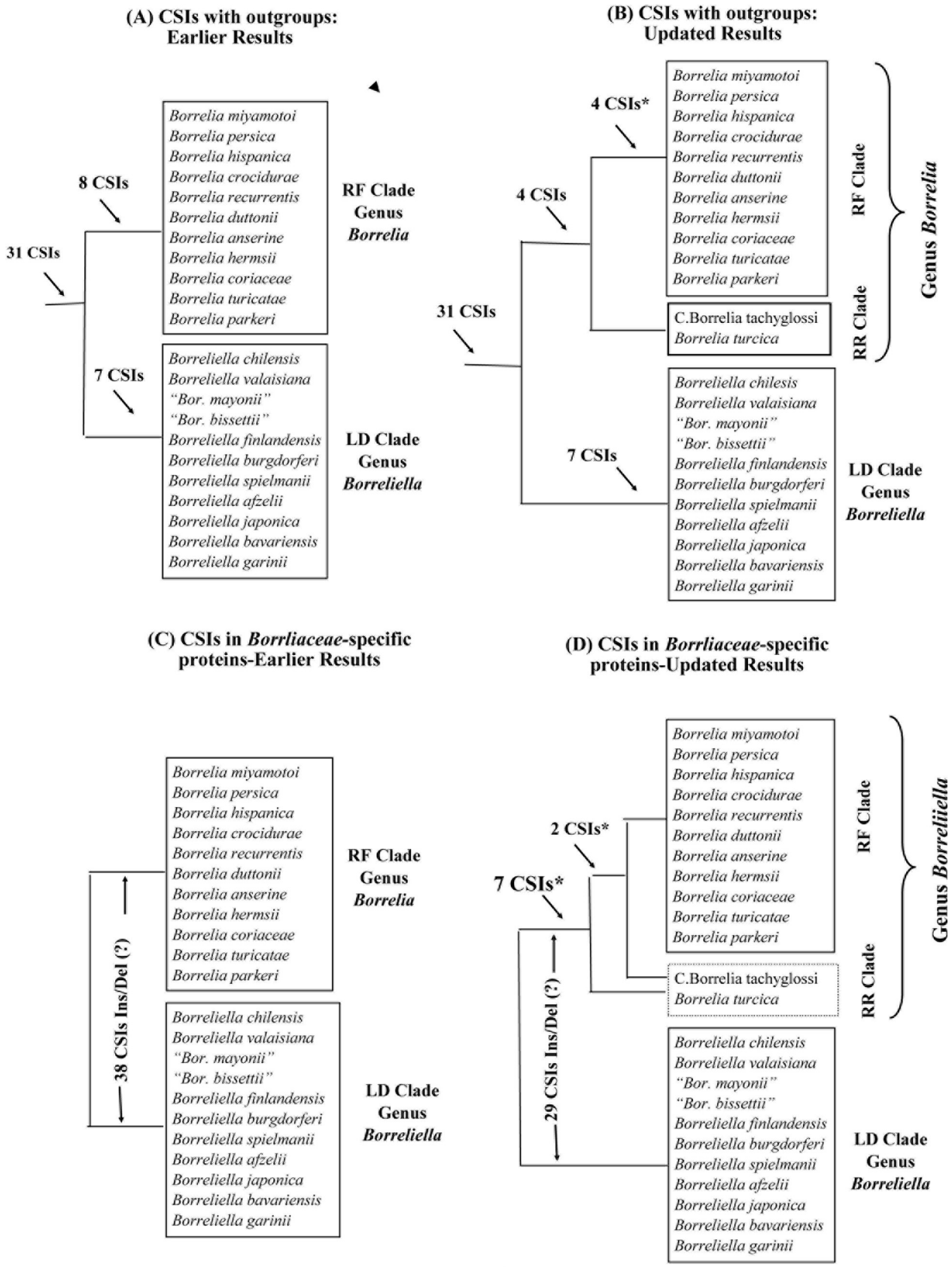

(A) CSIs with outgroups: Earlier Results
8 CSIs
31 CSIs
7 CSIs
Borrelia miyamotoi
Borrelia persica
Borrelia hispanica
Borrelia crocidurae
Borrelia recurrentis
Borrelia duttonii
Borrelia anserine
Borrelia hermsii
Borrelia coriaceae
Borrelia turicatae
Borrelia parkeri
RF Clade Genus Borrelia
Borreliella chilensis
Borreliella valaisiana
"Bor. mayonii"
"Bor. bissettii"
Borreliella finlandensis
Borreliella burgdorferi
Borreliella spielmanii
Borreliella afzelii
Borreliella japonica
Borreliella bavariensis
Borreliella garinii
LD Clade Genus Borreliella
(B) CSIs with outgroups: Updated Results
4 CSIs*
4 CSIs
31 CSIs
7 CSIs
Borrelia miyamotoi
Borrelia persica
Borrelia hispanica
Borrelia crocidurae
Borrelia recurrentis
Borrelia duttonii
Borrelia anserine
Borrelia hermsii
Borrelia coriaceae
Borrelia turicatae
Borrelia parkeri
C.Borrelia tachyglossi
Borrelia turcica
RF Clade
RR Clade
Genus Borrelia
Borreliella chilesis
Borreliella valaisiana
"Bor. mayonii"
"Bor. bissettii"
Borreliella finlandensis
Borreliella burgdorferi
Borreliella spielmanii
Borreliella afzelii
Borreliella japonica
Borreliella bavariensis
Borreliella garinii
LD Clade Genus Borreliella
(C) CSIs in Borrliaceae-specific proteins-Earlier Results
38 CSIs Ins/Del (?)
Borrelia miyamotoi
Borrelia persica
Borrelia hispanica
Borrelia crocidurae
Borrelia recurrentis
Borrelia duttonii
Borrelia anserine
Borrelia hermsii
Borrelia coriaceae
Borrelia turicatae
Borrelia parkeri
RF Clade Genus Borrelia
Borreliella chilensis
Borreliella valaisiana
"Bor. mayonii"
"Bor. bissettii"
Borreliella finlandensis
Borreliella burgdorferi
Borreliella spielmanii
Borreliella afzelii
Borreliella japonica
Borreliella bavariensis
Borreliella garinii
LD Clade Genus Borreliella
(D) CSIs in Borrliaceae-specific proteins-Updated Results
2 CSIs*
7 CSIs*
29 CSIs Ins/Del (?)
Borrelia miyamotoi
Borrelia persica
Borrelia hispanica
Borrelia crocidurae
Borrelia recurrentis
Borrelia duttonii
Borrelia anserine
Borrelia hermsii
Borrelia coriaceae
Borrelia turicatae
Borrelia parkeri
C.Borrelia tachyglossi
Borrelia turcica
RF Clade
RR Clade
Genus Borreliella
Borreliella chilensis
Borreliella valaisiana
"Bor. mayonii"
"Bor. bissettii"
Borreliella finlandensis
Borreliella burgdorferi
Borreliella spielmanii
Borreliella afzelii
Borreliella japonica
Borreliella bavariensis
Borreliella garinii
LD Clade Genus Borreliella

Fig 4. A summary diagram showing the species specificities of different CSIs reported in our earlier work [6]. The CSIs described in our earlier work were of two kinds. Panels (A) and (B) present the results for CSIs, where sequence information for outgroup species was available, whereas panels (C) and (D) show results for CSIs which are found in proteins that are limited to the *Borreliaceae* species (i.e. no homologs in any outgroup species). Panels (A) and (C) show the results as reported earlier [6], whereas panels (B) and (D) show how the observed specificities of the CSIs have been affected upon inclusion of sequences for *B. turcica* and Candidatus Borrelia tachyglossi. The asterisks (*) marks the CSIs whose specificities have been questioned by Margos *et al.* [10]. As shown here and as discussed in the text, these CSIs remain specific for the RF group (genus *Borrelia*) in addition to providing important information regarding the branching or phylogenetic placement of *B. turcica* and Candidatus Borrelia tachyglossi within the genus *Borrelia* and family *Borreliaceae*.

https://doi.org/10.1371/journal.pone.0221397.g004

tachyglossi), then such a CSI will be present in the RF group of species, but absent in *B. turcica* and Candidatus Borrelia tachyglossi as well as the LD group of species. There were 7 CSIs, which showed this type of pattern (listed at the bottom of Table 3). One example of a CSI showing this type of pattern is shown in Fig 6B. However, as indicated in Fig 4, the genetic changes in this CSI or other CSIs of this kind should not be interpreted as showing that *B. turcica* and Candidatus Borrelia tachyglossi are specifically related to the LD group of species, as these CSIs, due to the occurrence of genetic changes in a common ancestor of the RF group, are only distinguishing the RF group of species from other *Borreliaceae* species. Further, as noted earlier and shown in Fig 3, although in phylogenetic trees based on large datasets of proteins, *B. turcica* and Candidatus Borrelia tachyglossi form a deeper-branching clade, the grouping together of these two species/strain is not seen in trees based on several individual protein sequences and also in 16S rRNA trees (Fig 3B and 3C and unpublished results). Due to this, in some cases the genetic change leading to the CSI can also occur in an RF-group ancestor inclusive of *B. turcica* (or Candidatus Borrelia tachyglossi) but after the branching of Candidatus Borrelia tachyglossi (or *B. turcica*). The genetic changes in two of the CSIs in *Borreliaceae*-specific proteins (viz. a membrane protein and DNA polymerase III subunit delta) described in our earlier work [6] appeared to have occurred at these stages of evolution. Sequence information for one of these CSIs is presented in Fig 6C. In this case, the described CSI is present in LD clade of species and Cand. Borrelia tachyglossi whereas *B. turcica* and the RF group of species are lacking this CSI. However, in this case, it will again be incorrect to interpret that the presence of this CSI in Cand. Borrelia tachyglossi and the LD group of species indicates that this species is specifically related to the LD group of species. A summary of the distribution pattern of different CSIs in the second category before and after the inclusion of results from *B. turcica* and Cand. Borrelia tachyglossi is presented in Fig 4C and 4D. The CSIs whose specificities are questioned by Margos *et al.* [10] are marked by asterisk (*) in Fig 4. Based on the correct interpretation of the genetic and evolutionary significance of these

Table 1. Conserved signature indels (CSIs) found in widely distributed proteins that are specific for the members of the Lyme disease Borrelia (i.e. genus *Borreliella*).

Protein Name	GI Number	*B.burgdorferi* B31 locus	Indel Size	Indel Position
Recombinase A	492960118	BB_0131	1 aa ins	228–272
Trigger factor Tig	386854012	BB_0347	2 aa ins	106–142
Chemotaxis protein CheY	15594760	BB_0415	1 aa del	197–231
DNA polymerase III subunit beta	410679212	BB_0438	1 aa del	135–176
Translation factor Sua5	15595079	BB_0610	2 aa ins	149–182
Ferrous iron transporter A	51598605	BB_0730	1 aa del	88–126
Glucose-6-phosphate isomerase	493478887	BB_0734	1 aa ins	81–134

These 7 CSIs described in our earlier work [6] are uniquely shared by different genome-sequenced *Borreliella* (or LD group) species. Updating of sequence information for these CSIs show that they are not found in any species from the genus *Borrelia* (RF group) including *Borrelia turcica* and Candidatus Borrelia tachyglossi. Sequence alignments for these CSIs have been presented in earlier work [6].

https://doi.org/10.1371/journal.pone.0221397.t001

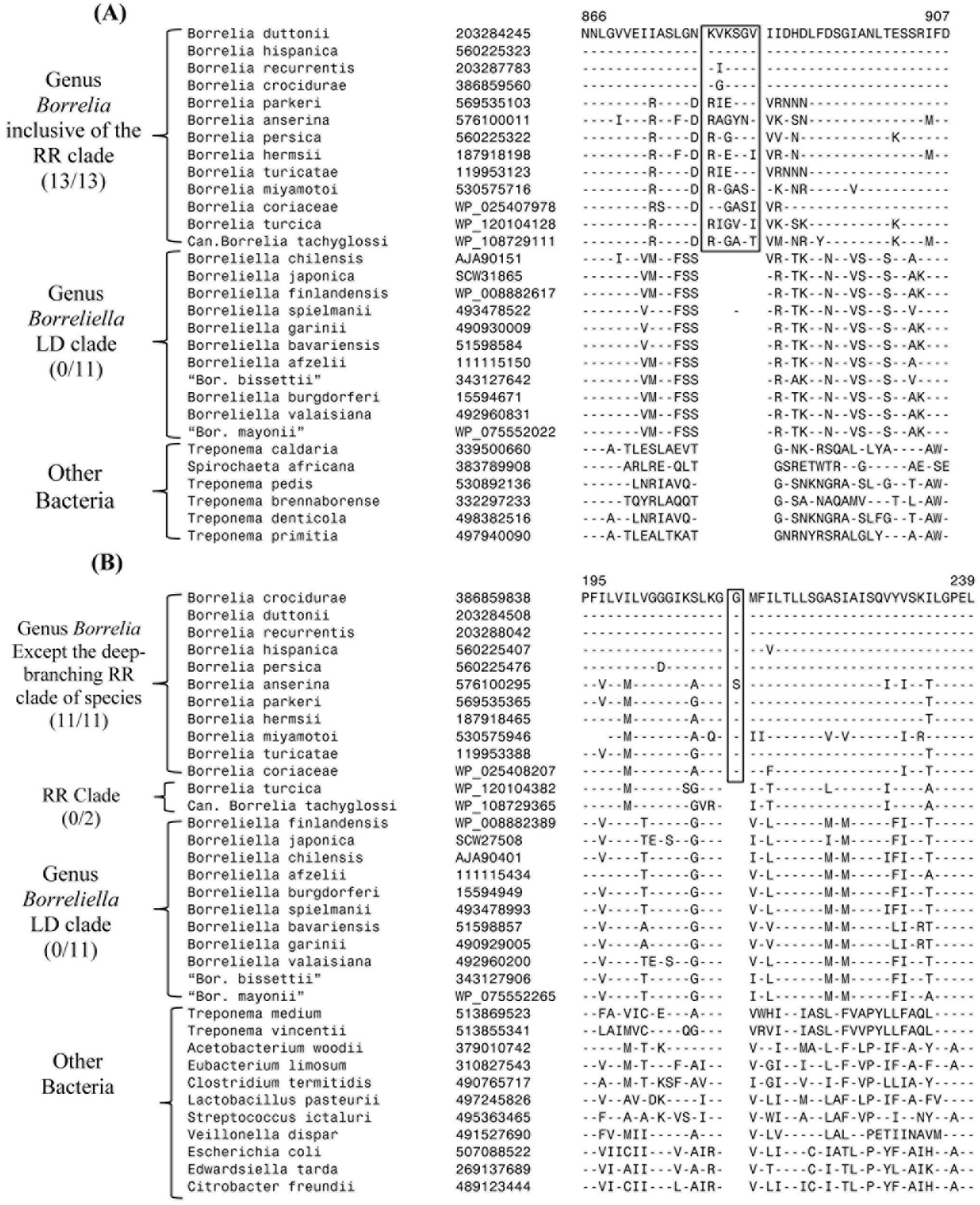

(A)
866
907
Genus Borrelia inclusive of the RR clade (13/13)
Genus Borreliella LD clade (0/11)
Other Bacteria
Borrelia duttonii 203284245 NNLGVVEIIASLGN KVKSGV IIDHDLFDSGIANLTESSRIFD
Borrelia hispanica 560225323
Borrelia recurrentis 203287783
Borrelia crocidurae 386859560
Borrelia parkeri 569535103
Borrelia anserina 576100011
Borrelia persica 560225322
Borrelia hermsii 187918198
Borrelia turicatae 119953123
Borrelia miyamotoi 530575716
Borrelia coriaceae WP_025407978
Borrelia turcica WP_120104128
Can.Borrelia tachyglossi WP_108729111
Borreliella chilensis AJA90151
Borreliella japonica SCW31865
Borreliella finlandensis WP_008882617
Borreliella spielmanii 493478522
Borreliella garinii 490930009
Borreliella bavariensis 51598584
Borreliella afzelii 111115150
"Bor. bissettii" 343127642
Borreliella burgdorferi 15594671
Borreliella valaisiana 492960831
"Bor. mayonii" WP_075552022
Treponema caldaria 339500660
Spirochaeta africana 383789908
Treponema pedis 530892136
Treponema brennaborense 332297233
Treponema denticola 498382516
Treponema primitia 497940090
(B)
195
239
Genus Borrelia Except the deep-branching RR clade of species (11/11)
RR Clade (0/2)
Genus Borreliella LD clade (0/11)
Other Bacteria
Borrelia crocidurae 386859838 PFILVILVGGGIKSLKG G MFILTLLSGASIAISQVYVSKILGPEL
Borrelia duttonii 203284508
Borrelia recurrentis 203288042
Borrelia hispanica 560225407
Borrelia persica 560225476
Borrelia anserina 576100295
Borrelia parkeri 569535365
Borrelia hermsii 187918465
Borrelia miyamotoi 530575946
Borrelia turicatae 119953388
Borrelia coriaceae WP_025408207
Borrelia turcica WP_120104382
Can. Borrelia tachyglossi WP_108729365
Borreliella finlandensis WP_008882389
Borreliella japonica SCW27508
Borreliella chilensis AJA90401
Borreliella afzelii 111115434
Borreliella burgdorferi 15594949
Borreliella spielmanii 493478993
Borreliella bavariensis 51598857
Borreliella garinii 490929005
Borreliella valaisiana 492960200
"Bor. bissettii" 343127906
"Bor. mayonii" WP_075552265
Treponema medium 513869523
Treponema vincentii 513855341
Acetobacterium woodii 379010742
Eubacterium limosum 310827543
Clostridium termitidis 490765717
Lactobacillus pasteurii 497245826
Streptococcus ictaluri 495363465
Veillonella dispar 491527690
Escherichia coli 507088522
Edwardsiella tarda 269137689
Citrobacter freundii 489123444

Fig 5. Partial sequence alignments of two CSIs in proteins with outgroup species that were previously reported as specific for the RF clade. Panel (A) shows a 6 aa insert in a hypothetical protein BDU327 (BB_0326) that is specifically found in all members of the genus *Borrelia* including *B. turcica* and Candidatus Borrelia tachyglossi. (B) This panel shows a 1 aa insert in the L-lactate permease protein, which is only shared by all RF clade species but is absent in the *B. turcica* and Candidatus Borrelia tachyglossi homologs, which are deeper branching members of the genus *Borrelia* (see Figs 3 and 4). Dashes (-) in all alignments shows sequence identity with the amino acids on the top line.

https://doi.org/10.1371/journal.pone.0221397.g005

CSIs, as shown in Fig 4, it is clear that these CSIs are also highly specific characteristics of most members of the genus *Borrelia*. In addition, they are also clarifying the phylogenetic placement of the species *B. turcica* and Cand. Borrelia tachyglossi within this genus and the family *Borreliaceae*.

Based on the evidence presented above, specifically, the correct interpretations of the results for the specificities of the CSIs and the inadequacy of genomic similarity (POCP threshold) as a criterion for genus level differentiation [12], it should be clear that the concerns raised by Margos *et al.* [10] to challenge the division of the genus *Borrelia* into two genera are not justified. In a recent publication, Estrada-Peña and Cabezas-Cruz [49] based on their examination of presence or absence of different biological processes in spirochetes species have inferred that members of the genus *Borrelia* and *Borreliella* are more similar to each other than other free-living (viz. *Sediminispirochaeta*, *Spirochaeta* and *Sphaerochaeta*) or pathogenic spirochetes such as *Leptospira*, *Treponema* and *Brachyspira*. However, their results are not surprising, as both *Borrelia* and *Borreliella* are part of the family *Borreliaceae* whose members exhibit very similar life cycle and vector(s)-host transmission characteristics [2,5,7,40]. With the exception of *B. recurrentis*, all other *Borrelieaceae* species have a tick-stage in their life cycle [2,5]. Thus, members of the genera *Borrelia* and *Borreliella* have coevolved intracellularly within their natural animal host-reservoir organisms for a long period of time. Due to this it is expected that all members of the family *Borreliaceae* (i.e. *Borrelia* and *Borreliella* genera) will share large number of biological processes and characteristics in common [5,47]. In our own work [6,7,41], we have described 31 CSIs and 82 CSPs which are uniquely shared by the members of these two genera. However, these shared characteristics are properties of the family and they reflect the multiple biological and phenotypic characteristics that the members of this family share in common. These shared properties and biological processes of the family

Table 2. Conserved signature indels in proteins that are specific for either all members of the Genus *Borrelia* or those lacking in the deeper branching *Borrelia turcica* and Candidatus Borrelia tachyglossi.

Protein Name	GI Number	Indel Size	*B.burgdorferi* B31 locus	Indel Position
(A) CSIs specific for all members of the Genus *Borrelia*				
Hypothetical protein BRE16	203287484	3 aa ins	BB_0011	64–98
Hypothetical protein BDU327	203284245	6 aa ins	BB_0326	866–907
1-phosphofructokinase	203288064	1 aa del	BB_0630	101–139
GTP-binding protein	203288075	2 aa ins	BB_0643	42–87
(B) CSIs specific for the Genus *Borrelia* except deeper branching *B. turcica* and Candidatus Borrelia tachyglossi				
Nicotinamide-nucleotide adenylyltransferase	187918635	1 aa del	BB_0782	31–61
Hypothetical protein BT0471[1]	119953261	1 aa del	BB_0471	216–261
L-lactate permease	386859838	1 aa ins	BB_0604	195–239
Sodium/panthothenate symporter	119953591	1 aa ins	BB_0814	421–454

These 8 CSIs were described as specific for the RF clade of *Borrelia* species in our earlier work [6]. Upon inclusion of sequence information for including *Borrelia turcica* and Candidatus Borrelia tachyglossi, 4 of these CSIs are also uniquely shared by these two species, whereas the other four CSIs listed below are absent in these two deeper branching species. These CSIs provide evidence that both *Borrelia turcica* and Candidatus Borrelia tachyglossi are members of the genus *Borrelia* and constitute deeper branching lineages of this genus (see summary Fig 4). Sequence alignment of one CSIs' of each kind is presented in Fig 5.

https://doi.org/10.1371/journal.pone.0221397.t002

Table 3. Conserved signature indels in *Borreliaceae*-specific proteins distinguishing *Borrelia* and *Borreliella* and showing deeper branching of the RR group of species within the genus *Borrelia*.

			Borrelia		*Borreliella*	
Protein Name	***B.B* 31 Locus**	**Indel Size**	**RF Clade**	**RR Species**	**LD Clade**	**Interpretation**
Hypothetical protein	BB_0028	2 aa	-	-	+	Distinguishes two genera
Hypothetical protein	BB_0028	1 aa	-	-	+	Distinguishes two genera
Hypothetical protein BRE47	BB_0044	5 aa	-	-	+	Distinguishes two genera
L-proline transport system ATP-binding protein	BB_0146	1 aa	-	-	+	Distinguishes two genera
Penicillin-binding protein	BB_0136	1 aa	-	-	+	Distinguishes two genera
Hypothetical protein Q7M131	BB_0125	1 aa	+	+	-	Distinguishes two genera
Hypothetical protein BT0110	BB_0110	2 aa	-	-	+	Distinguishes two genera
Hypothetical protein BT0110	BB_0110	2 aa	-	N/A	+	Distinguishes two genera
Glutamate racemase	BB_0100	6 aa	-	-	+	Distinguishes two genera
RNA methyltransferase RsmE	BB_0062	1 aa	+	+	-	Distinguishes two genera
DNA mismatch repair protein mutL	BB_0211	4 aa	+	+	-	Distinguishes two genera
Hypothetical protein BRE314	BB_0227	1 aa	+	+	-	Distinguishes two genera
Methylgalactoside ABC transporter ATP-binding protein	BB_0318	1 aa	-	-	+	Distinguishes two genera
Sensory transduction histidine kinase	BB_0420	1 aa	-	-	+	Distinguishes two genera
Hypothetical protein Q7M860	BB_0455	2 aa	+	+	-	Distinguishes two genera
Hypothetical protein KK90081	BB_0083	1 aa	-	-	+	Distinguishes two genera
Outer membrane protein	BB_0167	1 aa	-	-	+	Distinguishes two genera
Transglycosylase SLT domain-containing protein	BB_0259	1 aa	-	-	+	Distinguishes two genera
Cell division protein FtsZ	BB_0299	1 aa	-	-	+	Distinguishes two genera
Excinuclease ABC subunit C	BB_0457	1 aa	-	-	+	Distinguishes two genera
Hypothetical protein BG0519	BB_0507	1 aa	-	-	+	Distinguishes two genera
Hypothetical protein BBIDN1270545	BB_0543	4 aa	-	-	+	Distinguishes two genera
Hypothetical protein BBUN400354	BB_0354	3 aa	-	-	+	Distinguishes two genera
Hypothetical protein BBUZS70553	BB_0543	1 aa	-	-	+	Distinguishes two genera
Hypothetical protein BB0554	BB_0554	1 aa	-	-	+	Distinguishes two genera
Hypothetical protein BB0554	BB_0554	2 aa	-	-	+	Distinguishes two genera
Hypothetical protein BBUCA803285	BB_0664	1 aa	-	-	+	Distinguishes two genera
Chemotaxis protein	BB_0681	1 aa	-	-	+	Distinguishes two genera
Hypothetical protein L14403475	BB_0707	1 aa	-	-	+	Distinguishes two genera
Membrane protein	BB_0234	1 aa	+	#	-	Insertion occurred after the branching of *B. turcica*
DNA polymerase III subunit -ta	BB_0455	2 aa	-	#	+	Insertion occurred after the branching of *B. turcica*
Hypothetical protein BB0838	BB_0838	3 aa	-	+	+	Deletion in the RF clade
Putative lipoprotein	BB_0227	3 aa	+	-	-	Insertion in the RF Clade
Hypothetical protein BRE355	BB_0353	1 aa	+	-	-	Insertion in the RF Clade
Hypothetical protein Q7M140	BB_0134	2 aa	+	-	-	Insertion in the RF Clade
Hypothetical protein BG0159	BB_0161	1 aa	-	+	+	Deletion in the RF clade
Methyl-accepting chemotaxis protein	BB_0681	2 aa	+	-	-	Insertion in the RF Clade
Chemotaxis protein	BB_0681	1 aa	-	+	+	Deletion in the RF clade

These CSIs were previously indicated to differentiate members of the genus *Borrelia* and *Borreliella* [6]. With the inclusion of sequence information for *B. turcica* and Candidadus Borrelia tachyglossi, these CSIs still differentiate the members of these two genera; however, some of them also show the deep branching of the RR group of species in comparison to the other *Borrelia* species. Abbreviations: RR–refer to the repitles- and echidna- related species *B. turcica* and Candidadus Borrelia tachyglossi RF–Relapsing Fever Clade; LD–Lyme Disease Clade; + = presence of insert;— = absence of insert

Candidadus Borrelia tachyglossi contains the insert but it is absent in *B. turcica*.

https://doi.org/10.1371/journal.pone.0221397.t003

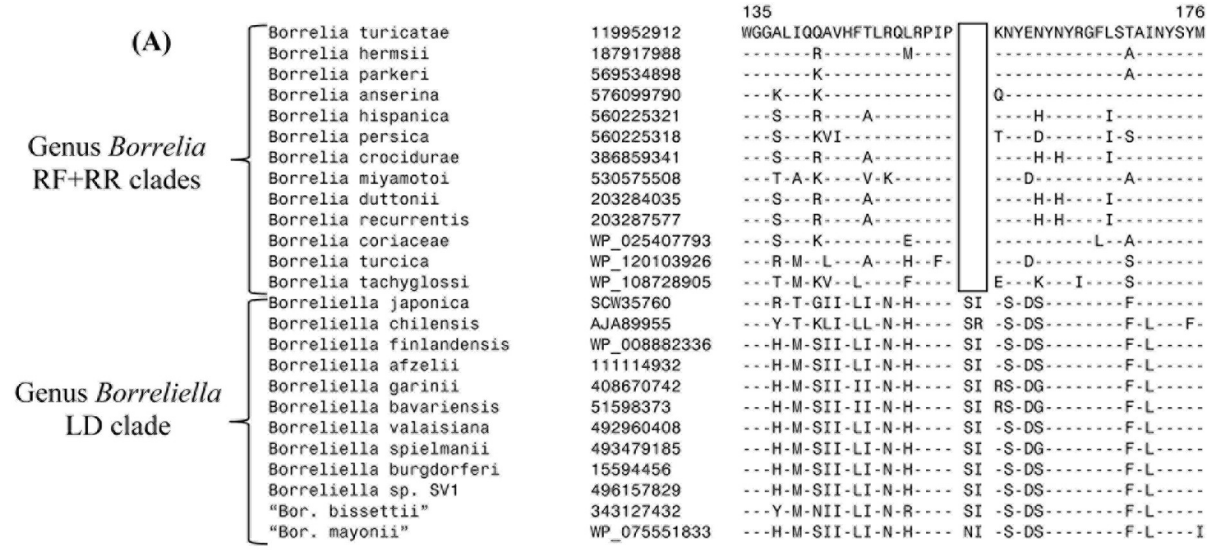

(A)
Genus Borrelia RF+RR clades
Genus Borreliella LD clade
135
176
Borrelia turicatae 119952912 WGGALIQQAVHFTLRQLRPIP KNYENYNYRGFLSTAINYSYM
Borrelia hermsii 187917988 -------R--------M---- -------------A-------
Borrelia parkeri 569534898 -------K------------- -------------A-------
Borrelia anserina 576099790 ---K---K------------- Q-------------------
Borrelia hispanica 560225321 ---S---R----A-------- ----H------I---------
Borrelia persica 560225318 ---S---KVI---------- T---D-------I-S-------
Borrelia crocidurae 386859341 ---S---R----A-------- ----H-H----I---------
Borrelia miyamotoi 530575508 ---T-A-K----V-K------ ---D---------A-------
Borrelia duttonii 203284035 ---S---R----A-------- ----H-H----I---------
Borrelia recurrentis 203287577 ---S---R----A-------- ----H-H----I---------
Borrelia coriaceae WP_025407793 ---S---K--------E---- ----------L--A-------
Borrelia turcica WP_120103926 ---R-M--L---A---H--F- ---D---------S-------
Borrelia tachyglossi WP_108728905 ---T-M-KV--L----F---- E---K---I----S-------
Borreliella japonica SCW35760 ---R-T-GII-LI-N-H---- SI -S-DS--------F-------
Borreliella chilensis AJA89955 ---Y-T-KLI-LL-N-H---- SR -S-DS--------F-L---F-
Borreliella finlandensis WP_008882336 ---H-M-SII-LI-N-H---- SI -S-DS--------F-L-----
Borreliella afzelii 111114932 ---H-M-SII-LI-N-H---- SI -S-DS--------F-L-----
Borreliella garinii 408670742 ---H-M-SII-II-N-H---- SI RS-DG--------F-L-----
Borreliella bavariensis 51598373 ---H-M-SII-II-N-H---- SI RS-DG--------F-L-----
Borreliella valaisiana 492960408 ---H-M-SII-LI-N-H---- SI -S-DS--------F-L-----
Borreliella spielmanii 493479185 ---H-M-SII-LI-N-H---- SI -S-DG--------F-L-----
Borreliella burgdorferi 15594456 ---H-M-SII-LI-N-H---- SI -S-DS--------F-L-----
Borreliella sp. SV1 496157829 ---H-M-SII-LI-N-H---- SI -S-DS--------F-L-----
"Bor. bissettii" 343127432 ---Y-M-NII-LI-N-R---- SI -S-DS--------F-L-----
"Bor. mayonii" WP_075551833 ---H-M-SII-LI-N-H---- NI -S-DS--------F-L----I
(B)
Genus Borrelia RF clade
Genus Borrelia RR clade
Genus Borreliella LD clade
172
204
Borrelia recurrentis 203287684 EAFPYFTHKIKEKI TTT YQNKLLFDLQQAVIYF
Borrelia crocidurae 386859452 ---------- --- ---E------
Borrelia duttonii 203284146 ---------- --- ---E------
Borrelia parkeri 576099167 -------QR----- EQ- -S------
Borrelia anserina 576099907 K------Q------ EK- QS------
Borrelia hispanica 560225323 ---------- --- ---E------I---
Borrelia persica 560225322 ----------V -K- ---E------
Borrelia miyamotoi 530575614 -S-----Q-----V SK- QLSE-------S----
Borrelia hermsii 187918099 -------QR---Q- EK- -P------
Borrelia turicatae 119953023 -------QR----- EQ- -P------
Borrelia coriaceae WP_025407890 -------QR---E- DKM -P-N--------V--
Borrelia turcica WP_120104034 -------Q-----A QS-D------
Borrelia tachyglossi WP_108729017 -------QR----T QP------
Borrelia chilensis AJA90057 ----F-IK------ LP-E------
Borreliella japonica SCW33224 ----F--K------ LP-E--------A---
Borreliella finlandensis WP_008882546 ----F-IK------ LP-E------
Borreliella spielmanii 493478704 ----F--K------ SP-E--------I---
Borreliella garinii 408670854 ----F-IK------ LP-E------
Borreliella bavariensis 51598488 ----F-IK------ LP-E------
Borreliella valaisiana 492960292 ----F-IK------ LP-E------
Borreliella afzelii 111115051 ----F-IK------ LP-E------
Borreliella burgdorferi 15594572 ----F-IK------ LP-E------
Borreliella sp. SV1 496158039 ----F-IK------ LP-E------
"Bor. bissettii" 343127545 ----F-IK------ LP-E------
"Bor. mayonii" WP_075551937 ----F-IK------ LP-E------
(C)
Genus Borreliella LD clade + C. Borrelia tachyglossi
Genus Borreliella RF clade + B. turcica
11
43
Borreliella burgdorferi 15594800 EQGLKEAYLKELLIKMDA FK SEVSVTKIFLSELS
Borreliella sp. SV1 496157963 ---------- -- -------V------
"Bor. bissettii" 343127763 ----------N- -- -K------
Borreliella valaisiana 492960772 ---------- -- ---------
Borreliella afzelii 111115283 ---------- -- -K----------I-
Borreliella garinii 386853861 ---------- -- -K-L--------
Borreliella bavariensis 51598711 ----------V---- -- -K-L--------
Borreliella spielmanii 493478951 ---------- -- -K-L--------
Borreliella finlandensis EEH00502 ---------- -- -------V------
Borreliella japonica SCW29049 ----------L-- -- ---------
Borreliella chilensis AJA90267 ---------- -- -G-L---L------
"Bor. mayonii" WP_075552134 ---------- -- ---L--------
Borrelia tachyglossi WP_108729228 ------D--NNI-G-LN- LH NN-M-R-L------
Borrelia turcica WP_120104249 ---------NDI-S-LGD LL-N-----D--
Borrelia coriaceae WP_028328331 ------S--NDI-S-LNV DGIC---F-M----
Borrelia turicatae 119953244 ---------NNIFS-LN- DNIC---L-M----
Borrelia parkeri 569535221 ---------NDIFS-LN- DDICI--L-M----
Borrelia anserina 576100150 ------T--NTIFSNLNT DDICI--L------
Borrelia hispanica 560225323 ---------NDIFR-LN- DDIHI--L-M----
Borrelia persica 560225476 ---------NDIFS-FHI DDLH---L-M----
Borrelia crocidurae 386859686 ---------NDIFS-LN- DDIHI--L-M----
Borrelia duttonii 203284365 ---------NDIFS-LN- DDIHI--L-M----
Borrelia hermsii 187918321 ---------NDIFN-LN- DDICI--L-M----
Borrelia recurrentis 203287899 ---------NDIFS-LN- DDIHI--L-M----
Borrelia miyamotoi 530575825 ---------NDIYS-LGT DD-C---L-M----

Fig 6. Partial sequence alignments of three CSIs in proteins found only in the *Borreliaceae* species providing differentiation among members of the genera *Borrelia* and *Borreliella*. (A) This panel shows a 2 aa CSI in a hypothetical protein BT0110 that differentiates the members of the genera *Borrelia* and *Borreliella*. Twenty nine other CSIs also show a similar species distribution (Table 3). Due to the absence of outgroup species it is difficult to infer whether this CSI is an insert in the genus *Borrelia* or a deletion in the genus *Borreliella*. (B) A 3 aa CSI in a putative lipoprotein that is specific for the RF clade of species. Due to the absence of this CSI in the LD clade as well as in *B. turcica* and Candidatus Borrelia tachyglossi homologs this CSI is an insert in the RF clade of species (see Fig 4). (C) A 2 aa CSI in DNA polymerase III subunit delta, which is commonly shared by the LD clade of species and Cand. Borrelia tachyglossi, but absent in *B. turcica* and the RF group of species. Based on its species distribution, this CSI is inferred to be an insert in a common ancestor of the RF clade and *B. turcica* (see Fig 4 for additional information).

https://doi.org/10.1371/journal.pone.0221397.g006

Borreliaceae, which have been better studied, have likely led to the inference by Estrada-Peña and Cabezas-Cruz [49] that the members of these two genera are more closely related to each other than other spirochetes groups/genera. However, the observed similarity between these two genera, which are the shared properties of the family *Borreliaceae*, does not in any way minimizes or reduces the significance of large numbers of molecular, phenotypic and clinical differences that exist between the members of these two genera that are summarized in this work and which forms the basis of dividing this family into two different genera [6,7]. Estrada-Peña and Cabezas-Cruz [49] have not questioned the validity or significance of any these described characteristics and thus their resistance to splitting the family *Borreliaceae* is not justified.

To further clearly illustrate the differences between members of the genera *Borrelia* and *Borreliella*, in Table 4, I present a summary of some of the characteristics which distinguish members of these two genera. A number of other characteristics, which also distinguish these genera, are noted by Barbour [5] in a recent publication on the family *Borreliaceae*. The characteristics which distinguish members of these two genera include their different disease spectrums, multiple important differences in their epidemiology and phenotypic properties [5,7], and the clear differentiation and demarcation of these two groups based on genomic similarity and numerous molecular sequence based characteristics. Based on this evidence, it will be accurate to state that the distinction between these two groups of spirochetes is supported by more numerous and distinct types of characteristics than has been reported/observed for any two closely related groups (genera) of prokaryotes. Hence, we urge critics of this division to keep in mind the strong and incontrovertible evidence supporting the distinctness of these two groups of spirochetes.

As noted earlier, the species which are now part of the genus *Borreliella* are widely referred to by the name "*Borrelia burgdorferi sensu lato*", recognizing their distinctness from other *Borreliaceae* species, which are members of the genus *Borrelia* [1,2,5,9]. However, the meanings of the terms "*Borrelia burgdorferi sensu lato*" or "RF clade", or which *Borrelia* species are part of each of these groups, or the species which fall outside of these two groups (viz. RR group of species), are not clearly understood by many scientists and others professionals working in this as well as other related fields. Hence, the substitution of these poorly understood terms with more precise and unambiguous names (*Borrelia* and *Borreliella*), which clearly differentiates the relapsing fever encompassing group of species from the different Lyme disease-causing and related microorganisms [3,4,46], should be highly beneficial to the field in terms of advancing our understanding of the molecular, biochemical and biological differences that underlie these two unique disease-causing groups of microorganisms.

Subsequent to our earlier work [6], a number of new species belonging to the family *Borreliaceae* have been described [55–58]. Of these species, *Borrelia bissettiae*, *Borrelia californiensis*, *Borrelia lanei*, *Borrelia mayonii* and *Borrelia yangtzensis* group reliably with the members of the genus *Borreliella* (LD-group) in 16S rRNA trees [55] (Fig 3C), or where genome sequence information is available based on uniquely shared molecular characteristics with other

Table 4. Clinical, molecular and phenotypic differences between members of the RF-RR group (genus *Borrelia*) and the Lyme disease group (genus *Borreliella*).

Characteristics	Genus *Borrelia* (RF+RR species)	Genus *Borreliella* (LD group)	References
Clinical Spectrum			
Relapsing fever causing bacteria	Encompasses All	None	See [1,2,5,50,51]
Lyme-disease causing bacteria	None	Encompasses All	See [1,2,5,50,52]
Phylogenetic, Molecular and Genomic Characteristics			
Branching in phylogenetic trees based on 16S rRNA and other genes/proteins sequences	In all phylogenetic trees, members of the RF+RR group and the LD group form strongly-supported clades clearly separated from each other.		[5–7,10,53] present study
Average Nucleotide Identity (ANI) Matrix based on Genomes	Members of the RF+RR group and the LD group are clearly differentiated based on higher similarity seen between the members of each group.		[6]
Average Amino Acid Identity (AAI) Matrix based on Genomes	Members of the RF+RR group and the LD group are clearly differentiated based on higher similarity seen between the members of each group.		[7]
Percentage of Conserved Protein (POCP) Matrix based on Genome Sequences	Members of the RF+RR group and the LD group are clearly differentiated based on higher similarity seen between the members of each group.		Present study
Conserved Signature Indels (CSIs) in widely-distributed proteins	4 CSIs exclusively found in different members of this group. 4 other CSIs are also specific for the RF clade but lacking in deeper branching RR group of species.	7 CSIs exclusively found in the LD group differentiating it from the RF+RR group of species.	[6,7] and present study
Conserved Signature Indels (CSIs) in Borreliaceae-specific proteins	29 CSIs provide clear differentiation between these two groups of species + 9 CSIs specific for the RF clade and show deeper branching of the RR group of species.		Present study
Conserved Signature Proteins	4 CSPs exclusively found in most members of this group.	17 CSPs exclusively (or mainly) found in members of this group.	[6][#]
Phenotypic Characteristics			
Arthropod vectors	Argasid ticks, prostriate and metastriate ixodid ticks and human body louse[*]	Primarily prostriate ticks of the genus *Ixodes*	See [1,2,5,50–52,54]
Density of Spirochetes in blood of infected humans/animals	High	Low	See [5,7,51,52]
Average number of flagella at one end of cells	Mostly in the range of 15–20	Generally in the range of 7–11	See [5,7,51,52]

In addition to the characteristics noted in this Table, some other molecular and phenotypic differences between members of these two genera have been summarized by Barbour [5].

[#] Based on updated sequence information.

[*] only *B. recurrentis* is transmitted via a louse.

https://doi.org/10.1371/journal.pone.0221397.t004

members of the genus *Borreliella* [8,9] (see Table 1). Hence, new name combinations for these species are described below.

Description of *Borreliella bissettiae* comb. nov. (bis.set´ti.ae. N.L. gen. n. *bissettiae*, of Bissett, named after Marjorie L. Bissett, who isolated and described this spirochaete with her co-worker Warren Hill) Basonym: *Borrelia bissettiae* Margos et al. 2016

The strain for this species was isolated by Bissett and Hill [59] and the description of this species is as provided by Margos *et al.* [56] for *Borrelia bissettiae*

Type strain: DN127 = DSM 17990 = CIP 109136.

Description of *Borreliella californiensis* comb. nov. (ca.li.for.ni.en´sis. N.L. fem. adj. *californiensis*, belonging to California, from where the type strain was isolated)

Basonym: *Borrelia californiensis* Margos et al. 2016

The description of this species is the same as provided by Postic *et al.* [60] and Margos *et al.* [56] for *Borrelia californiensis*

Type strain: CA446 = DSM 17989 = ATCC BAA-2689.

Description of *Borreliella lanei* comb. nov. (la.ne′i. N.L. gen. n. *lanei*, in honour of Professor Robert S. Lane for his outstanding contributions to *Borrelia* and *Ixodes* research)

Basonym: *Borrelia lanei* Margos et al. 2017

The description of this species is the same as provided by Margos *et al*. [55] for *Borrelia lanei*

Type strain: (see also StrainInfo.net) CA28-91 = DSM 17992 = CIP 109135.

Description of *Borreliella mayonii* comb. nov. (ma.yo′ni.i. N.L. gen. n. *mayonii*, after William James Mayo and Charles Horace Mayo, founders of the Mayo Clinic).

Basonym: *Borrelia mayonii* Pritt et al. 2016

The description of this species is the same as provided by Pritt *et al*. [58] for *Borrelia mayonii*

Type strain: MN14-1420 = ATCC BAA-2743 = DSM 10281.

Description of *Borreliella yangtzensis* comb. nov. (yang.tzen′sis. N.L. fem. adj. *yangtzensis*, referring to the Yangtze River valley in China, where these organisms were first isolated.

Basonym: *Borrelia yangtzensis* Margos et al. 2015

The description of this species is the same as provided by Margos *et al*. [57] for *Borrelia yangtzensis*

Type strain: Okinawa-CW62 = DSM 24625 = JCM 17189.

Supporting information

S1 Table. Species and genome sequence information for *Enterobacteriaceae* species used in POCP analysis.
(PDF)

S2 Table. Species and genome sequence information for *Morganellaceae* and *Cystobacteraceae* species used in POCP analysis.
(PDF)

S3 Table. Averages of the intrageneric and intergeneric POCP values for species from the family *Morganellaceae* and *Cystobacteraceae*.
(PDF)

Acknowledgments

This work was supported by research grant No. 249924 from the Natural Science and Engineering Research Council of Canada to RSG. I thank Sudip Patel for assistance in this work and Dr. Alan Barbour for helpful comments on the manuscript.

Author Contributions

Conceptualization: Radhey S. Gupta.

Data curation: Radhey S. Gupta.

Formal analysis: Radhey S. Gupta.

Funding acquisition: Radhey S. Gupta.

Investigation: Radhey S. Gupta.

Methodology: Radhey S. Gupta.

Project administration: Radhey S. Gupta.

Resources: Radhey S. Gupta.

Software: Radhey S. Gupta.

Supervision: Radhey S. Gupta.

Validation: Radhey S. Gupta.

Writing – original draft: Radhey S. Gupta.

Writing – review & editing: Radhey S. Gupta.

References

1. Barbour AG, Hayes SF. Biology of *Borrelia* species. Microbiol Rev.1986; 50: 381–400. PMID: 3540570
2. Wang G, Schwartz I: Genus II. *Borrelia* Swellengrebel 1907, 582^{AL}. In Bergey's Manual of Systematic Bacteriology Volume 4: The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes. Edited by Brenner DJ, Krieg NR, Garrity GM, Staley JT. New York: Springer Verlag; 2011:484–498.
3. Talagrand-Reboul E, Boyer PH, Bergstrom S, Vial L, Boulanger N. Relapsing Fevers: Neglected Tick-Borne Diseases. Front Cell Infect Microbiol.2018; 8: 98. https://doi.org/10.3389/fcimb.2018.00098 PMID: 29670860
4. Bergstrom S, Normark J. Microbiological features distinguishing Lyme disease and relapsing fever spirochetes. Wien Klin Wochenschr.2018; 130: 484–490. https://doi.org/10.1007/s00508-018-1368-2 PMID: 30074091
5. Barbour AG: Borreliaceae. In Bergey Manual of Systematics of Bacteria and Archaea. John Wiley & Sons, Inc. Bergey's Manual Trust; 2019.
6. Adeolu M, Gupta RS. A phylogenomic and molecular marker based proposal for the division of the genus *Borrelia* into two genera: The emended genus *Borrelia* containing only the members of the relapsing fever *Borrelia*, and the genus *Borreliella* gen. nov. containing the members of the Lyme disease *Borrelia* (*Borrelia burgdorferi* sensu lato complex). Antonie van Leeuwenhoek.2014; 105: 1049–1072. https://doi.org/10.1007/s10482-014-0164-x PMID: 24744012
7. Barbour AG, Adeolu M, Gupta RS. Division of the genus *Borrelia* into two genera (corresponding to Lyme disease and relapsing fever groups) reflects their genetic and phenotypic distinctiveness and will lead to a better understanding of these two groups of microbes (Margos et al. (2016) There is inadequate evidence to support the division of the genus *Borrelia*. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/ijsem.0.001717). Int J Syst Evol Microbiol.2017.
8. Barbour AG. Phylogeny of a relapsing fever *Borrelia* species transmitted by the hard tick *Ixodes scapularis*. Infect Genet Evol.2014; 27: 551–558. https://doi.org/10.1016/j.meegid.2014.04.022 PMID: 24813576
9. Margos G, Vollmer SA, Ogden NH, Fish D. Population genetics, taxonomy, phylogeny and evolution of *Borrelia burgdorferi* sensu lato. Infect Genet Evol.2011; 11: 1545–1563. https://doi.org/10.1016/j.meegid.2011.07.022 PMID: 21843658
10. Margos G, Gofton A, Wibberg D, Dangel A, Marosevic D, Loh SM, et al. The genus *Borrelia* reloaded. PLoS ONE.2018; 13: e0208432. https://doi.org/10.1371/journal.pone.0208432 PMID: 30586413
11. Gofton AW, Margos G, Fingerle V, Hepner S, Loh SM, Ryan U, et al. Genome-wide analysis of *Borrelia turcica* and 'Candidatus Borrelia tachyglossi' shows relapsing fever-like genomes with unique genomic links to Lyme disease *Borrelia*. Infect Genet Evol.2018; 66: 72–81. https://doi.org/10.1016/j.meegid.2018.09.013 PMID: 30240834
12. Qin QL, Xie BB, Zhang XY, Chen XL, Zhou BC, Zhou J, et al. A proposed genus boundary for the prokaryotes based on genomic insights. J Bacteriol.2014; 196: 2210–2215. https://doi.org/10.1128/JB.01688-14 PMID: 24706738
13. Gupta RS, Lo B, Son J. Phylogenomics and Comparative Genomic Studies Robustly Support Division of the Genus *Mycobacterium* into an Emended Genus *Mycobacterium* and Four Novel Genera. Front Microbiol.2018; 9: 67. https://doi.org/10.3389/fmicb.2018.00067 PMID: 29497402
14. Adeolu M, Alnajar S, Naushad S, Gupta RS. Genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for *Enterobacterales* ord. nov. divided into the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., Morganellaceae fam. nov., and *Budviciaceae* fam. nov. Int J Syst Evol Microbiol.2016; 66: 5575–5599. https://doi.org/10.1099/ijsem.0.001485 PMID: 27620848

15. Alnajar S, Gupta RS. Phylogenomics and comparative genomic studies delineate six main clades within the family *Enterobacteriaceae* and support the reclassification of several polyphyletic members of the family. Infect Genet Evol.2017; 54: 108–127. https://doi.org/10.1016/j.meegid.2017.06.024 PMID: 28658607

16. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics.2012; 28: 3150–3152. https://doi.org/10.1093/bioinformatics/bts565 PMID: 23060610

17. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol.2011; 7: 539. https://doi.org/10.1038/msb.2011.75 PMID: 21988835

18. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics.2009; 25: 1972–1973. https://doi.org/10.1093/bioinformatics/btp348 PMID: 19505945

19. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Res.2014; 42: D643–D648. https://doi.org/10.1093/nar/gkt1209 PMID: 24293649

20. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol.2013; 30: 2725–2729. https://doi.org/10.1093/molbev/mst197 PMID: 24132122

21. Gupta RS: Identification of Conserved Indels that are Useful for Classification and Evolutionary Studies. In Bacterial Taxonomy, Methods in Microbiology *Volume* 41. Edited by Goodfellow M, Sutcliffe IC, Chun J. London: Elsevier; 2014:153–182.

22. Oren A: Microbial Systematics. In Handbook of Environmental Engineering, *Vol.* 10: Environmental Biotechnology. Edited by Wang LK, Ivanov V, Tay J-H, Hung YT. New York: Springer Science + Business Media; 2010:81–120.

23. Oren A, Garrity GM. Then and now: a systematic review of the systematics of prokaryotes in the last 80 years 13. Antonie van Leeuwenhoek.2014; 106: 43–56. https://doi.org/10.1007/s10482-013-0084-1 PMID: 24306768

24. Parker CT, Tindall BJ, Garrity GM. International Code of Nomenclature of Prokaryotes. Int J Syst Evol Microbiol.2015.

25. Sutcliffe IC. Challenging the anthropocentric emphasis on phenotypic testing in prokaryotic species descriptions: rip it up and start again 1. Front Genet.2015; 6: 218. https://doi.org/10.3389/fgene.2015.00218 PMID: 26136772

26. Yarza P, Yilmaz P, Pruesse E, Glockner FO, Ludwig W, Schleifer KH, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences 1. Nat Rev Microbiol.2014; 12: 635–645. https://doi.org/10.1038/nrmicro3330 PMID: 25118885

27. Rossello-Mora R, Amann R. Past and future species definitions for Bacteria and Archaea. Syst Appl Microbiol.2015; 38: 209–216. https://doi.org/10.1016/j.syapm.2015.02.001 PMID: 25747618

28. Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC, et al. Microbial species delineation using whole genome sequences. Nucleic Acids Res.2015; 43: 6761–6771. https://doi.org/10.1093/nar/gkv657 PMID: 26150420

29. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol.2007; 57: 81–91. https://doi.org/10.1099/ijs.0.64483-0 PMID: 17220447

30. Konstantinidis KT, Tiedje JM. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A.2005; 102: 2567–2572. https://doi.org/10.1073/pnas.0409727102 PMID: 15701695

31. Klenk HP, Goker M. En route to a genome-based classification of Archaea and Bacteria? 1. Syst Appl Microbiol.2010; 33: 175–182. https://doi.org/10.1016/j.syapm.2010.03.003 PMID: 20409658

32. Gupta RS. Impact of genomics on the understanding of microbial evolution and classification: the importance of Darwin's views on classification. FEMS Microbiol Rev.2016; 40: 520–553. https://doi.org/10.1093/femsre/fuw011 PMID: 27279642

33. Sutcliffe IC, Trujillo ME, Whitman WB, Goodfellow M. A call to action for the International Committee on Systematics of Prokaryotes. Trends Microbiol.2013; 21: 51–52. https://doi.org/10.1016/j.tim.2012.11.004 PMID: 23238020

34. Sangal V, Goodfellow M, Jones AL, Schwalbe EC, Blom J, Hoskisson PA, et al. Next-generation systematics: An innovative approach to resolve the structure of complex prokaryotic taxa. Sci Rep.2016; 6: 38392. https://doi.org/10.1038/srep38392 PMID: 27924912

35. Oren A: Concepts about Phylogeny of Microorganisms—an Historical Overview. In *Molecular* Phylogeny of Microorganisms. Edited by Oren A, Papke RT. Norfolk: Caister Academic Press; 2010:1–21.

36. Bhandari V, Gupta RS. Molecular signatures for the phylum (class) *Thermotogae* and a proposal for its division into three orders (*Thermotogales*, Kosmotogales ord. nov. and *Petrotogales* ord. nov.) containing four families (*Thermotogaceae*, *Fervidobacteriaceae* fam. nov., Kosmotogaceae fam. nov. and *Petrotogaceae* fam. nov.) and a new genus *Pseudothermotoga* gen. nov. with five new combinations. Antonie van Leeuwenhoek.2014; 105: 143–168. https://doi.org/10.1007/s10482-013-0062-7 PMID: 24166034

37. Yutin N, Galperin MY. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. Environ Microbiol.2013; 15: 2631–2641. https://doi.org/10.1111/1462-2920.12173 PMID: 23834245

38. Lawson PA, Citron DM, Tyrrell KL, Finegold SM. Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prevot 1938. Anaerobe.2016; 40: 95–99. https://doi.org/10.1016/j.anaerobe.2016.06.008 PMID: 27370902

39. Awal RP, Garcia R, Gemperlein K, Wink J, Kunwar B, Parajuli N, et al. *Vitiosangium cumulatum* gen. nov., sp. nov. and *Vitiosangium subalbum* sp. nov., soil myxobacteria, and emended descriptions of the genera *Archangium* and *Angiococcus*, and of the family *Cystobacteraceae*. Int J Syst Evol Microbiol.2017; 67: 1422–1430. https://doi.org/10.1099/ijsem.0.001829 PMID: 28141508

40. Paster BJ: "Phylum XV. Spirochaetes Garrity and Holt (2001)". In *Bergey's Manual of Systematic Bacteriology Volume 4: The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*. Edited by Brenner DJ, Krieg NR, Garrity GM, Staley JT. New York: Springer Verlag; 2010.

41. Gupta RS, Mahmood S, Adeolu M. A phylogenomic and molecular signature based approach for characterization of the phylum Spirochaetes and its major clades: proposal for a taxonomic revision of the phylum. Front Microbiol.2013; 4: 217. https://doi.org/10.3389/fmicb.2013.00217 PMID: 23908650

42. Norris SJ, Paster BJ, Smibert RM: Genus IV. Treponema Schaudinn 1905, 1728^{AL} (*Spironema* Vuillemin 1905, 1568; *Microspironema* Stiles and Pfender 1905, 936). In *Bergey's Manual of Systematic Bacteriology Volume 4: The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*. Edited by Brenner DJ, Krieg NR, Garrity GM, Staley JT. New York: Springer Verlag; 2010:501–531.

43. Miller SC, Porcella SF, Raffel SJ, Schwan TG, Barbour AG. Large linear plasmids of *Borrelia* species that cause relapsing fever. J Bacteriol.2013; 195: 3629–3639. https://doi.org/10.1128/JB.00347-13 PMID: 23749977

44. Casjens SR, Mongodin EF, Qiu WG, Luft BJ, Schutzer SE, Gilcrease EB, et al. Genome stability of Lyme disease spirochetes: comparative genomics of Borrelia burgdorferi plasmids. PLoS ONE.2012; 7: e33280. https://doi.org/10.1371/journal.pone.0033280 PMID: 22432010

45. Schuler W, Bunikis I, Weber-Lehman J, Comstedt P, Kutschan-Bunikis S, Stanek G, et al. Complete genome sequence of *Borrelia afzelii* K78 and comparative genome analysis. PLoS ONE.2015; 10: e0120548. https://doi.org/10.1371/journal.pone.0120548 PMID: 25798594

46. Casjens SR, Di L, Akther S, Mongodin EF, Luft BJ, Schutzer SE, et al. Primordial origin and diversification of plasmids in Lyme disease agent bacteria. BMC Genomics.2018; 19: 218. https://doi.org/10.1186/s12864-018-4597-x PMID: 29580205

47. Margos G, Marosevic D, Cutler S, Derdakova M, Diuk-Wasser M, Emler S, et al. There is inadequate evidence to support the division of the genus *Borrelia*. Int J Syst Evol Microbiol.2017; 67: 1081–1084. https://doi.org/10.1099/ijsem.0.001717 PMID: 27930271

48. Gao B, Gupta RS. Microbial systematics in the post-genomics era. Antonie van Leeuwenhoek.2012; 101: 45–54. https://doi.org/10.1007/s10482-011-9663-1 PMID: 22048742

49. Estrada-Pena A, Cabezas-Cruz A. Phyloproteomic and functional analyses do not support a split in the genus Borrelia (phylum Spirochaetes). BMC Evol Biol.2019; 19: 54. https://doi.org/10.1186/s12862-019-1379-2 PMID: 30760200

50. Barbour AG: Relapsing fever. In *Harrison's Principle of Internal Medicine*. Edited by Kasper DL, Longo DL, Jameson JL, Loscalzo. New York: McGraw Hill; 2015:1146–1149.

51. Barbour AG, Schwan TG: *Borrelia*, Swellengrebel 1907, 582^{AL}. In *Bergey Manual of Systematics of Bacteria and Archaea, Online @ 2015 Bergey's Manual Trust*. J ohn Wiley & Sons, Inc. Bergey's Manual Trust; 2019:1–22.

52. Barbour AG, Weigang Q: *Borreliella*, Adeolu and Gupta 2014, 1014^{VP}. In *Bergey Manual of Systematics of Bacteria and Archaea, Online @ 2015 Bergey's Manual Trust*. John Wiley & Sons, Inc. Bergey's Manual Trust; 2019:1–23.

53. Wang Z, Wu M. A phylum-level bacterial phylogenetic marker database. Mol Biol Evol.2013; 30: 1258–1262. https://doi.org/10.1093/molbev/mst059 PMID: 23519313

54. Barbour AG: Specificity of *Borrelia*-tick vector relationships. In *SGM Symposium 63: Microbe-Vector Interactions in Vector-Borne Disease*. Edited by Gillispie S, Smith G, Osbourn A. Cambridge: Cambridge University Press; 2004:75–90.

55. Margos G, Fedorova N, Kleinjan JE, Hartberger C, Schwan TG, Sing A, et al. *Borrelia lanei* sp. nov. extends the diversity of *Borrelia* species in California. Int J Syst Evol Microbiol.2017; 67: 3872–3876. https://doi.org/10.1099/ijsem.0.002214 PMID: 28884668

56. Margos G, Lane RS, Fedorova N, Koloczek J, Piesman J, Hojgaard A, et al. *Borrelia bissettiae* sp. nov. and *Borrelia californiensis* sp. nov. prevail in diverse enzootic transmission cycles. Int J Syst Evol Microbiol.2016; 66: 1447–1452. https://doi.org/10.1099/ijsem.0.000897 PMID: 26813420

57. Margos G, Chu CY, Takano A, Jiang BG, Liu W, Kurtenbach K, et al. *Borrelia yangtzensis* sp. nov., a rodent-associated species in Asia, is related to *Borrelia valaisiana*. Int J Syst Evol Microbiol.2015; 65: 3836–3840. https://doi.org/10.1099/ijsem.0.000491 PMID: 28875924

58. Pritt BS, Mead PS, Johnson DK, Neitzel DF, Respicio-Kingry LB, Davis JP, et al. Identification of a novel pathogenic *Borrelia* species causing Lyme borreliosis with unusually high spirochaetaemia: a descriptive study. Lancet Infect Dis.2016; 16: 556–564. https://doi.org/10.1016/S1473-3099(15)00464-8 PMID: 26856777

59. Bissett ML, Hill W. Characterization of *Borrelia burgdorferi* strains isolated from Ixodes pacificus ticks in California. J Clin Microbiol.1987; 25: 2296–2301. PMID: 3323225

60. Postic D, Garnier M, Baranton G. Multilocus sequence analysis of atypical *Borrelia burgdorferi* sensu lato isolates—description of *Borrelia californiensis* sp. nov., and genomospecies 1 and 2. Int J Med Microbiol.2007; 297: 263–271. https://doi.org/10.1016/j.ijmm.2007.01.006 PMID: 17374507