

Insertion sequence polymorphism and genomic rearrangements uncover hidden *Wolbachia* diversity in *Drosophila suzukii* and *D. subpulchrella*

Rupinder Kaur^{1,2*}, Stefanos Siozios^{3*}, Wolfgang J. Miller², Omar Rota-Stabelli^{1@}

Table S1. Primers sequences of Insertion sequence (IS) site polymorphism and Genomic rearrangement (GR) based markers

IS family	Annotation	Locus ID	Primer name	Primer sequences (5' => 3')	Coordinates in wRi genome	Estimated size in wRi (bps)	Estimated size in wSuz (bps)
IS5 family	ISWpi1	IS1	ISpol1F ISpol1R	TTCGCTACCCTGTTGTTCCC GGTTCAAGCTAAACACCGCG	385903-388478	2576	1600
	ISWpi1	IS2	ISpol2F ISpol2R	ACTGCGTAGACGTTCCCTGTG GTAGGGGCTGCAAGGCTTAA	765749-767702	1954	1035
	ISWpi1	IS3	ISpol3F ISpol3R	TCGTGCAACTCCTGGTCTTC GTGGGACGAGCAACAAGAGA	855201-856998	1798	879
	ISWpi1	IS4	ISpol4F ISpol4R	TTGAATGTGTGCCAAACCGC CGTTGCTCTTGGTTGGTGTG	929665-931529	1865	946
	ISWpi1	IS5	ISpol5F ISpol5R	ACGCCTTAAATGGGGCTTCA TTTAGTCGAAGCACGCCTGT	1041226-1043005	1780	861
	ISWpi1	IS6	ISpol6F ISpol6R	TGAACTGCCGCGTTCTAACT GTGGGTAAACTGGGTTTCGA	1166783-1168253	1471	552
	ISWpi1	IS7	ISpol7F ISpol7R	ACCGCATGCGTCTCTTCCCC TGCACACCGCTGCACCTTGC	582773-584203 / 1089232-1090662	1431	512
	ISWpi1	IS31 (wSuz-specific)	ISpol31F ISpol31R	GGGCTTTTTGCGAGTGTAGCTGT CATCTTGCAGCAGGAGTGTCGT	823636-824033	398	1292
IS66 family	ISWpi5	IS8	ISpol8F ISpol8R	ACCAGTATTGATGCTGCGCT TGCAAAGCACAGAAACCTTCG	679259-681204	1946	456

ISWpi5	IS9	ISpol9F ISpol9R	GCAGTTTCTGGTTCTGGTGC TAGTGCAGCATCTTGTGGCA	837678-839632	1955	330
ISWpi5	IS10	ISpol10F ISpol10R	CCGCTAGCCCCCTTATCATG CACACGCCCGCTATTCCTAT	936354-938508	2155	665
ISWpi5	IS11	ISpol11F ISpol11R	GATACTGCGGCGGTATGACA TAATCGCTATGCTCCTGCCG	1192806-1194739	1934	444
ISWpi5	IS12	ISpol12F ISpol12R	GGGCAGATTACCGATCAGGG TTAAGCAAGAGGTTCGCAGCA	1253149-1255132	1984	494
ISWpi5	IS13	ISpol13F ISpol13R	ACACTGGGATGACAAACGGT GGTGCTGCAATGCGTAAGTT	1258576-1260575	2000	510
ISWpi5	IS14	ISpol14F ISpol14R	AGCCAAGTTGTCACCTTCTGT GGTCACCAGCCAGGAAAAAG	317383-319382	2000	512
ISWpi5	IS15	ISpol15F ISpol15R	CTGACCCTGAAAGCACTGGT GAGGGCAATAGCAGAACCGA	866390-868283	1894	264
ISWpi5	IS16	ISpol16F ISpol16R	TGAGCACGATCCTGAAAAGGT AACGCACCTACGGTATAGCG	877110-879053	1944	317
ISWpi5	IS17	ISpol17F ISpol17R	CGCTCTGGTTGGCAAGAAAA CCCTCATCAATATCCACAGGAGT	1038164-1040941	2778	234
ISWpi5	IS18	ISpol18F ISpol18R	TGCTGCTCGCATAAATTCTGCCTGT GTGGCAACCGATGCGCACTT	980518-982570	2028	357
ISWpi5	IS19	ISpol19F ISpol19R	TCGCATCAGCTTCACCCTCCCT TCCTGTGTTTTTACGCGTGTTTCCA	198153-200132	1980	490
ISWpi5	IS20	ISpol20F ISpol20R	ACGTATTCGTGATGAACGGGCTGT ACAGCAAGATGCGAAGGAGAGACT	638019-639891	1873	383

ISWpi5	IS21	ISpol21F ISpol21R	CCCTTCTTGCTGCTCTTCCTTTGGC TTTGCTTGATACTGGAGCAACCGAT	231991-233923	1933	443
ISWpi5	IS22	ISpol22F ISpol22R	AGCCAGCAAGCACAAGGTGTCA GCATGGATGTTGCGATTCTCAAGGC	237160-238979	1820	330
ISWpi5	IS23	ISpol23F ISpol23R	TGCTTGGTAGGTAGATGTGTGCAAG AAGTGCAATAACGGTCATAACCGCA	256923-258881	1959	469
ISWpi5	IS24	ISpol24F ISpol24R	CAGCGGCAGCACTCCCAACA AGCCAACCACGGGGCTTCTT	308087-310060	1974	484
ISWpi5	IS25	ISpol25F ISpol25R	CCAGCGTCACGCGCTGGAAT GCTACCACGAGCGGCACTGA	560111-562072	1962	472
ISWpi5	IS26	ISpol26F ISpol26R	TGCCATGGTGC GGTTACTGTGG ACAGCAGGAGTTGCTGTGGGT	391590-393346	1757	270
ISWpi5	IS27	ISpol27F ISpol27R	TTAGCCAACCACGAGGCTTC GGATGTCATTCCAGTGCCCA	110357-112124	1768	278
ISWpi5	IS28	ISpol28F ISpol28R	CCTATAGCTATAACCGCCGCG GTTGCTGCTCTTTTTTCGGCA	1430759-1435759	4892	720
ISWpi5	IS29	ISpol29F ISpol29R	ACGGTGGGGGTTCTCTTGGCA ACCCTTTAGCACCTGTAATGCCCT	607466-609421 / 1113925-1115880	1959	466
ISWpi5	IS32 (wSuz-specific)	ISpol32F ISpol32R	TGGGTCTTGCTGCTGGACTGA ACATATCGCGCAGTCGGCAC	296839-297132	304	1794
IS110 family	ISWpi7	ISpol30F ISpol30R	TGATGGAGTGGAGCATGCAG TCGATGAAGCGGTCCAAACA	1404496-1406245	1750	260

Genomic Rearrangements*

Rearrangement 1	GR1	c65_F c65_R	GGGAGAGTAGGTCGCCGCCA AGCCAAGCGACGCAACAAACA	191445 – 191464 556008 – 556028	no amplification in wRi
Rearrangement 2	GR2	c151_F c151_R	CTGCATTGCGCTATAGAGTTTGA TGGCATAGATCGGAGGCTTAC	828612 – 828634 779789 – 779809	no amplification in wRi

* since wSuz genome is not closed, we could only validate with PCR one of the breakpoints for each rearrangement.

Table S2. 34 character-state matrix designed on IS- and GR-based *Wolbachia* strain-typing pattern

wRi	11111110111111111111111110111011100
wSpc	10000000000001000000010000000001
wSuz_CHN1	00000001000000000000000000000011
wSuz_CHN2	00000001000000000000000000000011
wSuz_JPN1	00000001000000000000000000000011
wSuz_JPN2	00000001000000000000000000000011
wSuz_AUT	000000010000000000000000200000011
wSuz_ITA1	000000010000000000000000200000011
wSuz_ITA2	000000010000000000000000200000011
wSuz_FRA	000000010000000000000000200000011
wSuz_GBR	000000010000000000000000200000011
wSuz_ESP	000000010000000000000000200000011
wSuz_SVN	000000010000000000000000200000011
wSuz_USA	00000001000000000000000000000013
wSuz_CAN	00000001000000000000000000000013

**Figure S1. Transposable element insertion at IS32 locus in *wSuzEU* compared to homologous gene *wRi_002820* in *wRi* genome.
L-TIR: Left Terminal inverted repeat**

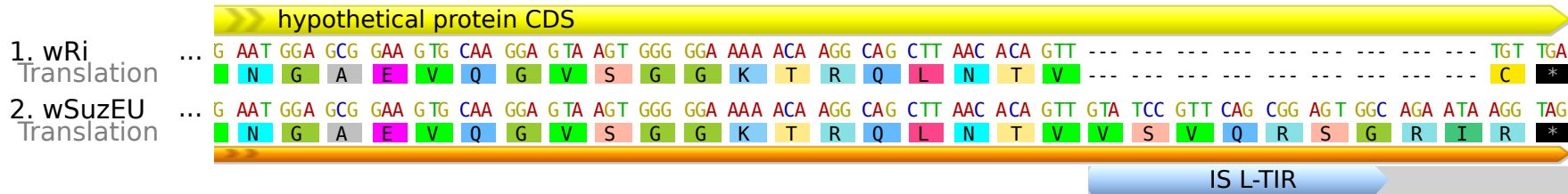


Figure S2. Full-length gel pictures.

Following gels correspond to IS and GR polymorphism from Figures 2A-2E and 3A, 3B in the main text. Lanes 1–10: Marker (bp) GeneRuler DNA Ladder Mix (Fermentas Life Sciences), *wRi*, *wSpc*, *wSuz_ITA*, *wSuz_FRA*, *wSuz_JPN*, *wSuz_CHN*, *wSuz_USA*, *wSuz_CAN* and STC-*Wolbachia* negative control.

