

## *Supplementary Material*

### **1 Supplementary Methods**

#### **1.1 PCR protocol for amplicon run 0, 1, 2, and 3:**

Three PCR amplifications were carried out for each sample, the first using Bac 341F and Bac 805R primers to amplify variable regions V3 and V4 (Herlemann et al., 2011), the second adding overhang adapters with modified primers and finally an index PCR with Nextera XT index primers (Illumina) for sample identification. Each PCR had a total reaction volume of 25  $\mu$ l containing 2.5  $\mu$ l DNA sample, 0.5  $\mu$ l of each primer (primer stock concentration = 10pmol/ $\mu$ l), 12.5  $\mu$ l 2x KAPA HiFi HotStart Ready Mix and 9  $\mu$ l sterile H<sub>2</sub>O. The PCRs were performed on a Veriti® 96-Well Thermal Cycler (Applied Biosystems®) with the following program: Denaturation at 95°C for 3 min. Cycles of 95°C, 55°C and 72°C for 30 secs each, and finally a 5 min extension step at 72°C. PCR 1 with 20 cycles, PCR 2 with 10 cycles and PCR 3 with 8 cycles. After each PCR, samples were purified using AMPure XP beads (Illumina).

#### **1.2 PCR protocol for amplicon run 4:**

Run 4 was a trouble shooting run, redoing PCRs and sequencing for samples which had failed to provide adequate read numbers in previous runs. Based on several trials and optimization iterations, the PCR protocol for samples in amplicon run 4 was carried out differently for different groups of samples. Table S7 summarizes these differences while the text below explains the details.

For several samples PCR 1 and PCR 2 were carried out normally as described for amplicon run 0-3, But some samples went through the following optimized protocol: PCR 1 was run with a diluted template and added BSA giving a total reaction volume of 25  $\mu$ l containing 2.5  $\mu$ l DNA sample (diluted 1:50), 0.5  $\mu$ l of each primer (primer stock concentration = 10pmol/ $\mu$ l), 0.5  $\mu$ M BSA (10mg/ml), 12.5  $\mu$ l 2x KAPA HiFi HotStart Ready Mix and 8.5  $\mu$ l sterile H<sub>2</sub>O. This was run for 30 cycles rather than the 20 cycles normally run for PCR1. PCR 2 was run in triplicate for each sample with 3  $\mu$ l DNA sample, 0.5  $\mu$ l of each primer (primer stock concentration = 10pmol/ $\mu$ l), 12.5  $\mu$ l 2x KAPA HiFi HotStart Ready Mix and 8.5  $\mu$ l sterile H<sub>2</sub>O, and for 12 cycles instead of the normal 10.

For all samples in amplicon run 4, PCR 2 products were size separated by agarose gel electrophoresis and the desired band was extracted from the gels using an E.Z.N.A. Gel Extraction Kit (Omega Bio-Tek). Several samples (see Table S7) in run 4 had such low DNA concentration after gel extraction that AMPure XP beads (Illumina) were used to bind and resuspend in a smaller volume. PCR 3 adding index primers was run as described for amplicon run 0-3, but with differing amounts of template depending on DNA concentration (see Table S7).

Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., & Andersson, A. F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME Journal*, 5(10), 1571–1579. doi: 10.1038/ismej.2011.41

## 2 Supplementary Figures and Tables

**Table S1** Summary statistics of samples, sequencing, and community data

	<i>S. dumicola</i>	<i>S. mimosarum</i>	<i>S. sarasinorum</i>
<b>Samples</b>	58	60	98
<b>Nests</b>	28	23	34
<b>Populations</b>	5	5	8
<b>Raw read number<sup>a,c</sup></b>	43079 (5278-675594)	51551 (14601-96639)	44425 (3959-280637)
<b>Filtered read number<sup>a,d</sup></b>	30270 (3131-438346)	30368 (4059-68412)	31146 (3004-195228)
<b>ASV number<sup>d</sup></b>	772	1470	1567
<b>ASV number per sample<sup>a,d</sup></b>	24.52 (4-152)	44.73 (3-211)	46.97 (3-343)
<b>Shannon diversity<sup>b,d</sup></b>	0.56 ± 0.67	1.58 ± 1.30	1.71 ± 1.11
<b>Simpson diversity<sup>b,d</sup></b>	0.77 ± 0.25	0.48 ± 0.38	0.40 ± 0.29
<b>DMN<sup>b,d</sup></b>	0.93 ± 0.13	0.68 ± 0.32	0.68 ± 0.27

Note: Based on samples with minimum 3000 reads after filtering

a: Mean (minimum-maximum)

b: Mean ± standard deviation

c: Unfiltered forward read numbers from Miseq

d: After quality filtering, denoising, paired-end merging, chimera finding, length filtering (minimum 400bp) and taxonomic filtering

**Table S2** R scripts used for analysis and data visualization

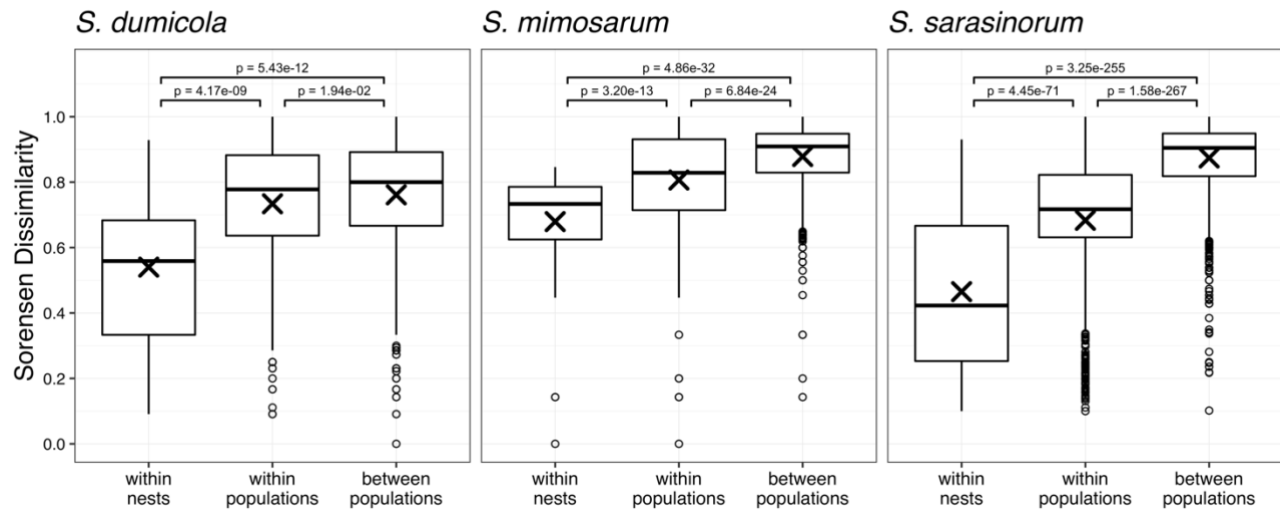
Use	Figure	GitHub link
<b>Basic DADA2 analysis</b>	-	<a href="https://github.com/Mettetron/3Species/blob/master/DADA2_basicAnalysis.R">https://github.com/Mettetron/3Species/blob/master/DADA2_basicAnalysis.R</a>
<b>Merging of ASV data from several sequencing runs</b>	-	<a href="https://github.com/Mettetron/3Species/blob/master/DADA2_combRuns_tax.R">https://github.com/Mettetron/3Species/blob/master/DADA2_combRuns_tax.R</a>
<b>Data formatting</b>	-	<a href="https://github.com/Mettetron/3Species/blob/master/DADA2_to_Phyloseq.R">https://github.com/Mettetron/3Species/blob/master/DADA2_to_Phyloseq.R</a>
<b>Decontamination and normalization</b>	-	<a href="https://github.com/Mettetron/3Species/blob/master/DADA2_filterAndNorm.R">https://github.com/Mettetron/3Species/blob/master/DADA2_filterAndNorm.R</a>
<b>Richness, alpha diversity and dominance boxplots</b>	2	<a href="https://github.com/Mettetron/3Species/blob/master/ASV_divDom_box.R">https://github.com/Mettetron/3Species/blob/master/ASV_divDom_box.R</a>
<b>Prevalence vs. abundance scatter plots</b>	3	<a href="https://github.com/Mettetron/3Species/blob/master/ASV_coreDom_scatter.R">https://github.com/Mettetron/3Species/blob/master/ASV_coreDom_scatter.R</a>
<b>Prevalence heatmap</b>	4	<a href="https://github.com/Mettetron/3Species/blob/master/ASV_prevalenceHeatmap.R">https://github.com/Mettetron/3Species/blob/master/ASV_prevalenceHeatmap.R</a>
<b>Diversity boxplots</b>	7	<a href="https://github.com/Mettetron/3Species/blob/master/ASV_Bdiv_boxplots.R">https://github.com/Mettetron/3Species/blob/master/ASV_Bdiv_boxplots.R</a>
<b>Ordination</b>	8	<a href="https://github.com/Mettetron/3Species/blob/master/ASV_NMDS_ordination.R">https://github.com/Mettetron/3Species/blob/master/ASV_NMDS_ordination.R</a>
<b>Stacked barchart for individual hosts</b>	S1	<a href="https://github.com/Mettetron/3Species/blob/master/ASV_barchartsIndividuals_dumicola.R">https://github.com/Mettetron/3Species/blob/master/ASV_barchartsIndividuals_dumicola.R</a>
<b>Diversity vs. Geographical distance</b>	S4	<a href="https://github.com/Mettetron/3Species/blob/master/ASV_betaDist_vs_geoDist.R">https://github.com/Mettetron/3Species/blob/master/ASV_betaDist_vs_geoDist.R</a>

Note: Example data for running scripts also available on GitHub <https://github.com/Mettetron/3Species>

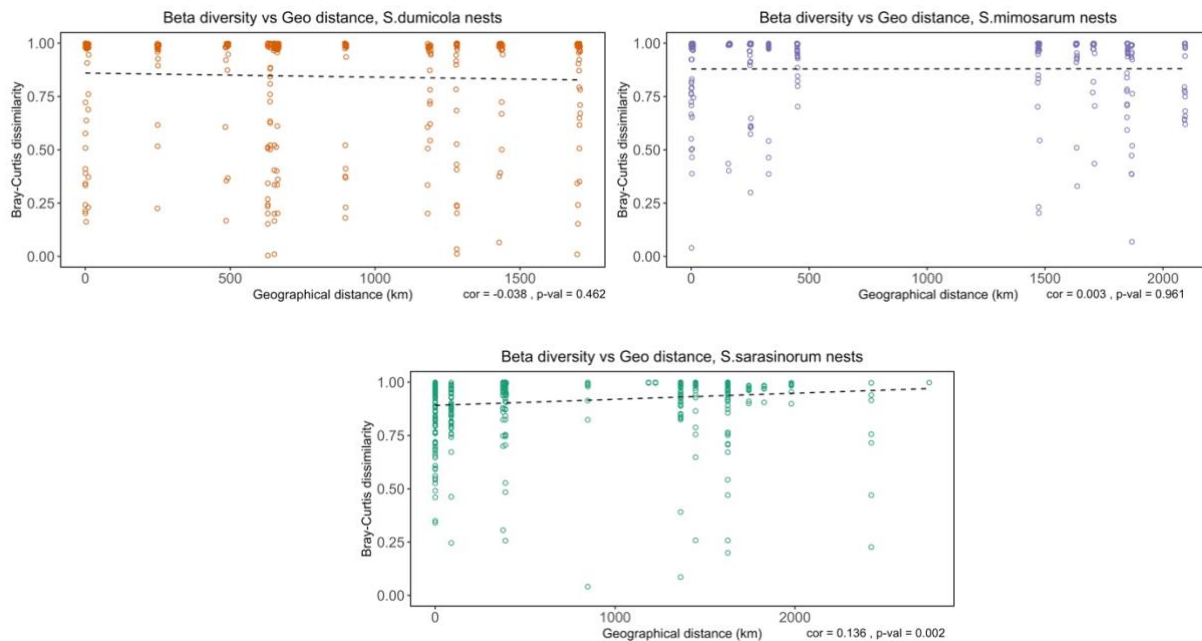




**Supplementary Figure S2.** Phylogenetic trees with dominant and core *Stegodyphus* symbionts. (A) *Diplorickettsia*, (B) *Rickettsia*/Rickettsiaceae, (C) *Acaricomes*, (D) Weeksellaceae, (E) *Entomoplasma*/*Spiroplasma*/*Mycoplasma*, (F) *Brevibacterium*. Bayesian inference consensus trees based on full length 16S rRNA sequences from bacteria closely related to social *Stegodyphus* symbionts found in the silva database (v128). Numbers on nodes show posterior probabilities. Blue text indicates sequences from this study. Short sequences (indicated by dotted lines, e.g. of ASVs) were added without changing the tree topology. Tip labels: Accession number, **full name**, and (host) of bacteria. Scalebar: 0.1 estimated substitutions per site. The placement of *Brevibacterium* ASV\_8 was ambiguous and could have been placed with all tips marked with X.



**Supplementary Figure S3.** Beta diversity at three different host organization levels (within nest, within population, and between populations) in the three species of social *Stegodyphus*. Each data point compares the microbiomes of two individual spiders. For estimating the Sørensen dissimilarities the data were subsampled to 3000 ASVs per spider. Group means are indicated with “x” and compared using one-way ANOVA and pairwise t-test with Benjamini-Hochberg adjusted p-values.

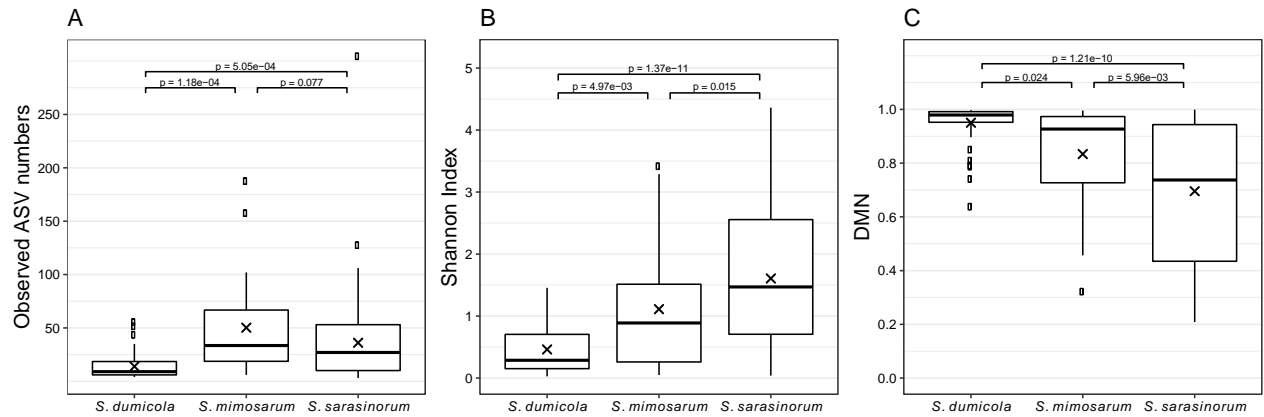


**Supplementary Figure S4.** Correlation of beta diversity and geographical distance between social *Stegodyphus* nests. Each data point represents comparison between average microbiomes of two nests. Beta diversity measures based on subsampled ASV data. Pearson’s product-moment correlation.

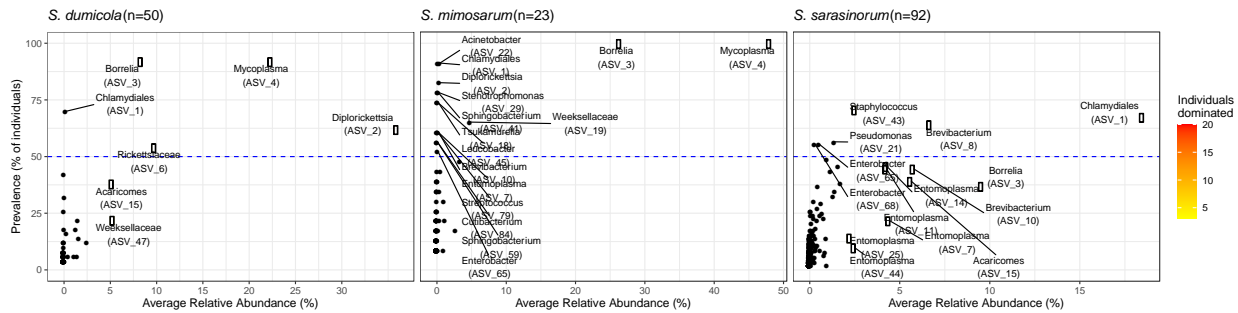
Table S8 – ASV BLAST table

ASVs	Best BLAST match in the nucleotide collection database			
	Seq. ID <sup>†</sup>	Name	Source/Host	Accession
<b>ASV_1</b> <b>Chlamydiales</b>	89% (381/428)	Uncultured <i>Chlamydia</i> sp.	Soil (Scotland)	LN831101.1
<b>ASV_2</b> <b>Diplorickettsia</b>	99% (426/427)	Uncultured <i>Diplorickettsia</i> sp.	Tick ( <i>Ixodes scapularis</i> )	MN192053.1
<b>ASV_3</b> <i>Borrelia</i>	97% (414/427)	Uncultured bacterium clone	Scorpion ( <i>Centruroides limpidus</i> )	KM978314.1
<b>Clone seq.</b> <b>MH627299</b>	95% (1426/1500)	Uncultured bacterium clone	Scorpion ( <i>Centruroides limpidus</i> )	KM978314.1
<b>ASV_4</b> <b>Mycoplasma</b>	91% (389/426)	Uncultured bacterium clone	Plant ( <i>Tamarix nilotica</i> )	JF302694.1
<b>Clone seq.</b> <b>MH627308</b>	86% (1276/1481)	Uncultured bacterium clone	Scorpion ( <i>Vaejovis smithi</i> )	MG813921.1
<b>ASV_5</b> <i>Rickettsia</i>	100% (402/402)	<i>Candidatus Rickettsia tarasevichiae</i>	Tick ( <i>Ixodes persulcatus</i> )	MN446745.1
<b>ASV_6</b> <b>Rickettsiaceae</b>	99% (398/402)	Rickettsiales bacterium Ac37b	Tick ( <i>Amblyomma cajennense</i> )	CP009217.2
<b>ASV_7</b> <b>Entomoplasma</b>	98% (417/427)	<i>Entomoplasma freundtii</i>	Beetle (Cicindelidae)	CP024962.1
<b>ASV_11</b> <b>Entomoplasma</b>	98% (418/427)	<i>Entomoplasma freundtii</i>	Beetle (Cicindelidae)	CP024962.1
<b>ASV_14</b> <b>Entomoplasma</b>	97% (416/427)	<i>Entomoplasma freundtii</i>	Beetle (Cicindelidae)	CP024962.1
<b>ASV_25</b> <b>Entomoplasma</b>	97% (416/427)	<i>Entomoplasma somnilux</i>	Beetle ( <i>Pyractomena angulata</i> )	CP024965.1
<b>ASV_44</b> <b>Entomoplasma</b>	97% (415/427)	<i>Entomoplasma freundtii</i>	Beetle (Cicindelidae)	CP024962.1
<b>ASV_8, ASV_10</b> <b>Brevibacterium</b>	100% (407/407)	100+ sequences. E.g. <i>B. sediminis</i> , <i>B. epidermidis</i> , <i>B. sanguinis</i>	E.g. Silkworm, Cell phone, Deep-sea sediment, Human	NA
<b>ASV_12</b> <b>Weeksellaceae</b>	97% (408/422)	Uncultured bacterium clone	Shrimp ( <i>Penaeus vannamei</i> )	KP953663.1
<b>Clone seq.</b> <b>MH627331</b>	95% (1416/1487)	<i>Bergeyella zoohelcum</i>	Cat ( <i>Felis catus</i> )	KM461977
<b>ASV_47</b> <b>Weeksellaceae</b>	96% (407/422)	Uncultured bacterium clone	Shrimp ( <i>Penaeus vannamei</i> )	KP953663.1
<b>ASV_15</b> <b>Acaricomes</b>	99% (413/416)	<i>Acaricomes phytoseiuli</i>	Plant ( <i>Vitis vinifera</i> ) / Mite?	HM803939.1
<b>Clone seq.</b> <b>MH627364</b>	98% (1454/1483)	<i>Acaricomes phytoseiuli</i>	Mite ( <i>Phytoseiulus persimilis</i> )	NR_042334
<b>ASV_16</b> <b>Delftia</b>	100% (427/427)	100+ sequences. E.g. <i>D. Tsuruhatensis</i> , <i>D. lacustris</i> , <i>D. siamensis</i>	E.g. Bollworm, Mosquito larvae, Louse, Fresh water, Activated sludge,	NA
<b>ASV_21</b> <b>Pseudomonas</b>	100% (427/427)	100+ sequences. E.g. <i>P. parafulva</i> , <i>P. fulva</i> , <i>P. punonensis</i> , <i>P. straminea</i>	E.g. Moth larvae, Soil, Marine algae, Rice paddy, Drosophila	NA
<b>ASV_43</b> <b>Staphylococcus</b>	100% (427/427)	100+ sequences. E.g. <i>S. sciuri</i> , <i>S. aureus</i>	E.g. Cow, Monitor lizard, Human, Chili, Fish, Banana, Bollworm	NA
<b>ASV_45</b> <b>Leucobacter</b>	100% (409/409)	29 sequences. E.g. <i>L. tardus</i>	E.g. Sea cucumber, Fish, Citrus tree, Moth, Human	NA
<b>ASV_65</b> <b>Enterobacter</b>	100% (427/427)	95 sequences. E.g. <i>E. asburiae</i> , <i>E. tabaci</i> , <i>E. cloacae</i> , <i>E. hormaechei</i>	E.g. Potato root, Moth larve, Fish, Honey bee, River water	NA
<b>ASV_68</b> <b>Enterobacter</b>	100% (427/427)	100+ sequences. E.g. <i>E. bugandensis</i> , <i>E. hormaechei</i> , <i>Klebsiella pneumoniae</i>	E.g. Soil, Rice paddy, Mosquito pupae, Bee	NA

<sup>†</sup> Percentage match (identical bases/alignment length)

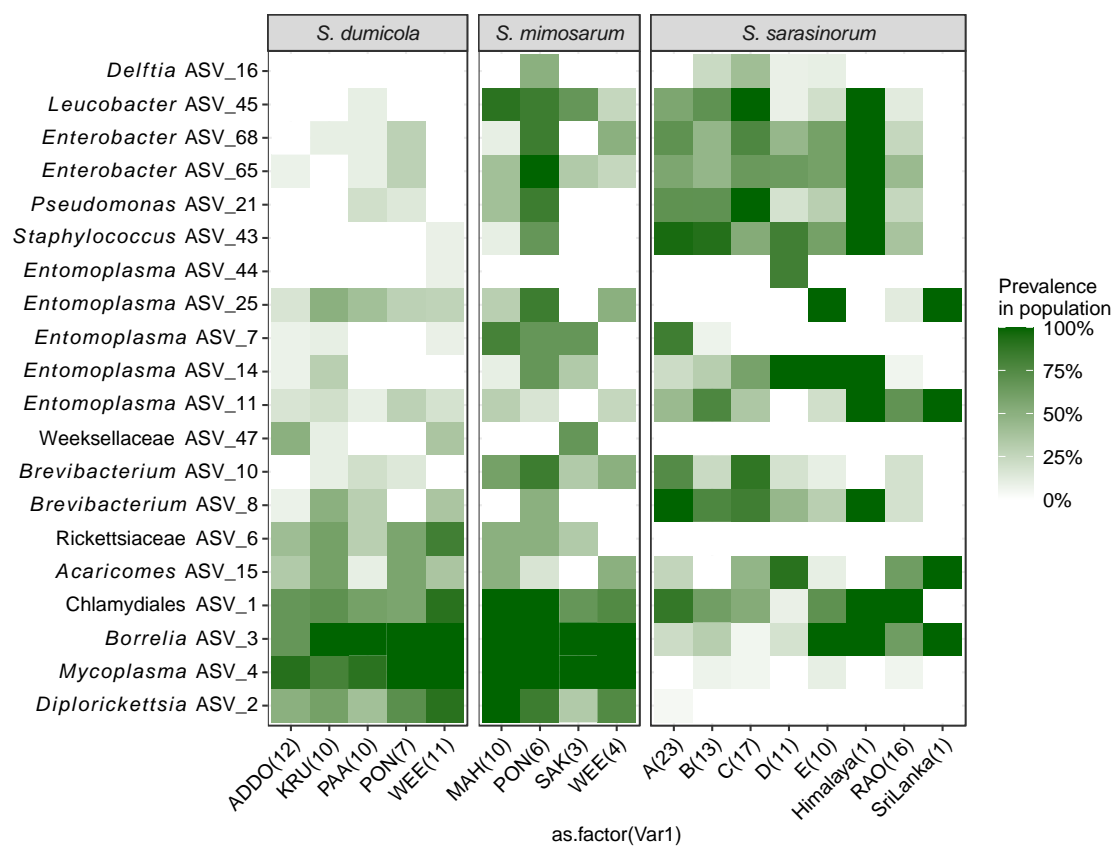


**Supplementary Figure S5.** Same as figure 2 in main text, but excluding all samples from sequencing run 4 (re-run and optimized samples).

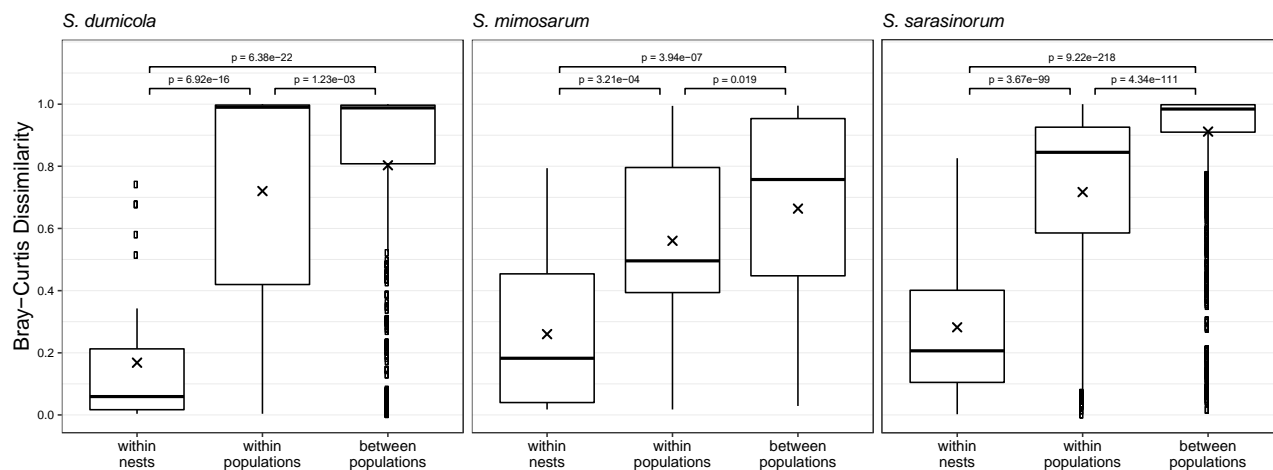


**Supplementary Figure S6.** Same as figure 3 in main text, but excluding all samples from sequencing run 4 (re-run and optimized samples).

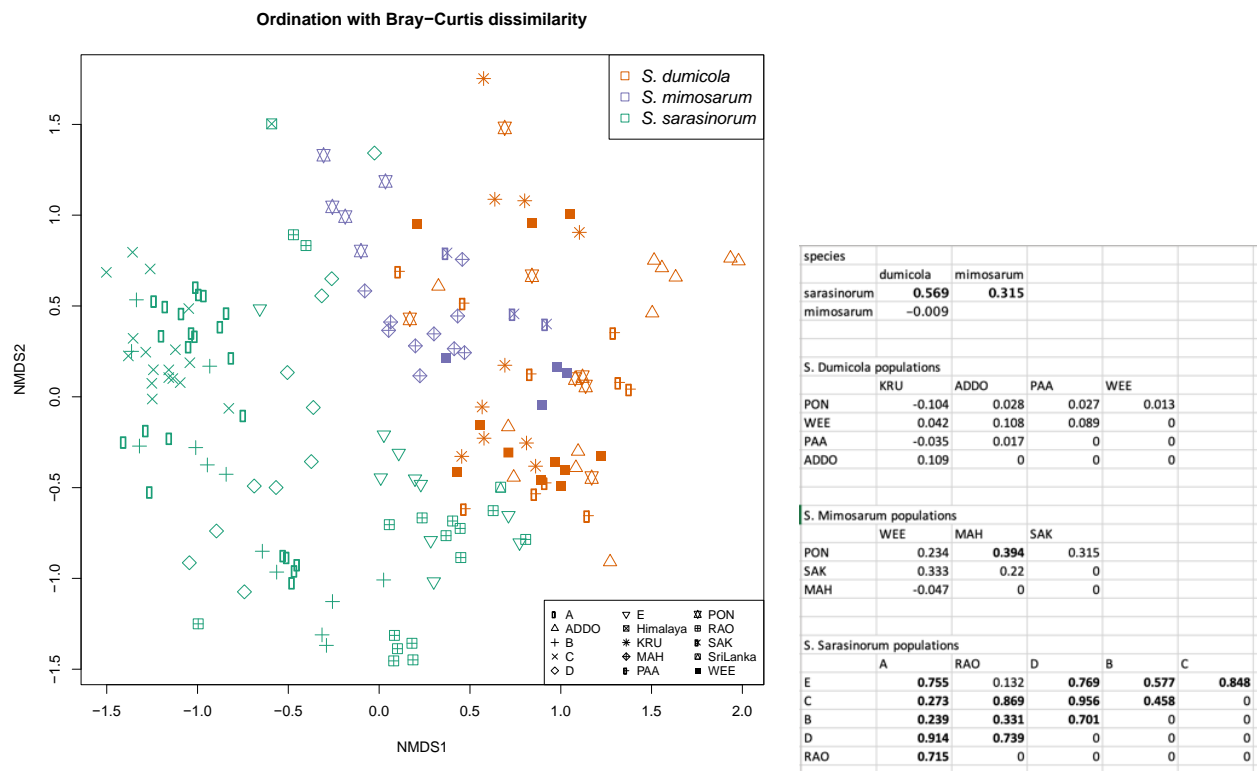




**Supplementary Figure S7.** Same as figure 4 in main text, but excluding all samples from sequencing run 4 (re-run and optimized samples).



**Supplementary Figure S8.** Same as figure 7 in main text, but excluding all samples from sequencing run 4 (re-run and optimized samples).



**Supplementary Figure S9.** Same as figure 8 in main text, but excluding all samples from sequencing run 4 (re-run and optimized samples).