Multiple evolutionary trajectories have led to the emergence of races in Fusarium oxysporum f. sp. lycopersici

Chellappan Biju, V.; Fokkens, L.; Houterman, P.M.; Rep, M.; Cornelissen, B.J.C.

DOI
10.1128/AEM.02548-16

Publication date
2017

Document Version
Other version

Published in
Applied and Environmental Microbiology

Supplemental material for "Multiple evolutionary trajectories have led to the emergence of races in *Fusarium oxysporum* f. sp. *lycopersici*.", Biju V.C. et al.

Supplemental material and methods

Details of the assembly of clone 9G3

Sequencing of BAC clone 9G3 and *de novo* assembly resulted in three scaffolds, notably 47 (approximately 35 kb in length), 56 (31 kb) and 53 (22 kb), hence in a contiguous sequence with two gaps (Supplemental figure 2A). The three scaffolds could easily be ordered because of the presence of BAC vector sequences at one end of both scaffold 47 and 53. Using primer pairs 4241/4242 and 4239/4240 corresponding to sequences flanking the two gaps in the 9G3 sequence (Supplemental figure 2A and Supplemental table 1), fragments could be PCR-amplified (Supplemental figure 2B) and sequenced.

The gap between scaffolds 56 and 53 appeared to be a sequence of 6122 bp representing a *Helitron* (*HelB*). The gap between scaffolds 56 and 47 was found not to be a real gap but rather the result of a mis-assembly, most likely due to the presence of another *Helitron* copy (*HelA*) at the end of scaffold 47. Comparing the 9G3 gap closed sequence with the genome sequence of the Fol reference strain (Fol4287) revealed that it fully aligns with a genomic region in the lineage specific (LS) chromosome 14, namely with supercontig (sc) 2.22: 651200 to 712754, except for a unique fragment containing *AVR1* (Supplemental figure 2A).

Sc2.22 is composed of a large number of contigs of which four (partly) align with 9G3, namely contigs 852, 853, 854 and 855 (Supplemental figure 2A). These contigs are separated by sequence gaps and the availability of 9G3 allowed closing of these gaps. Using primer pair 4618/4619 (Supplemental table 1) a 6.4 kb fragment could be amplified from both Fol4287 genomic DNA and 9G3 DNA (data not shown). Sequencing confirmed the identity of the amplified fragments and let us conclude that the gap between contigs 852 and 853 was the result of a mis-assembly due to the presence in this region of an *NHT2*-like retrotransposon. Similarly, PCR analysis using primer pair 4620/4621 (Supplemental table 1) and sequencing indicated that the gap between contigs 853 and 854 can be explained by a mis-assembly due to the presence of a *Fot5* DNA transposon (data not shown).

The third gap (between contigs 854 and 855) appeared to be located within a copy of transposable element *Yaret2* that precedes a second *Yaret2* copy on contig 855, suggesting the presence of two *Yaret2* copies in a row. However, in 9G3 only one copy was identified.

Using primer set 4625/4630 (Supplemental table 1) a 3.7 kb fragment could be amplified from genomic DNA of both Fol4287 and Fol004 as well as from and 9G3 DNA (data not shown). Sequence analysis confirmed the presence of two *Yaret2* copies in tandem sharing the LTR that separates the open reading frame of the two copies. This suggests that during assembly of the 9G3 insert sequence one copy was missed due to a high level of sequence similarity. The length of the full 9G3 insert was found to be 98694 nucleotides in total.

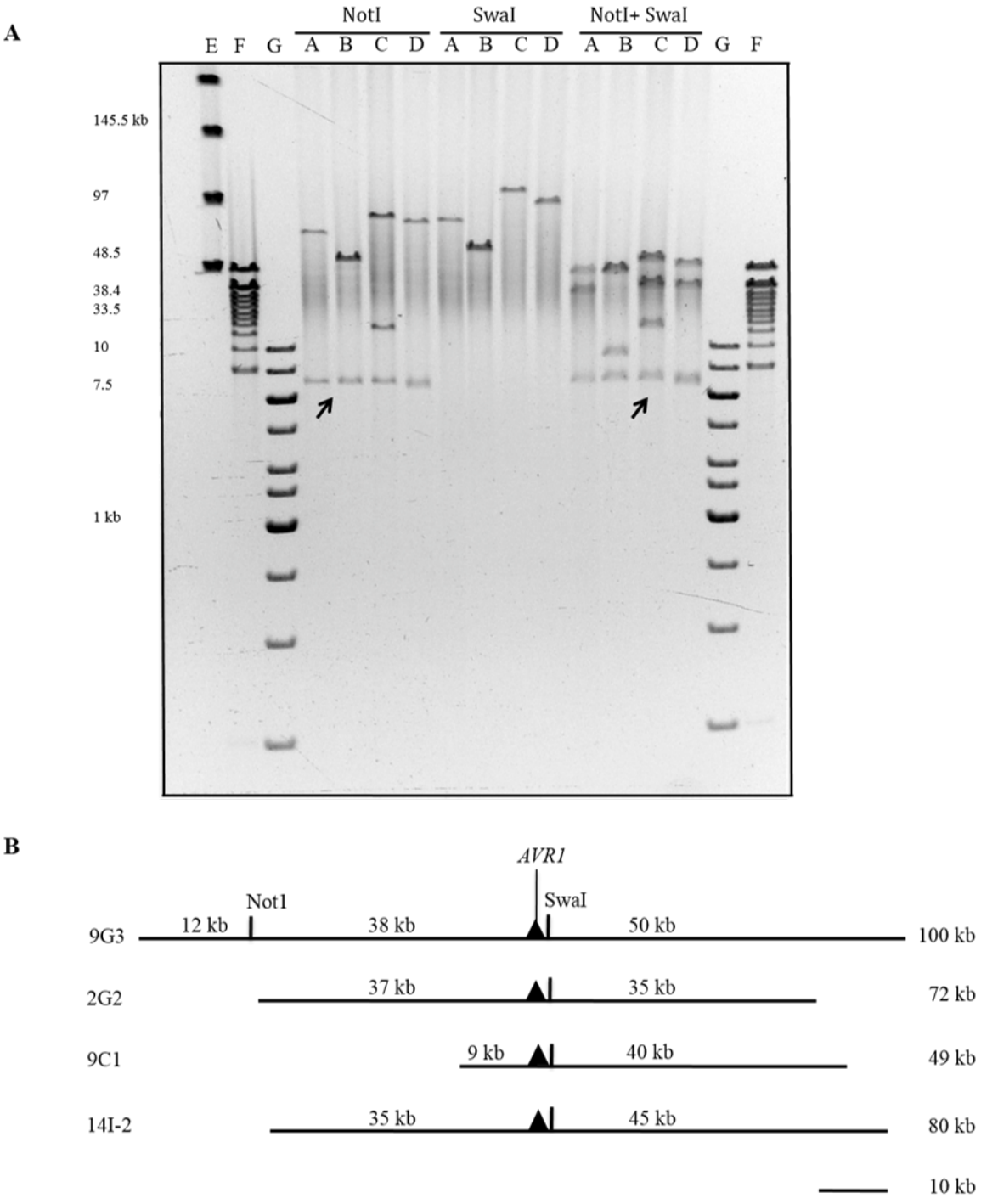

Figure S1. Analysis of BAC clones with an insert containing the *AVR1* locus

(A) DNA of BAC clones 2G2, 9C1, 9G3 and 14I-2, selected from a Fol004-BAC library, was digested with either NotI or SwaI alone or with both enzymes. NotI was chosen to cut out the insert from the BAC vector (the vector pBleoBAC11 contains two NotI sites in the region flanking the insert). SwaI was chosen to estimate the approximate location of *AVR1* in the insert (The analysis of 2.8 kb *AVR1* genomic region revealed a SwaI site 294 bp downstream of *AVR1* stop codon; the vector pBleoBAC11 does not carry a SwaI site). DNA fragments were separated on a 1% agarose CHEF gel at 5- to 15-s linear ramp time, 6 V/cm, 14°C in 0.5× TBE buffer for 18 h, and stained with ethidium bromide. The 7.5 kb band present in the NotI digests (indicated by arrows) corresponds to the cloning vector. A: clone 2G2, B: clone 9C1, C: clone 9G3, D: clone 14I-2, E: Lamda ladder, F: 8-48 kb ladder, G: GeneRuler 1 kb DNA Ladder (0.25 – 10 kb). (B) Schematic representation of the relative positions of the BAC inserts containing *AVR1*. The insert sizes and the positions of NotI and SwaI sites are deduced from the restriction analysis. The position of *AVR1* was inferred from the position of the SwaI site.

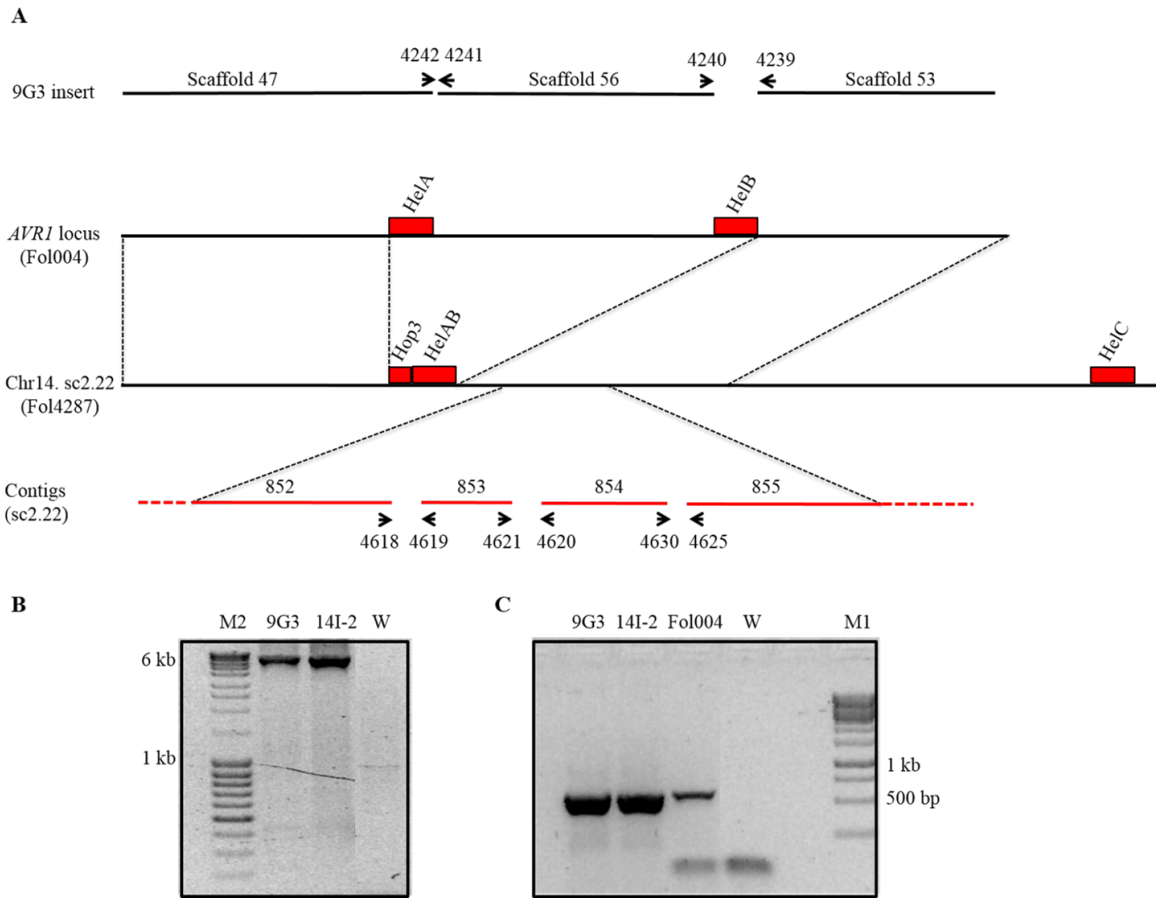

Figure S2. Assembly of the 9G3 insert.
(A) Schematic representation of 9G3 insert mapped to a genomic region in supercontig (sc) 2.22 of chromosome 14 of Fol4287. The order of three scaffolds corresponding to the sequence of the 9G3 insert is shown above. Arrows indicate the location of the primers used to fill the gaps between the scaffolds. Red lines indicate the contigs of sc2.22 corresponding to this genomic region. Arrows indicate the location of the primers used to fill the gaps between the contigs. (B) Left: amplification of a PCR product of 6.2 kb using the primer pair 4240/4239. Right: amplification of a PCR product of 600 bp using the primer pair 4242/4241. W: water control, M1: GeneRuler 1 kb DNA Ladder (0.25 - 10 kb), M2: MassRuler DNA Ladder Mix (0.08 - 10 kb), W: Water control.

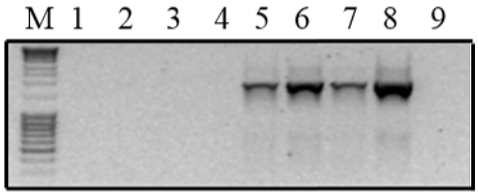

Figure S3. **Confirmation of the absence of *Helitron B* (*HelB*) in some race 1 isolates**.
PCR experiment showing the presence/absence of a 1.8 kb fragment in the genomic DNA of different race 1 isolates (lanes 1-8) using primer pair 4345/4340 (these primers correspond to the 5' and 3' flanking sequences of *HelB*, respectively). Lane 1: Fol001, lane 2: Fol003, lane 3: Fol004, lane 4: Fol006, lane 5: Fol009, lane 6: Fol010, lane 7: Fol011, lane 8: Fol016, lane 9: water control, M: MassRuler™ Low Range DNA Ladder (80-1031 bp).

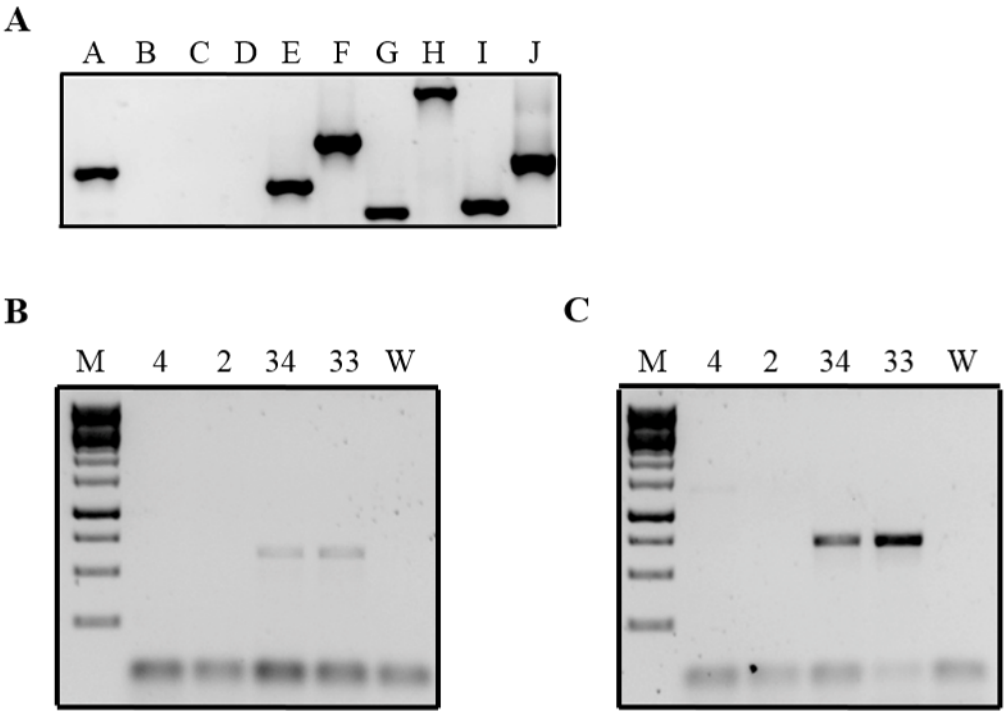

Figure S4. Fol4287 and Fol033 contain a *Hop3* insertion at the 5' end of *HelAB*.
(A) Amplification of PCR fragments from the genomic DNA of Fol4287 with primer pairs A-J. The absence of a PCR product with primer pair B indicates a *Hop3* insertion at the 5' end of *HelAB*. (B) A product of 624 bp was amplified only from Fol4287 and Fol033 using primer pair 4304/4298. (C) A product of 723 bp was amplified only from Fol4287 and Fol033 using primer pair 4303/4297. M: Marker, 1: Fol004, 2: Fol002, 3: Fol4287, 4: Fol033, W: water control.

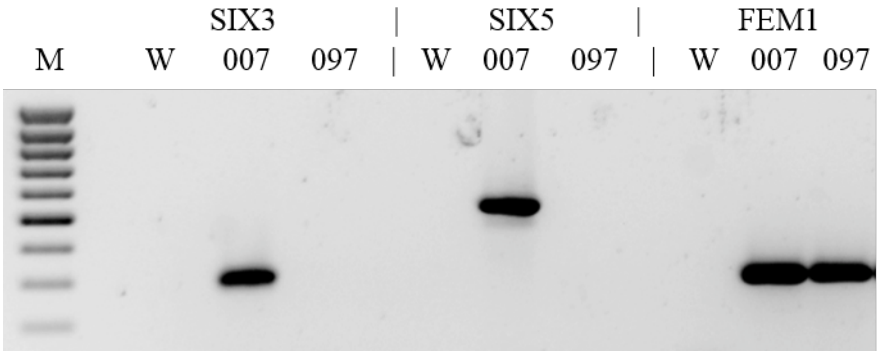

Figure S5: Six5 is absent in Fol097.
We used PCR to determine whether, in addition to *AVR2*, Six5 is absent from Fol097. The left three lanes show presence and absence of amplicons obtained using primer pair 962+963 (corresponding to a 273 bp fragment) in water (W, negative control), Fol007 and Fol097. We confirmed the absence of *AVR2* (SIX3) in Fol097. The middle three lanes show presence and absence of amplicons obtained using primer pair 1993+1994 (corresponding to a 524 bp fragment) in water, Fol007 and Fol097. We show the absence of Six5 in Fol097. The right three lanes are our positive control: primer pairs 157+158 correspond to a 274 bp region in the Fem1 promoter. Primers are listed in Suplemental table 1 and we used MassRuler DNA Ladder Mix (ThermoFisher) as a marker for fragment sizes.

Table S1. Primers used in this study

Number	Sequence	Targeting gene/genomic position in 4287
4539	AAGCGAGAGAAAACGGAAGC	5956 bp upstream HelA
4540	AATGTTTAGGACGGCAATACC	5158 bp upstream HelA
4298	TGAAGCACAAGTAGCTGAGG	316 bp upstream HelA
4297	TGCCTCTTTGCTCTGAAGG	Specific to 5' end of Helitrons
4242	ACAAGTCACAAAGCATCAC	Specific to 3' end of Helitrons
4241	TTGACGACACGTTCAACATC	201 bp downstream HelA
4345	TAGCTGGCGCATTTGATAG	429 bp upstream HelB
4340	ACAACGGGGACATTGATGGC	1382 bp downstream HelB
4355	AAGAGTGGTTAAGCGACTTC	2333 bp downstream HelB
4354	TGGGTACTGCATGGTAATG	3369 bp downstream HelB
4309	TTGTGGAGGCAGCCGTTTGG	14369 bp downstream HelB
4306	TGTCATACATTGAGGATGG	14772 bp downstream HelB
4395	TCCTCACAATGCTGACATCG	1550 bp upstream HelC
4371	AGCGTGGACTTGAGTTCTGC	38 bp downstream HelC
4541	TGTGCTCAGCCACATCAGC	824 bp downstream HelC
4542	TCGGTCGCAATCAAAGCACC	1545 bp downstream HelC
4866	TAAGGTCTCAATGGTCCTTCG	Chr14: SC22:97075 -97096
4867	TTGCGGACGAAGTTATCAG	Chr14: SC22:97606 -97625
4890	TGATAATGTCGTTGCAAGAC	Chr14: SC22:98193 -98213
4891	ACTGGATTTCTCGTCGACTGC	Chr14: SC22:99021 -99041
4967	TCAGAAGACTTGTTATTTGC	Chr14: SC22:106083 -106103
4968	TTGCCCACTGGAAGACTGC	Chr14: SC22:106420 -106439
5023	AATCCATCTTCCAGAATCAAC	Chr14: SC22:219341 -219362
5024	TTATCAGACATACTCGATTC	Chr14: SC22:219847 -219867
5025	TCAGACAACAAGCACATTC	Chr14: SC22:229717 -229736
5026	ATGGCCGTTACTGTCGAGAC	Chr14: SC22:230051 -230071
4874	ATCTTGGCTGGTGGCGAAACG	Chr14: SC22:243580 -243601
4875	TTTATTAGCGAAGGGAGAGC	Chr14: SC22:243869 -243889
4749	TGTATTCAAGGCTCACATCGG	Chr14: SC22:532710 -:532731
4750	TGCTACATATGCCCTAGTACG	Chr14: SC22:533499 -533519
4880	AGGCGTCGCTTAGGTATTGG	Chr14: SC22:556590 -556610
4881	TATCGTCGTAATAGAAGTCC	Chr14: SC22:557050 -557070
4751	TGGTATCATATGCCATTACGG	Chr14: SC22:561892 -561913
4752	TTATTAAGAGCTGAGAACCG	Chr14: SC22:562803 -562823
4618	TGCCCAATTTCACTTACACAG	Chr14: SC22:692515 -692536
4732	TCTCGCTCTCGTAGATAGCC	Chr14: SC22:693395 -693415
4733	AACCACGCATCCTCTGTGAC	Chr14: SC22:697926 -697946
4626	TTCATCCTCTGTATAAGACACG	Chr14: SC22:69813 -698533
4309	TTGTGGAGGCAGCCGTTTGG	Chr14: SC22:703559 -703579
4306	TGTCATACATTGAGGATGG	Chr14: SC22:703941 -703961
4616	TTCGGTTGGAATCGATCCAG	Chr14: SC22:711336 -711356
4617	TCATTCACTCTCTCGTGTCC	Chr14: SC22:711679-711699
4395	TCCTCACAATGCTGACATCG	1550 bp upstream HelC
4297	TGCCTCTTTGCTCTGAAGG	Specific to 5' end of Helitrons
1091	TCAGGCTTCACTTAGCATAC	*AVR1* ORF
4239	TGTTGCATACAGACAGCTGAG	19278 bp downstream *AVR1* stop codon
4240	AATCAGGAACTCACGCTTCG	12747 bp downstream *AVR1* stop codon
4619	ATCATACACGTTAGCTCAATTC	28921 bp downstream *AVR1* stop codon
4620	ACATAGCCATCCACTCATCC	31210 bp downstream *AVR1* stop codon
4621	TGACTATGAATTGAGCTAACG	28915 bp downstream *AVR1* stop codon
4625	ACCGTGGTACTGTCATACATTG	33849 bp downstream *AVR1* stop codon
4630	TATGGACAATACAGAGACG	34853 bp downstream *AVR1* stop codon
4303	ACTTCCCAGTGACAAACGC	Specific to 5’ end of Hop3
4304	TACTCGAACGATAAACTGG	Specific to 5’ end of Hop3
1033	GCCGACCGAAAAACCCTAA	*AVR1* ORF
2934	CCAGCCAGAAGGCCAGTTT	*AVR2* ORF
964	GGCAATTAACCACTCTGCC	*AVR2* ORF
1002	TATCCCTCCGGATTTTGAGC	*AVR3* ORF
363	AATAGAGCCTGCAAAGCATG	*AVR3* ORF
1416	GGAAGTACCAGTGATCATGTT	*EF-1α* ORF
889	TCGTCGTCATCGGCCACGTC	*EF-1a* ORF
1723	CGATGCCTTGACCGAAAGTT	*AVR2*-upstream
1236	AGTGGTAAATGTTTAGGCAAG	*AVR2*-upstream
1237	TTCTGTGGCAGTTCCCCTT	*AVR2*-downstream
1238	GGTGTGTTGAACAGGTGCT	*AVR2*-downstream

157	ATGAAGTACACTCTCGCTACC	*FEM1* ORF
158	GGTGAAAGTGAAAGAGTCACC	*FEM1* ORF
962	TGAGCGGGCTGGCAATTC	*AVR2* ORF
963	CAATCCTCTGAGATAGTAAG	*AVR2* ORF
1993	GCGCTTCGAGTACATCTCTG	*SIX5 ORF*
1994	CTAGGCCGCATCACAATAGA	*SIX5 ORF*

Table S2. Transposable elements on the insert of 9G3

Classification								
Order	**Superfamily**	**Family**	**Designation**	**Size**	**LTR**	**TIR**	**TSD**	**Coordinates**
ClassI								
LTR	Gypsy\Ty3		MAGGY-like	5717	152			44144 - 49860
	Copia\Ty1		NHT2-like	5610	176		8	75725 - 81334
			NHT2-like fragment	211				14114 - 4324
	Unclassified		Yaret2	4840	195			86539 - 91378
			Yaret2	4840	195			91184 - 96083
	Solo-LTR		MAGGY-like solo-LTR	152				74512 - 74663
			Skippy-solo-LTR	431				59963 - 60393
LINE			MGR583 like LiNE element	5353				8403 - 13755
			MGR583-like LINE element	665				17893 - 18557
			MGR583-like LINE element	86				6588 - 6673
			MGR583-like LINE element fragment	59				6944 - 7002
SINE			Foxy fragment	156				7778 - 7933
Unrelated			Marsu	2328				20822 - 23149
Class II								
Subclass 1								
TIR	TC1\mariner	Pogo	Fot5	1869		42	2	53449 - 55317
			Fot5	1860		44	2	81947 - 83806
			Fot3-partial	247				75478 - 75724
			Fot3-partial	603				81337 - 81939
		hAT	hAT-1	3093		11	8	55469 - 58559
			Tfo1-partial	1644				49861 - 51504
			Hormin	759		15	8	26669 - 27427
			YahAT7 fragment	71				28744 - 28814
			YahAT7 fragment	528				34923 - 35450
		Mutator	Hop6 fragment	2038				24631 - 26668
			Hop6 fragment	1308				27436 - 28743
		MITE	mimp3	215		27	2	24102 - 24316
			mimp1	222		27	2	4879 - 5101
			mimp1	223		27	2	5546 - 5767
			mimp4-partial	86				51652 - 51737
Class II								
Subclass 2								
			Helitron	6108				28815 - 34922
			Helitron	6123				65981 - 72103
Class II								
Unclassified								
			Unclassified	233		26	2	3885 - 4120
			Unclassified	859		21		19969 - 20827
Total size of the TEs				58537				

Table S3. Non-transposable ORFs on the insert of 9G3

ORF	Size	Position in 9G3	Homologous gene	Position in strain 4287	Remarks
ORF1	2727	1140 - 3866	*FOXG_14233*	Chr14: Supercontig 22: 652339-655065 -	
ORF2	2600	15079 - 17678	*FOXG_14234*	Chr14: Supercontig 22: 666276-668875 +	Bifunctional catalase-
ORF3	6114	36089 - 42202	*FOXG_06805* *FOXG_07365* *FOXG_16388*	Chr03: Supercontig 7: 1924520-1931400 Chr06: Supercontig 9: 2323584-2330452 - Unpositioned: Supercontig 34: 451087-457955 -	Encode protein with Helicase domain and Rep domain
AVR1	793	52222 - 53014	No homologue		Avirulence gene corresponding to R gene *I* or *I-1*
ORF4	828	62898 - 63725	*FOXG_22916* *FOXG_14128* *FOXG_12409*	Unpositioned: Supercontig 68: 1286-2722 - Chr14: Supercont2.51:71494-72306 – Chr03: Supercontig 18: 68914-70292 -	Unknown
ORF5	2189	72260 - 74448	*FOXG_14236*	Chr14: Supercontig 22:689228-691416 +	Highly similar to secreted oxidoreductase *ORX1*
ORF6	942	74643 - 75584	*FOXG_14237*	Chr14: Supercontig 22: 691611-692552 +	recQ family helicase
ORF7	1136	85183 - 86318	*FOXG_14238*	Chr14: Supercontig 22: 702268-703400 +	recQ family helicase
ORF8	477	97014 - 97490	*FOXG_14240*	Chr14: Supercontig 22: 711073-711549 -	Unknown