Systematic and Applied Microbiology 39 (2016) 223–225

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

Reply to letter to the editor

A response to Lindsey et al. "*Wolbachia pipientis* should not be split into multiple species: A response to Ramírez-Puebla et al."

In Ramírez-Puebla et al. [18] we compared 34 *Wolbachia* genomes and constructed phylogenetic trees using genomic data. In general, our results were congruent with previously reported phylogenetic trees [5,9]. Our datasets were carefully selected, checked and analyzed avoiding horizontally transferred genes. In the case of the *w*Ana genome we did not use the raw data, but the assembled genome [22] and 31 genes were used to compare in a dataset of conserved proteins. To confirm our conclusions a new phylogenomic analysis was performed excluding the *w*Ana strain in the dataset (Fig. 1). The same topology was obtained, therefore indicating that the results were not affected by the presence of this particular strain.

The genome sequences of strains *w*DacA (BioProject PRJNA274701, GenBank accession number LSYX00000000) and *w*DacB (BioProject PRJNA274698, GenBank accession number LSYY00000000) are in the process of being released. To obtain them, three sequencing methods were used: Illumina, 454 pyrosequencing and PacBio. In a directed strategy to assemble *Wolbachia* genomes, Illumina and 454 reads were mapped to *Wolbachia* genomes using Nucleotide-Nucleotide BLAST 2.2.30. The recovered Illumina reads were assembled using IDBA-UD [17] with the default parameters. Pyrosequencing reads were assembled with Newbler v3.0. Contigs from both assemblies were merged by GARM software [23] and SSPACE software [4] was used for scaffolding. Additionally, the whole metagenome was assembled with SPAdes [2] using Illumina reads and PacBio filtered subreads, as well as 454 reads supplied to the assembler as non-trusted contigs. To recover the two strains from the metagenome, the assembled contigs were aligned by Nucmer [12] against all available *Wolbachia* genomes and those producing significant matches (>50% coverage and >90% identity) were retained. Two contig groups were identified, one showing the highest similarity to sequences from supergroup A strains and other most similar to supergroup B strains. Those groups were independently used as seeds to capture additional contigs with the binning program Clams [16] using tetranucleotide signatures for similarity matching. This process was iteratively repeated until no more contigs were recovered. All recovered contigs were aligned with BLASTX against the nr GenBank database to discard false positives, i.e. those contigs matching bacteria other than *Wolbachia*. The low values of ANI identity and DDH of *w*DacA were not caused by chimeric binning but inherent differences in this strain.

The *Wolbachia* strain naming system in use has focused basically on the host [24]. It is not a practical one and it is not easy to determine if a strain is a symbiont of arthropods or nematodes. When accessing databases like NCBI and searching for the taxonomy of *Wolbachia* one gets lost. It takes a considerable effort to get a glimpse of the immense diversity of *Wolbachia* and its complex and diverse biology. A universal agreement on the designation of species in bacteria has proved to be practical. That is why we proposed to focus on genomes to categorize all *Wolbachia* into different species according to genomic metric systems that have been widely adopted and that facilitate taxonomy [21]. The aim of our study [18] was not to describe *Wolbachia* biology, recombination, phenotypes or all possible hosts, but to designate different species based on genomic comparisons.

Recently, the identification and classification of bacteria have been mainly based on different gene sequences. However, each gene phylogeny may render a particular evolutionary history not always reflecting the species phylogenies. For example, few *Burkholderia pseudomallei* strains genetically unrelated on a whole genome basis were classified as belonging to the same multilocus sequence types. Thus MLST and moreover single marker based phylogenies are considered to have limitations as genotyping methods while whole genome sequencing seems to become the preferred strategy for this [7]. Considering such single gene anomalies, genomic-based approaches have been recognized as providing a better basis for taxonomy studies [21]. A calibration of species distinctiveness was needed and determined using global genomic data [10,19]. From extensive comparisons of recognized species, the overall nucleotide identity was calculated and these values were correlated with the taxonomy gold standard, DNA–DNA hybridization [10]. Novel limits to delineate species were defined as average nucleotide identity, ANI and have been evaluated in different bacterial groups. In other groups of bacteria, phylogenomic approaches have been used to distinguish species and ANI has been used to review reported species [15]. It was subsequently recommended to test the novel ANI estimates and limits with bacteria that may not be cultured and are considered as candidates [11]. Thus, we applied such methods to *Wolbachia* and complemented the study with in *silico* DNA–DNA hybridization and synteny genome comparisons. Additionally, the impact of recombination or lateral gene transfer may be diminished by a genome based perspective. If there is recombination in some genes and genomic regions, the majority of other genes will balance the classification of species toward a more accurate position. Remarkably, it has been observed that there is limited recombination among some of the supergroups analyzed [8] and notably even in highly recombining bacteria, bacterial species names are maintained [3].

Undoubtedly, documented single gene molecular studies cover a larger number of *Wolbachia* lineages whose genomes are not presently available. Certainly those supergroups could not be included in the report at that moment and some of them may subsequently correspond to novel species once their genomes are available, which is to be expected given the current high throughput sequencing capacity. This will not demean our proposal but

DOI of original article: http://dx.doi.org/10.1016/j.syapm.2016.03.001.

http://dx.doi.org/10.1016/j.syapm.2016.03.004

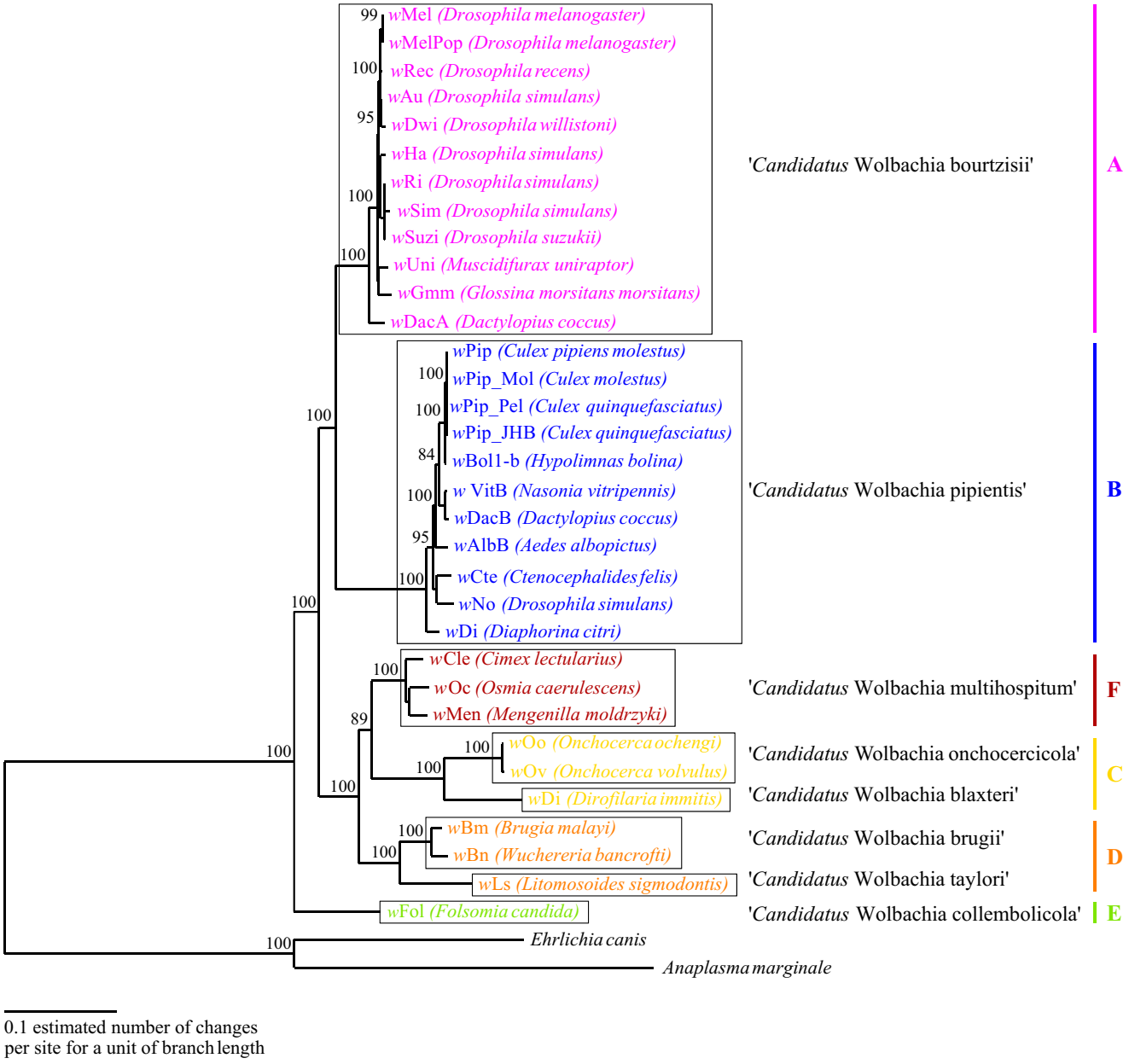

Fig. 1. Phylogenomic tree showing evolutionary relationships between *Wolbachia* strains inferred with PhyML based on a concatenated alignment of 31 marker proteins detected with AMPHORA2 and analyzed with the JTT substitution model. The phylogenic analysis was performed as in [18] but excluding *w*Ana from *Drosophila ananassae.*

rather enrich it. New species are proposed and incorporated every day for different bacterial taxa. Even among cultured bacteria, unnamed bacteria (recognized in many cases only by 16S rRNA gene sequences or by other molecular markers) may outnumber named bacteria. However, this has not precluded the use of species names in those groups. Naming species has advanced in a continuous mode and there is always space for species and genus amendments if properly supported. When first suggested, the use of supergroups for *Wolbachia* classification was presented as a mere proposal based on similarity of the *wsp* gene greater than 97.5%, and it was considered then as an arbitrary value [24]. It was discussed that the number of supergroups would increase as more strains were sequenced and, therefore adjusted within that nomenclature. In the same way, we expect that more species will be included as more genomes are sequenced.

The differences in ANI and in *silico* DDH are sufficiently large to support our conclusion that *Candidatus* Wolbachia blaxteri and *Candidatus* Wolbachia taylori represent different species. Additionally, it is possible to describe bacterial species with only one representative, as it has occurred with many endosymbiont species when they were first described (e.g. [1,6,14]).

For selection of *Wolbachia* species names, we considered that strains from supergroups C, D and E have been found in taxonomically related hosts, and that is why we selected names linked to the taxonomic position of the host. We are aware that not all *Wolbachia* from supergroup E (*Candidatus* Wolbachia collembolicola) are specific to Collembola and that not all *Wolbachia* from supergroup C (*Candidatus* Wolbachia onchocercicola) are specific to the genus *Onchocerca.* But the selection of species names in bacteria does not have to be strictly related to the host. Instead, it is frequently used the name of the host first reported.

The usefulness of naming species has been clearly demonstrated in many bacterial groups and although incipient as it may be in *Wolbachia*, our proposal is a good seed for further taxonomic studies. The identification of novel species within previously recognized species has occurred with novel genomic analysis in bacterial

species such as *Rhizobium* or *Bradyrhizobium*. As an example, *Rhizobium etli* has been split and *Rhizobium phaseoli* and other species were proposed to account for some formerly named *R. etli* strains [13]. Species delineation is not necessarily based on functional divergence. One species may have strain variations and rhizobia are classic examples of this (reviewed in [20]). The *Wolbachia pipientis* original description lacked some fundamental characteristics of other species that can be cultured but was a pioneer effort on the same line, and the name has not been discarded.

The aim of our study was to propose a congruent taxonomy. The designation of unnamed strains may be used until properly classified. In conclusion, our proposal should be freely considered by researchers; its suitability and advantages will be shown in the future.

Acknowledgements

We thank Michael Dunn and Mauro Degli Esposti for critical reading of this manuscript.

References

[1] Aksoy, S. (1995) *Wigglesworthia* gen. nov. and *Wigglesworthia glossinidia* sp. nov., taxa consisting of the mycetocyte-associated, primary endosymbionts of tsetse flies. Int. J. Syst. Bacteriol. 45, 848–851.
[2] Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.
[3] Bennett, J.S., Jolley, K.A., Earle, S.G., Corton, C., Bentley, S.D., Parkhill, J., Maiden, M.C. (2012) A genomic approach to bacterial taxonomy: an examination and proposed reclassification of species within the genus *Neisseria*. Microbiology 158, 1570–1580.
[4] Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D., Pirovano, W. (2011) Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27, 578–579.
[5] Comandatore, F., Sassera, D., Montagna, M., Kumar, S., Koutsovoulos, G., Thomas, G., Repton, C., Babayan, S.A., Gray, N., Cordaux, R., Darby, A., Makepeace, B., Blaxter, M. (2013) Phylogenomics and analysis of shared genes suggest a single transition to mutualism in *Wolbachia* of nematodes. Genome Biol. Evol. 5, 1668–1674.
[6] Dale, C., Maudlin, I. (1999) *Sodalis* gen. nov. and *Sodalis glossinidius* sp. nov., a microaerophilic secondary endosymbiont of the tsetse fly *Glossina morsitans morsitans*. Int. J. Syst. Bacteriol. 49 (Pt 1), 267–275.
[7] De Smet, B., Sarovich, D.S., Price, E.P., Mayo, M., Theobald, V., Kham, C., Heng, S., Thong, P., Holden, M.T., Parkhill, J., Peacock, S.J., Spratt, B.G., Jacobs, J.A., Vandamme, P., Currie, B.J. (2015) Whole-genome sequencing confirms that *Burkholderia pseudomallei* multilocus sequence types common to both Cambodia and Australia are due to homoplasy. J. Clin. Microbiol. 53, 323–326.
[8] Ellegaard, K.M., Klasson, L., Näslund, K., Bourtzis, K., Andersson, S.G. (2013) Comparative genomics of *Wolbachia* and the bacterial species concept. PLoS Genet. 9, e1003381.
[9] Gerth, M., Gansauge, M.-T., Weigert, A., Bleidron, C. (2014) Phylogenomic analyses uncover origin and spread of the *Wolbachia* pandemic. Nat. Commun. 5, 5117.
[10] Konstantinidis, K.T., Tiedje, J.M. (2004) Genomic insights that advance the species definition for prokaryotes. Proc. Natl. Acad. Sci. 102 (7), 2567–2572.
[11] Konstantinidis, K.T., Rosselló-Móra, R. (2015) Classifying the uncultivated microbial majority: a place for metagenomic data in the *Candidatus* proposal. Syst. Appl. Microbiol. 38 (4), 223–230.
[12] Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., Salzberg, S.L. (2004) Versatile and open software for comparing large genomes. Genome Biol. 5, R12.
[13] López-Guerrero, M.G., Ormeño-Orrillo, E., Velásquez, E., Rogel, M.A., Acosta, J.L., González, V., Martínez, J., Martínez-Romero, E. (2012) *Rhizobium etli* taxonomy revised with novel genomic data and analysis. Syst. Appl. Microbiol. 35 (6), 353–358.
[14] McCutcheon, J.P., von Dohlen, C.D. (2011) An interdependent metabolic patchwork in the nested symbiosis of mealybugs. Curr. Biol. 21, 1366–1372.
[15] Ormeño-Orrillo, E., Servín-Garciadueñas, L.E., Rogel, M.A., González, V., Peralta, H., Mora, J., Martínez-Romero, J., Martínez-Romero, E. (2014) Taxonomy of rhizobia and agrobacteria from the Rhizobiaceae family in light of genomics. Syst. Appl. Microbiol. 38 (4), 287–291.
[16] Pati, A., Heath, L.S., Kyrpides, N.C., Ivanova, N. (2011) ClaMS: a classifier for metagenomic sequences. Stand. Genomic Sci. 5, 248–253.
[17] Peng, Y., Leung, H.C., Yiu, S.M., Chin, F.Y. (2012) IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28 (11), 1420–1428.
[18] Ramírez-Puebla, S.T., Servín-Garciadueñas, L.E., Ormeño-Orrillo, E., Vera-Ponce de Leon, A., Rosenblueth, M., Delaye, L., Martínez, J., Martínez-Romero, E. (2015) Species in Wolbachia? Proposal for the designation of '*Candidatus* Wolbachia bourtzisii', '*Candidatus* Wolbachia onchocercicola', '*Candidatus* Wolbachia blaxteri', '*Candidatus* Wolbachia brugii', '*Candidatus* Wolbachia taylori', '*Candidatus* Wolbachia collembolicola' and '*Candidatus* Wolbachia multihospitum' for the different species within *Wolbachia* supergroups. Syst. Appl. Microbiol. 38 (6), 390–399.
[19] Richter, M., Rosselló-Móra, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. 106 (45), 19126–19131.
[20] Rogel, M.A., Ormeño-Orrillo, E., Martinez Romero, E. (2011) Symbiovars in rhizobia reflect bacterial adaptation to legumes. Syst. Appl. Microbiol. 34, 96–104.
[21] Rosselló-Móra, R., Amann, R. (2015) Past and future species definitions for Bacteria and Archaea. Syst. Appl. Microbiol. 38 (4), 209–216.
[22] Salzberg, S.L., Dunning Hotopp, J.C., Delcher, A.L., Pop, M., Smith, D.R., Eisen, M.B., Nelson, W.C. (2005) Serendipitous discovery of *Wolbachia* genomes in multiple *Drosophila* species. Genome Biol. 6, R23.
[23] Soto-Jimenez, L.M., Estrada, K., Sánchez-Flores, A. (2014) GARM: genome assembly, reconciliation and merging pipeline. Curr. Top. Med. Chem. 14 (3), 418–424.
[24] Zhou, W., Rousset, F., OíNeill, S. (1998) Phyllogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. Proc. R. Soc. Lond. 265, 509–515.

Shamayim T. Ramírez-Puebla
Centro de Ciencias Genómicas, UNAM, Cuernavaca, Morelos, Mexico

Luis E. Servín-Garcidueñas
Laboratorio Nacional de Análisis y Síntesis Ecológica para la Conservación de Recursos Genéticos, ENES-Morelia, UNAM, Morelia, Michoacán, Mexico

Ernesto Ormeño-Orrillo
Arturo Vera-Ponce de León
Mónica Rosenblueth
Centro de Ciencias Genómicas, UNAM, Cuernavaca, Morelos, Mexico

Luis Delaye
Departamento de Ingeniería Genética, CINVESTAV-Irapuato, Irapuato, Guanajuato, Mexico

Julio Martínez
Esperanza Martínez-Romero*
Centro de Ciencias Genómicas, UNAM, Cuernavaca, Morelos, Mexico

* Corresponding author. Tel.: +52 777 3131697.
E-mail address: emartine@ccg.unam.mx
(E. Martínez-Romero)