

1 **Classification of *Wolbachia* (Alphaproteobacteria, Rickettsiales): No evidence for a distinct**
2 **supergroup in cave spiders**

3

4 Michael Gerth

5

6

7 **Contact**

8 Michael Gerth

9 Institute for Integrative Biology

10 University of Liverpool

11 Biosciences Building

12 Liverpool L69 7ZB

13 United Kingdom

14 michael.gerth@liv.ac.uk

15

16

17 **Keywords**

18 *Wolbachia* supergroups, bacterial systematics, endosymbionts, molecular phylogeny

19 ***To the editor***

20 *Wolbachia* are intracellular, inherited Alphaproteobacteria present in a large proportion of arthropod
21 species and in filarial nematodes. *Wolbachia*–host interactions are many-faceted and include, but
22 are not limited to reproductive manipulations (Werren et al. 2008), nutritional mutualism (Nikoh et
23 al. 2014), and protection from pathogens (Hedges et al. 2008, Teixeira et al. 2008). Although there
24 is just a single *Wolbachia* species described, the genus is highly diverse with regard to its
25 distribution, phenotypes induced in the host, and genomic architecture. The generally accepted
26 classification scheme currently in use differentiates among genetically distinct, monophyletic
27 lineages named “supergroups” (Lindsey et al. 2016). To date, 16 of these *Wolbachia* lineages that –
28 based on multiple genetic markers – are clearly distinct from another were described (supergroups
29 A–F and H–Q, Glowska et al. 2015).

30 In a recent study, Wang et al. (2016, from here on “Wang et al.”) reported the discovery of a
31 novel, 17th *Wolbachia* supergroup (“R”) from cave spiders (*Telema* spp.). The authors base this
32 conclusion on phylogenetic analyses of sequences from three protein coding genes (*ftsZ*, *coxA*,
33 *groEL*) and the 16S rRNA gene. Here, I re-analyse these data and show that *Wolbachia* from *Telema*
34 spiders clusters with supergroup A strains and thus, the creation of a novel *Wolbachia* supergroup
35 for these strains is without justification.

36 Using the sequence data generated by Wang et al., I first performed an online BLAST search
37 of all *ftsZ*, *coxA*, *groEL*, and 16S sequences (2, 6, 1, and 5 haplotypes, respectively) against the
38 nucleotide database of NCBI GenBank. Strikingly, all of the sequences match to known supergroup
39 A strains with an identity greater than 99%. Next, using Mafft version 7.245 (Katoh and Standley
40 2013) with automatic parameter settings, I aligned all novel sequences with the corresponding
41 orthologs extracted from the *Wolbachia* genomes wRi, wHa, wMel (all supergroup A), wPip & wNo
42 (supergroup B), and wCle (supergroup F). Using Aliview version 1.17.1 (Larsson 2014), the
43 datasets were then trimmed manually to include only positions that were sequenced by Wang et al.

44 All alignments are available under https://github.com/gerthmicha/supergroup_R. From these
45 alignments, raw genetic distances were calculated using the “ape” package within the R statistical
46 environment (Paradis et al. 2004, R Core Team 2016). All *Wolbachia* loci from *Telema* hosts are
47 very similar to each other and to the reference supergroup A loci (~0–1% distance, Fig. S1). In
48 general, the distances within supergroup A (including *Telema* strains) are almost one order of
49 magnitude smaller compared to the distances between supergroups (Fig. S1). The distance measures
50 are therefore in line with the conclusion that *Wolbachia* from cave spiders are supergroup A strains.

51 Finally, maximum likelihood trees were created for all single gene alignments and a
52 supermatrix of all concatenated genes using IQ-TREE multicore version 1.4.1 (Nguyen et al. 2015)
53 with automated model selection (“-m TEST”) and 1000 ultrafast bootstrap replicates. All of the
54 single gene trees and the tree based on the concatenated genes strongly support an association of
55 *Wolbachia* strains from *Telema* ssp. with supergroup A (Fig. 1). This becomes especially clear when
56 comparing the genetic distances within supergroup A and the distances between supergroups A, B,
57 and F (Fig. 1).

58 Visual comparisons of the trees presented in Fig. 1 and the ones generated by Wang et al.
59 reveal little discordance. Single gene trees from maximum likelihood analyses of *coxA*, *ftsZ*, and
60 *16S* sequences (Supplementary Figure 1 in Wang et al.) show a very close association of *Wolbachia*
61 sequences from *Telema* ssp. with supergroup A that was also recovered in the re-analyses presented
62 here (although for *16S*, no supergroup A reference sequences were included by Wang et al.).
63 However, the analysis of *groEL* sequences (Supplementary Figure 1D in Wang et al.) and the one of
64 the concatenated dataset resulted in a placement of *Wolbachia* from *Telema* ssp. distinct from other
65 supergroups. From the data available, these analyses are not replicable. Speculatively, an alignment
66 artefact has led to the erroneous placement of *Wolbachia* from *Telema* ssp. in the analysis of Wang
67 et al.

68 The discordance between single gene trees and the concatenated analysis was noted by

69 Wang et al. and led them to conclude that a single locus is not sufficient to resolve phylogenetic
70 relationships within *Wolbachia*. Whilst this is certainly true, the observation that each single gene
71 analysis consistently places *Wolbachia* from *Telema* ssp. within (or closely associated to)
72 supergroup A questions the validity of the supermatrix analysis. Further to this, the fact that
73 *Wolbachia* from *Telema* ssp. form a monophyletic group in most of the analyses does not imply that
74 they should be considered a separate supergroup, as Wang et al. suggest. Naturally, each *Wolbachia*
75 supergroup is composed of multiple monophyletic lineages. The reference sequences used here and
76 by Wang et al. do not reflect the full genetic diversity of supergroup A strains (which are very
77 common among arthropods), and while *Wolbachia* from *Telema* ssp. may seem distinct in these
78 analyses, this is not necessarily the case when considering the multitude of supergroup A alleles of
79 the loci in question. Finally, the detection of supergroup A-like prophage sequences in *Telema* hosts
80 provides further evidence for their inclusion in this supergroup.

81 In summary, the here presented similarity searches, genetic distance measures, and
82 phylogenetic analyses are all consistent and provide strong evidence for the association of
83 *Wolbachia* strains from *Telema* with supergroup A. Their placement within a novel supergroup, as
84 proposed by Wang et al., lacks any support and is likely artefactual. In order to minimize confusion,
85 it may be helpful not to assign the name “supergroup R” to any novel *Wolbachia* lineage potentially
86 to be discovered in the future.

87

88 ***Acknowledgements***

89 I am grateful to Guan-Hong Wang for providing sequence data and for discussions. I thank
90 Franziska Anni Franke for comments on the manuscript. My current position is funded by EMBO
91 (ALTF 48-2015) and co-funded by Marie-Curie Actions of the European Commission
92 (LTFCOFUND2013, GA-2013-609409).

93

94 **References**

- 95 Glowska, E., Dragun-Damian, A., Dabert, M. & Gerth, M. (2015). New *Wolbachia* supergroups
96 detected in quill mites (Acari: Syringophilidae). *Infect. Genet. Evol.*, 30, 140-146.
- 97 Hedges, L. M., Brownlie, J. C., O'Neill, S. L. & Johnson, K. N. (2008). *Wolbachia* and virus
98 protection in insects. *Science*, 322, 702.
- 99 Katoh, K. & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7:
100 improvements in performance and usability. *Mol. Biol. Evol.*, 30, 772-780.
- 101 Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets.
102 *Bioinformatics*, 30, 3276-3278.
- 103 Lindsey, A. R., Bordenstein, S. R., Newton, I. L. & Rasgon, J. L. (2016). *Wolbachia pipientis*
104 should not be split into multiple species: A response to Ramirez-Puebla et al., "Species in
105 *Wolbachia*? Proposal for the designation of 'Candidatus *Wolbachia bourtzisii*', 'Candidatus
106 *Wolbachia onchocercicola*', 'Candidatus *Wolbachia blaxteri*', 'Candidatus *Wolbachia*
107 *brugii*', 'Candidatus *Wolbachia taylori*', 'Candidatus *Wolbachia collembolicola*' and
108 'Candidatus *Wolbachia multihospitum*' for the different species within *Wolbachia*
109 supergroups". *Syst. Appl. Microbiol.*, doi:10.1016/j.syapm.2016.03.001.
- 110 Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. (2015). IQ-TREE: A fast and
111 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol.*
112 *Evol.*, 32, 268-274.
- 113 Nikoh, N., Hosokawa, T., Moriyama, M., Oshima, K., Hattori, M. & Fukatsu, T. (2014).
114 Evolutionary origin of insect-*Wolbachia* nutritional mutualism. *Proc. Natl. Acad. Sci. U.S.A.*,
115 111, 10257-10262.
- 116 Paradis, E., Claude, J. & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R
117 language. *Bioinformatics*, 20, 289-290.
- 118 R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for

119 Statistical Computing, Vienna, Austria.

120 Teixeira, L., Ferreira, A. & Ashburner, M. (2008). The bacterial symbiont *Wolbachia* induces
121 resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol.*, 6, e1000002.

122 Wang, G.-H., Jia, L.-Y., Xiao, J.-H. & Huang, D.-W. (2016). Discovery of a new *Wolbachia*
123 supergroup in cave spider species and the lateral transfer of phage WO among distant hosts.
124 *Infect. Genet. Evol.*, doi:10.1016/j.meegid.2016.03.015.

125 Werren, J. H., Baldo, L. & Clark, M. E. (2008). *Wolbachia*: master manipulators of invertebrate
126 biology. *Nat. Rev. Microbiol.*, 6, 741-751.

127

128

129 **Fig. 1** Phylogenetic analyses of *Wolbachia* sequences isolated from *Telema* ssp. (cave spiders).

130 Maximum likelihood trees are shown for each single gene analysis and for the concatenated
131 alignment of all four genes. Numbers on branches correspond to bootstrap values from 1000
132 pseudoreplicates. Only values above or equal to 70 are shown. Single gene trees are midpoint
133 rooted. Reference sequences are highlighted as follows: blue- supergroup A, red- supergroup B,
134 green- supergroup F. Naming of the sequences from *Telema* hosts follows Wang et al. (2016).

135

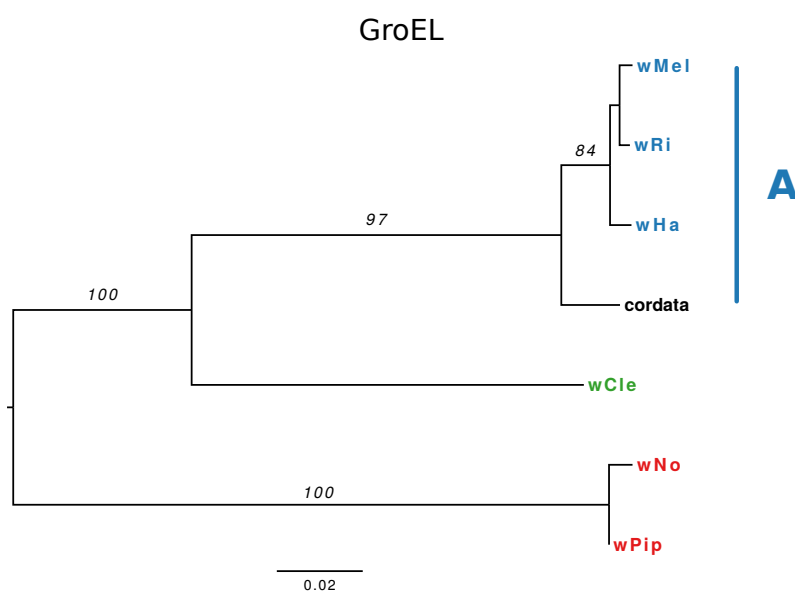
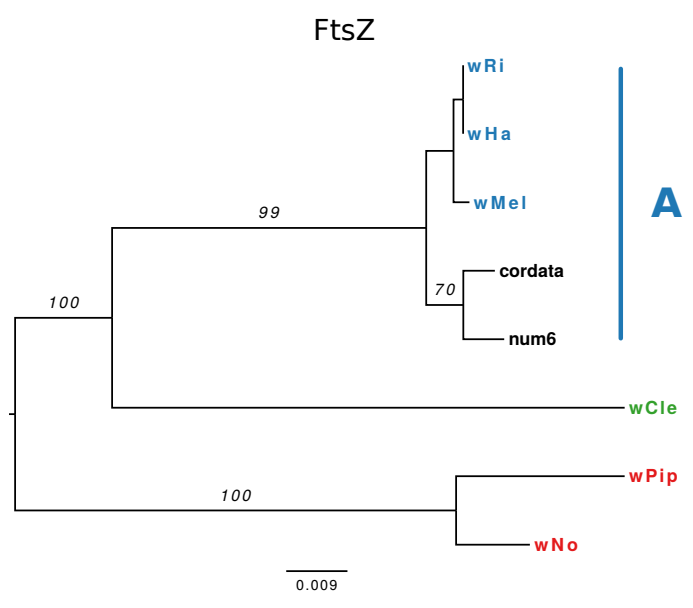
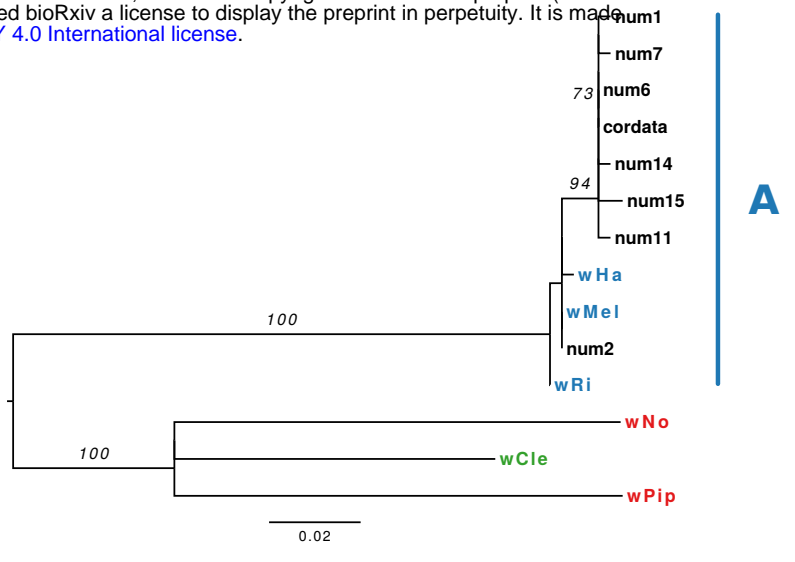
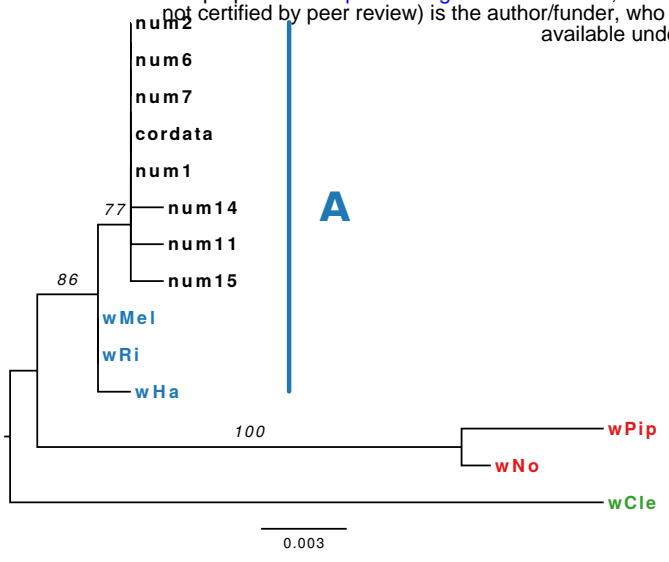
136 **Supplementary information**

137 **Fig. S1** Distance matrices (uncorrected p-distances) for each alignment analysed in this study.

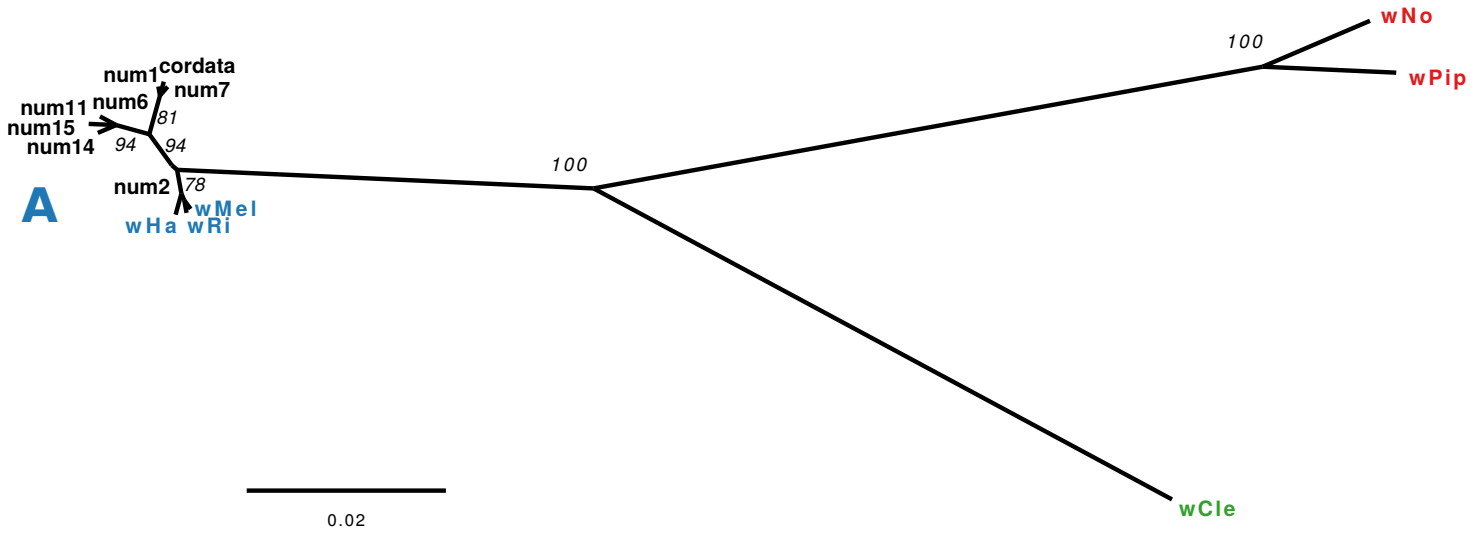
138 Sequences from *Telema* sp. are highlighted in bold.

165

CoxA



concatenated



Supergroup A	num1	num2	num6	num7	num11	num14	num15	cordata	wHa	wMel	wRi	wNo	wPip	
	num2	0.00%												
	num6	0.00%	0.00%											16S
	num7	0.00%	0.00%	0.00%										
	num11	0.11%	0.11%	0.11%	0.11%									
	num14	0.11%	0.11%	0.11%	0.11%	0.23%								
	num15	0.11%	0.11%	0.11%	0.11%	0.23%	0.23%							
	cordata	0.00%	0.00%	0.00%	0.00%	0.11%	0.11%	0.11%						
	wHa	0.23%	0.23%	0.23%	0.23%	0.34%	0.34%	0.34%	0.23%					
wMel	0.11%	0.11%	0.11%	0.11%	0.23%	0.23%	0.23%	0.11%	0.11%					
wRi	0.11%	0.11%	0.11%	0.11%	0.23%	0.23%	0.23%	0.11%	0.11%	0.00%				
F B	wNo	1.72%	1.72%	1.72%	1.72%	1.83%	1.83%	1.60%	1.72%	1.72%	1.60%			
	wPip	1.83%	1.83%	1.83%	1.83%	1.95%	1.95%	1.72%	1.83%	2.06%	1.95%	1.95%	0.57%	
	wCle	2.17%	2.17%	2.17%	2.17%	2.29%	2.29%	2.29%	2.17%	2.17%	2.06%	2.06%	3.20%	3.32%

Supergroup A	num1	num2	num6	num7	num11	num14	num15	cordata	wHa	wMel	wRi	wNo	wPip	
	num2	0.89%												
	num6	0.30%	0.59%											coxA
	num7	0.59%	0.89%	0.30%										
	num11	0.59%	0.89%	0.30%	0.59%									
	num14	0.59%	0.89%	0.30%	0.59%	0.59%								
	num15	0.89%	1.19%	0.59%	0.89%	0.89%	0.89%							
	cordata	0.30%	0.59%	0.00%	0.30%	0.30%	0.30%	0.59%						
	wHa	1.19%	0.30%	0.89%	1.19%	1.19%	1.19%	1.48%	0.89%					
wMel	0.89%	0.00%	0.59%	0.89%	0.89%	0.89%	1.19%	0.59%	0.30%					
wRi	1.19%	0.30%	0.89%	1.19%	1.19%	1.19%	1.48%	0.89%	0.30%	0.30%				
F B	wNo	13.35%	12.76%	13.06%	13.35%	13.35%	13.35%	13.65%	13.06%	12.46%	12.76%	12.46%		
	wPip	12.17%	11.57%	11.87%	12.17%	12.17%	12.17%	12.46%	11.87%	11.87%	11.57%	11.28%	11.28%	
	wCle	12.76%	11.87%	12.46%	12.76%	12.76%	12.76%	13.06%	12.46%	11.57%	11.87%	11.57%	10.39%	10.39%

A	num6	cordata	wHa	wMel	wRi	wNo	wPip	
	cordata	1.11%						
	wHa	1.33%	1.11%				ftsZ	
	wMel	1.11%	0.89%	0.22%				
F B	wRi	1.33%	1.11%	0.00%	0.22%			
	wNo	10.20%	10.20%	10.20%	9.98%	10.20%		
	wPip	10.42%	10.42%	10.42%	10.20%	10.42%	1.11%	
F	wCle	10.86%	10.64%	10.86%	10.64%	10.86%	13.08%	12.86%

A	cordata	wHa	wMel	wRi	wNo	wPip
	wHa	2.59%				
	wMel	2.72%	0.91%			
F B	wRi	2.46%	0.91%	0.52%		groEL
	wNo	12.45%	11.67%	11.93%	11.93%	
	wPip	12.19%	11.41%	11.67%	11.67%	0.52%
	wCle	10.64%	10.38%	10.64%	10.38%	12.71%

