

Table S2. Short sequence composition of the non-coding DNA flanking RAYTs.

Organism	Name of RAYT homologue	Pos ^a	Most abundant 16-mer in non-coding DNA flanking RAYTs	Freq	p-Value ^b	Palindromes (all 16-mers within the palindrome occur at least twice within the genome)
<i>P. fluorescens</i> SBW25	<i>yafM</i>	5'	gggcaagcccgcctcac	241	2.00E-06	gcggggcaagcccgc
<i>P. fluorescens</i> SBW25	<i>yafM</i>	3'	gggcaagcccgcctcac	241	2.00E-06	gggctgcttcgcagccc
<i>P. fluorescens</i> SBW25	<i>pflu2165</i>	5'	gagggagcttgctccc	208	3.55E-06	gggagcttgctccc
<i>P. fluorescens</i> SBW25	<i>pflu2165</i>	3'	gagggagcttgctccc	208	3.55E-06	gggagcttgctccc
<i>P. syringae phaseolicola</i> 1448A	<i>pspph_4464</i>	3'	gcaagctcgtcccac	28	4.57E-04	gcgagcaagctcgc
<i>P. syringae phaseolicola</i> 1448A	<i>pspph_4464</i>	5'	acgatcgcactttgcc	1	1	none
<i>P. syringae syringae</i> B728a	<i>psyr_4421</i>	3'	tcgcgagcaagctcgc	88	4.41E-06	gcgagcaagctcgc
<i>P. syringae syringae</i> B728a	<i>psyr_4421</i>	5'	atgtgattgtgatctc	1	1	none
<i>P. syringae syringae</i> B728a	<i>psyr_4707</i>	3'	ttcgcgaacaagttcg	201	3.40E-07	gcgaacaagttcgc
<i>P. syringae syringae</i> B728a	<i>psyr_4707</i>	5'	gtgtcgttgcgcaatg	1	1	none
<i>P. syringae phaseolicola</i> 1448A	<i>pspph_5043</i>	3'	acggcgtgccactgcg	1	1	none
<i>P. syringae phaseolicola</i> 1448A	<i>pspph_5043</i>	5'	ggagcggacttgcccg	42	2.40E-04	gcggacttgcccg
<i>P. syringae tomato</i> DC3000	<i>pspto_0262</i>	3'	gcgtgccgctgcgcaa	3	0.004	none
<i>P. syringae tomato</i> DC3000	<i>pspto_0262</i>	5'	gagcggacttgctccgc	39	1.81E-04	gcggacttgctccgc
<i>P. fluorescens</i> SBW25	<i>pflu3939</i>	5'	gcaagccccctcccac	618	1.54E-07	gggggcaagccccc
<i>P. fluorescens</i> SBW25	<i>pflu3939</i>	3'	gtgggagggggttc	618	1.54E-07	gggggcttgccccc
<i>P. fluorescens</i> Pf-5	<i>pfl_3160</i>	5'	tcgccggcaagccggc	358	1.48E-07	gccggcaagccggc
<i>P. fluorescens</i> Pf-5	<i>pfl_3160</i>	3'	tcgccggcaagccggc	358	1.48E-07	gccggcaagccggc
<i>P. entomophila</i> L48	<i>pseen5170</i>	5'	gtaggagccagcttgc	95	1.62E-05	gccagcttgctggcg
<i>P. entomophila</i> L48	<i>pseen5170</i>	3'	aacactttatccacag	2	0.04	none
<i>T. sp HL-EbGR7</i>	<i>tgr7_2777</i>	3'	tcggcctgaaggccga	50	1.59E-05	gtcggcctgaaggccgac
<i>T. sp HL-EbGR7</i>	<i>tgr7_2777</i>	5'	tcgggctgaagccgga	80	6.01E-07	gtcgggctgaagcccgac
<i>P. putida</i> W619	<i>pputw619_5047</i>	5'	gatcgcggcaagccg	20	2.66E-05	cggcaagccg
<i>P. putida</i> W619	<i>pputw619_5047</i>	3'	tcgccggcaagccggc	229	1.79E-07	ggcaagccggcttgcc
<i>P. entomophila</i> L48	<i>pseen4846</i>	5'	caaggccgctcccaca	181	4.51E-06	gggccgctgtgaggccc
<i>P. entomophila</i> L48	<i>pseen4846</i>	3'	caaggccgctcccaca	181	4.51E-06	gcgacacaaggccgc
<i>P. putida</i> KT2440	<i>pp_0568</i>	5'	tgtgggagcggccttg	54	9.65E-06	gcggccttgctcgc
<i>P. putida</i> KT2440	<i>pp_0568</i>	3'	caaggccgctcccaca	54	9.65E-06	gcgacacaaggccgc

<i>P. putida</i> F1	<i>pput_0607</i>	5'	atgagggcgagccct	2	0.03	tgagggcgaa g ccctca
<i>P. putida</i> F1	<i>pput_0607</i>	3'	caagggcgcctcccaca	140	3.50E-06	cgctcccaca g gga cc g cg
<i>P. putida</i> W619	<i>pputw619_4597</i>	3'	gcggccttggtcgcg	148	1.07E-06	ggggctgcctt g ca g ccc
<i>P. putida</i> W619	<i>pputw619_4597</i>	5'	caagggcgcctcctaca	119	3.21E-06	ccgctcctac ag ggg
<i>P. putida</i> GB1	<i>pputgb1_0613</i>	5'	tcgcgacacaagggccg	235	1.71E-07	ggggcgccttt g cg g ccc
<i>P. putida</i> GB1	<i>pputgb1_0613</i>	3'	tcgcgacacaagggccg	235	1.71E-07	cgcgacaca ag ggcgcct ac aggg at cg cg
<i>P. putida</i> GB1	<i>pputgb1_5236</i>	5'	aaccgcctcccacag	62	6.32E-06	g cg gg tg a acc gc
<i>P. putida</i> GB1	<i>pputgb1_5236</i>	3'	tcgcgggtaaacccgc	90	3.07E-06	g cg gg ta a acc gc
<i>P. putida</i> KT2440	<i>pp_5176</i>	5'	agcccgcgaagagcc	26	3.20E-05	ctcttcg cg gg g ga g ccc gc gaag
<i>P. putida</i> KT2440	<i>pp_5176</i>	3'	cctgtgggagcggcg	86	5.25E-06	g cg gg g ct g ccc gc
<i>P. putida</i> F1	<i>pput_5083</i>	5'	cgggagcgcgcgaa	33	4.07E-05	ggcctcttcg g ggcgagccc gc gaag agcc
<i>P. putida</i> F1	<i>pput_5083</i>	3'	gcccgctcccacaggg	70	1.19E-05	g cg gg cat gccc gc
<i>P. putida</i> GB1	<i>pputgb1_1364</i>	5'	gccgcccgcgcggcg	35	1.33E-05	ag cg cgcc gc gc g gg cg ct
<i>P. putida</i> GB1	<i>pputgb1_1364</i>	3'	gccgcccgcgcggcg	35	1.33E-05	g cg ccc gc gc g gg cg gc
<i>P. entomophila</i> L48	<i>pseen3227</i>	5'	gcggatccatccgcga	151	6.50E-06	g cg gat cc atcc gc
<i>P. entomophila</i> L48	<i>pseen3227</i>	3'	tcgcggatgaatccgc	151	6.50E-06	g cg gat ga atcc gc
<i>P. putida</i> F1	<i>pput_3919</i>	5'	cgggtttaccgcgaa	404	1.75E-07	g cg gg tt t acc gc
<i>P. putida</i> F1	<i>pput_3919</i>	3'	cctcaccagcgcg	2	0.03	None
<i>P. mendocina</i> ymp	<i>pmen_3135</i>	5'	ggtgcgcacggcgac	198	2.08E-07	g gt g cg ca cg g cg ca cc
<i>P. mendocina</i> ymp	<i>pmen_3135</i>	3'	ggtgcgcacggcgac	198	2.08E-07	g gt g cg ca cg g cg ca cc
<i>T. sp. HL-EbGR7</i>	<i>tgr7_1317</i>	5'	gtaggatggcгаааgс	14	1.43E-04	atgggcaa ag cgata g cg tg ccc at
<i>T. sp. HL-EbGR7</i>	<i>tgr7_1317</i>	3'	gtaggatggcгаааgс	14	1.43E-04	atgggcaa ag caa cg cg tg ccc at
<i>N. punctiforme</i> PCC 73102	<i>npun_f5543</i>	5'	gaggaaacgaaacccaa	13	4.39E-04	g tt g g g tt gagga ac gaa acc ca ac
<i>N. punctiforme</i> PCC 73102	<i>npun_f5543</i>	3'	atgttgggtttcgttc	13	4.39E-04	g tt g g g tt tcg tt cct ca acc ca ac
<i>P. mendocina</i> ymp	<i>pmen_0731</i>	5'	cggattgcatccgggc	93	3.12E-06	ccc g gatt gc atcc gg g
<i>P. mendocina</i> ymp	<i>pmen_0731</i>	3'	cggattgcatccgggc	93	3.12E-06	ccc g gatt gc atcc gg g
<i>P. stutzeri</i> A1501	<i>pst_1052</i>	5'	attagccgaagcgta	4	0.0029	t g gatt ag ccgaa g cg ta at cc g
<i>P. stutzeri</i> A1501	<i>pst_1052</i>	3'	aaacgacggaagcgcc	2	0.03	None
<i>S. enterica</i> serovar Paratyphi A AKU 12601	<i>sspa4070</i>	5'	cgcttaccgggcctac	18	7.57E-06	gccc g g tg cgctt cg ctt acc gg gc
<i>S. enterica</i> serovar Paratyphi A AKU 12601	<i>sspa4070</i>	3'	gtaggccgataaggc	57	2.23E-07	ag g cc g gata ag g cg t
<i>E. coli</i> K-12 DH10B	<i>yafM</i>	5'	tgctgatgacgct	77	2.26E-06	g cc t g at g ca cg ctg gc g ct ct at cat gc
<i>E. coli</i> K-12 DH10B	<i>yafM</i>	3'	gtaggccgataaggc	106	2.26E-07	ag g cc g gata ag g cg t
<i>P. aeruginosa</i> PAO1	<i>pa1154</i>	5'	gcgttattcgccctac	30	1.02E-06	g cg ttatt cg c
<i>P. aeruginosa</i> PAO1	<i>pa1154</i>	3'	gcgttattcgccctac	30	1.02E-06	g cg ttatt cg c
<i>P. aeruginosa</i> PA7	<i>pspa7_4226</i>	5'	gtaggcggaataacgc	7	3.15E-04	g cg gaata ac gc

<i>P. aeruginosa</i> PA7	<i>pspa7_4226</i>	3'	gtagggcgaataacgc	7	3.15E-04	g cg aataa cg c
<i>P. aeruginosa</i> LESB58	<i>pales_41671</i>	5'	gtagggcgaataacgc	26	8.11E-07	g cg aataa cg c
<i>P. aeruginosa</i> LESB58	<i>pales_41671</i>	3'	gtagggcgaataacgc	26	8.11E-07	g cg aataa cg c

All homologues are flanked by at least one 16-mer that is unusually over-represented within the respective genome of the bacterium. In all cases the 16-mer contains or is part of a palindrome or inverted repeat. Letters in red denote complementary base pairs. ^aDenotes whether the 16-mer was found in the extragenic space flanking the RAYT on the 5' or 3' side. ^bProportion of different 16-mers that occur equally or more often than the most abundant 16-mer from the non-coding DNA flanking the RAYT homologue.