

Genome	Complete, single copy	Fragment	Not found
Complete reference genomes			
wRi	180	2	39
wMel	180	2	39
wAu	181	2	38
wHa	179	3	39
wNo	181	4	36
Draft genome assemblies			
wRi_I14_18	182 ^a	2	37
wRi_I14_19	182	2	37
wRi_LZV15_057	181	3	37
wRi_LZV15_058	182	2	37
wRi_NMB15_030	182	2	37
wRi_USP16.124	182	2	37
wRi_USP16.125	182	2	37
wRi_Y14_29	183	2	36
wAna_Cebu	182	2	37
wAna_GB1	181	3	37
wAna_HNL0501	177	7	37
wAna_KMJ1	181	3	37
wAna_OGS_98K1	181	3	37
wAna_RC102	182	2	37
wAna_TBU136	182	3	36
wAna_VAV150	181	3	37
wAno_A29	181	3	37
wAno_CHC221	181	3	37
wPan_CHC1	182	2	37
wPan_CHG108	181	3	37

wPan_pl	181	3	37
wSuz	182	2	37
wSpc	182	2	37
wAur_NGN1	181	3	37
wAur_SP11-11	181	3	37
wTri	181	3	37

Table S1. Near-universal, single-copy proteobacteria genes, related to Figures 1 and 2 and Table 1. We used BUSCO v. 3.0.0 [S1] to identify complete and fragmented proteobacteria genes (out of 221) found in our *Wolbachia* genomes. For comparison, we provide results for reference-quality *Wolbachia* genomes.

^aThe anomalous result of finding two more of the proteobacteria target genes in most of our draft wRi genomes than in the wRi reference genome [S2] is an apparent artifact of BUSCO v. 3.0.0. In our draft wRi genomes, POG09087Z77 is generally found as complete and POG0908053X is generally found as fragmented. Both are present in the wRi reference with 100% identity but are not found by BUSCO v. 3.0.0.

Genes affected	Copy number change	Kolmogorov-Smirnov <i>P</i> -value	Affected genomes
<i>WD0631</i> , <i>WD0632</i>	2 → 1*	<0.0001	Both <i>w</i> Aur
<i>WD0631</i> , <i>WD0632</i>	2 → 4*	<0.0001	<i>w</i> Ana_HNL0501
<i>WRi_006710</i> , <i>WRi_006720</i>	1 → 0	<0.0001	Both <i>w</i> Aur, <i>w</i> Tri

Table S2. Copy-number variants in CI-causing loci, related to Table 2. All copy numbers are given relative to the *w*Ri reference [S2].

*This sequence is duplicated in the *w*Ri genome, so it was treated as diploid in our ControlFREEC v. 8.0 analysis [S3].

Species	Location	N	Infected	\hat{p} (confidence interval)
<i>D. ananassae</i>	Cairns, Australia	13	13	1.0 (0.75, 1.0)
	Townsville, Australia	1	1	1.0 (0.03, 1.0)
Total		14	14	1.0 (0.77, 1.0)
<i>D. anomalata</i>	Cairns, Australia	7	7	1.0 (0.59, 1.0)
	Townsville, Australia	1	1	1.0 (0.03, 1.0)
Total		8	8	1.0 (0.63, 1.0)
<i>D. auraria</i>	Japan	21	21	1.0 (0.84, 1.0)
<i>D. pandora</i>	Cairns, Australia	7	7	1.0 (0.59, 1.0)
<i>D. suzukii</i>	Haoping, Taibaishan, Shaanxi	15	8	0.53 (0.27, 0.79)
	Zhongyuanzi, Xinling, Badong, Hubei	1	1	1.0 (0.03, 1.0)
	Fushan Park, Qingdao, Shandong	3	3	1.0 (0.29, 1.0)
	Laoyan, Tiancheng, Wanzhou, Chongqing	2	2	1.0 (0.16, 1.0)
	Zhanggou, Gaoqiao, Emeishan, Sichuan	1	1	1.0 (0.03, 1.0)
	Dahebian, Banpo, Yixiang, Puer, Yunnan	28	26	0.93 (0.76, 0.99)
	Guanshan Nature Reserve, Yifeng, Jiangxi	30	25	0.83 (0.65, 0.94)
	Total		80	66
<i>D. subpulchrella</i>	Haoping, Taibaishan, Shaanxi	15	11	0.73 (0.45, 0.92)
	Fuhusi-Shanjuesi, Mt. Emei, Sichuan	16	8	0.50 (0.25, 0.75)
	Zhongyuanzi, Xinling, Badong, Hubei	8	3	0.38 (0.09, 0.76)
	Laoyingqing, Dahebian, Banpo, Yixiang, Puer, Yunnan	11	9	0.82 (0.48, 0.98)
Total		50	31	0.62 (0.47, 0.75)

Table S3. *Wolbachia* infection frequencies in natural populations, related to STAR Methods.

We provide data from *D. ananassae*, *D. anomalata*, *D. auraria*, *D. pandora*, *D. suzukii*, and *D. subpulchrella*. Sample sizes (N), estimated infection frequencies (\hat{p}) and exact 95% binomial confidence intervals for each collection and species.

Species	Infection status		Mean hatch rate (\pm s.d.)	<i>N</i>	<i>P</i>
	Female	Male			
<i>D. anomalata</i>	U	I	0.048 \pm 0.168	43	<0.001
	I	U	0.698 \pm 0.247	30	
<i>D. auraria</i>	U	I	0.344 \pm 0.184	17	<0.001
	I	U	0.899 \pm 0.093	18	
<i>D. pandora</i>	U	I	0.009 \pm 0.027	24	<0.001
	I	U	0.778 \pm 0.294	23	
<i>D. triauraria</i>	U	I	0.144 \pm 0.167	17	<0.001
	I	U	0.886 \pm 0.093	16	

Table S4. Cytoplasmic incompatibility assays, related to STAR Methods. We present data for *D. anomalata*, *D. auraria*, *D. pandora*, and *D. triauraria* infected with wRi-like *Wolbachia*. Data are mean egg hatch over five days for reciprocal crosses. *P* values are for one-tailed Mann-Whitney U tests.

U denotes *Wolbachia*-uninfected flies

I denotes *Wolbachia*-infected flies

N is the number of replicates

Model	Root age (95%)	wRi-clade age (95%)
Random tree shape	12009 (5193, 25183)	1588 (663, 3514)
Birth-death ($\rho = 0.1$)	13555 (5167, 27142)	1333 (425, 2856)
Birth-death ($\rho = 0.5$)	13985 (5737, 31378)	1304 (514, 2967)

Table S5. Age estimates for key nodes in Figure 1B. We show the effects of alternative node-age priors on estimated divergence times (and 95% posterior support intervals) for wRi-like *Wolbachia*.

	Model 1	Model 2	Model 3	Model 4
<i>D. suzukii</i>	-691419	-691420	-679434	-679434
<i>ananassae</i> , topology 1 ^a	-507233	-507220	-498620	-498608
<i>ananassae</i> , topology 2	-507229	-507215	-498628	-498612
<i>ananassae</i> , topology 3	-507234	-507219	-498630	-498613

Table S6. Model selection, related to Figure 2. Ln(marginal likelihood) for alternative models describing the joint evolution of mtDNA and *Wolbachia* within *D. suzukii* and three *ananassae* subgroup species (*cf.* Figure 2). See Methods for details of the calculations and descriptions of the models, which involve different partitions of the data (4 partitions, Models 1 and 2, versus 6 partitions, Models 3 and 4) and constant versus variable relative rates of mtDNA versus *Wolbachia* evolution (constant relative rates, Models 1 and 3, versus variable relative rates, Models 2 and 4). Preferred models indicated with bold.

The three topologies are the three alternative ways to resolve the trichotomy involving *D. pandora*, *D. anomalata* and three strains of *D. ananassae* shown in Figures 2C and 2D.

Branch	Topology 1 ^{a, b}	Topology 2 ^{a, b}	Topology 3 ^{a, b}
Suz ^c	(467, 566, 706)	(465, 573, 704)	(470, 579, 719)
Ana ¹	(205, 311, 491)	(295, 402, 551)	(341, 457, 616)
Ana ²	(52, 180, 653)	(53, 194, 736)	(55, 194, 658)
Ana ³	(340, 453, 609)	(335, 448, 608)	(293, 402, 558)
Ana ⁴	(410, 715, 1347)	(397, 706, 1340)	(413, 723, 1350)
Ana ⁵	(283, 396, 551)	(286, 398, 547)	(285, 393, 558)
Ana ⁶	(65, 116, 212)	(63, 113, 201)	(61, 113, 207)
Ana ⁷	(499, 853, 1535)	(506, 879, 1658)	(498, 872, 1598)
Ana ⁸	(171, 229, 316)	(168, 227, 308)	(171, 231, 315)
Ana ⁹	(126, 188, 284)	(129, 191, 285)	(129, 193, 285)
Ana ¹⁰	(846, 1470, 2716)	(876, 1471, 2720)	(848, 1480, 2725)
Ana ¹¹	(217, 412, 818)	(171, 334, 654)	(204, 394, 762)
Ano	(265, 693, 2177)	(275, 747, 2376)	(116, 399, 1462)
Pan	(301, 406, 560)	(200, 313, 489)	(203, 312, 492)

Table S7. Quartiles of the estimated ratio of mtDNA to *Wolbachia* substitutions, related to Figure 2. The rows correspond to the branches in Figures 2B and 2D. The alternative topologies refer to the alternative resolutions of the trichotomies displayed in these figures.

^aThe three alternative *D. sukii* topologies are the resolutions of the (Brazil, Italy, Italy) trichotomy (denoted as Suz¹, Suz² and Suz³ in Figure 2E).

^bThe three *D. ananassae* topologies are the resolutions of the (*pandora*, *anomalata*, 3X *ananassae*) trichotomy. Topology 1 is ((*anomalata*, *pandora*), *ananassae*); topology 2 is ((*pandora*, *ananassae*), *anomalata*); topology 3 is ((*anomalata*, *ananassae*), *pandora*).

^cThe favored model for *D. sukii* has all branches sharing the same ratio of rates.

Supplemental References

- [S1] Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva E.V., and Zdobnov E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212.
- [S2] Klasson, L., Westberg, J., Sapountzis, P., Näslund, K., Lutnaes, Y., Darby, A.C., Veneti, Z., Chen, L., Braig, H.R., and Garrett, R. (2009). The mosaic genome structure of the *Wolbachia* wRi strain infecting *Drosophila simulans*. *Proc. Natl. Acad. Sci. USA* 106, 5725–5730.
- [S3] Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappo, J., Schleiermacher, G., Janoueix-Lerosey, I., Delattre, O., and Barillot, E. (2012). Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 28, 423–425.