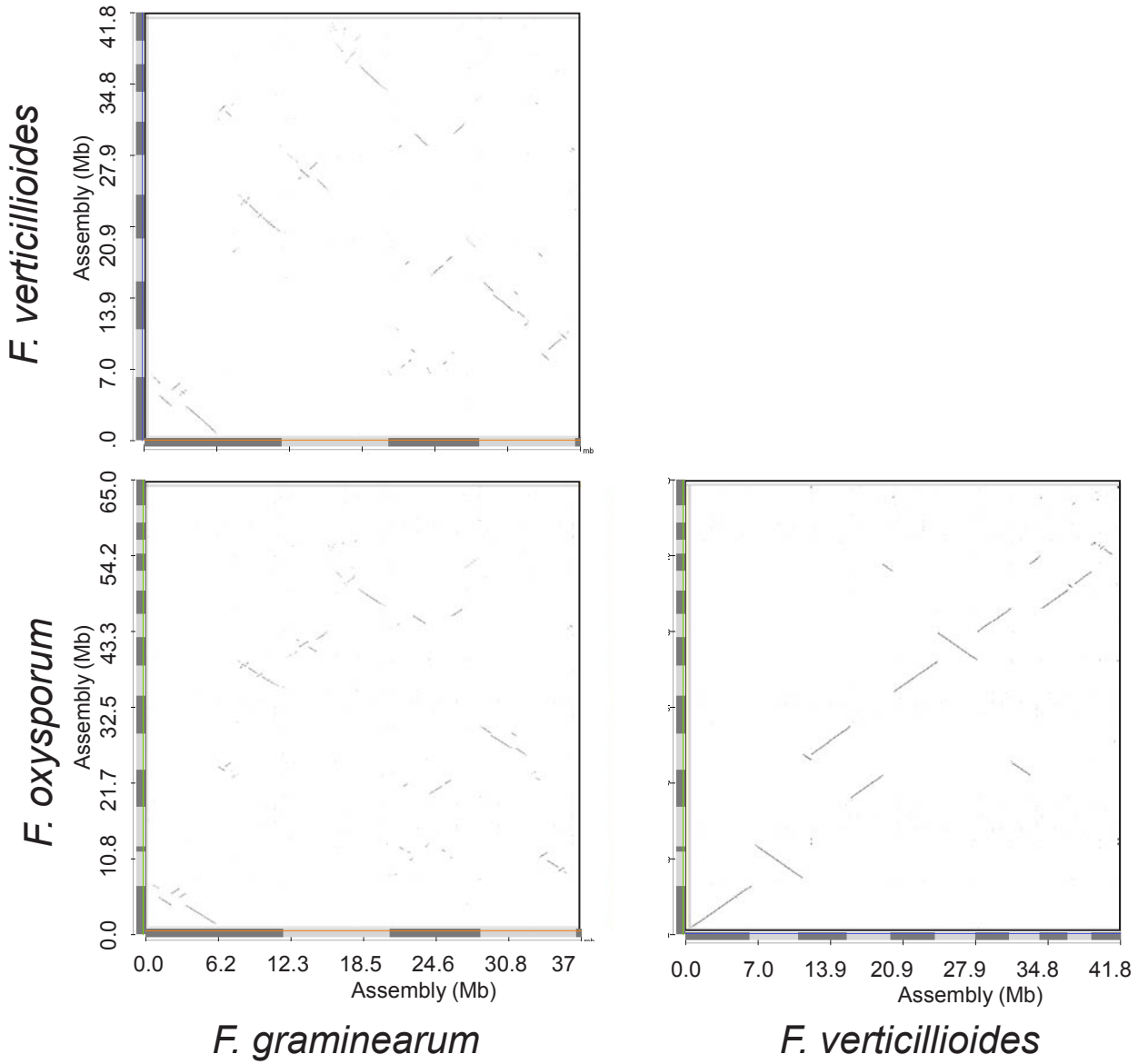

Figure S1. Pair-wise dotplots among three *Fusarium* genomes. The x and y axes represent assembled genomes. The altered gray and light grey bars are defined chromosomes in each assembly (See detailed mapping information at: http://www.broadinstitute.org/annotation/genome/fusarium_group/maps/Index.html). Comparisons are based on blastn alignments of genomic sequence (cutoff 1e-10) using *Fg* (against *Fv* and *Fo*) and *Fv* (against *Fo*) genomic DNA as query sequences.

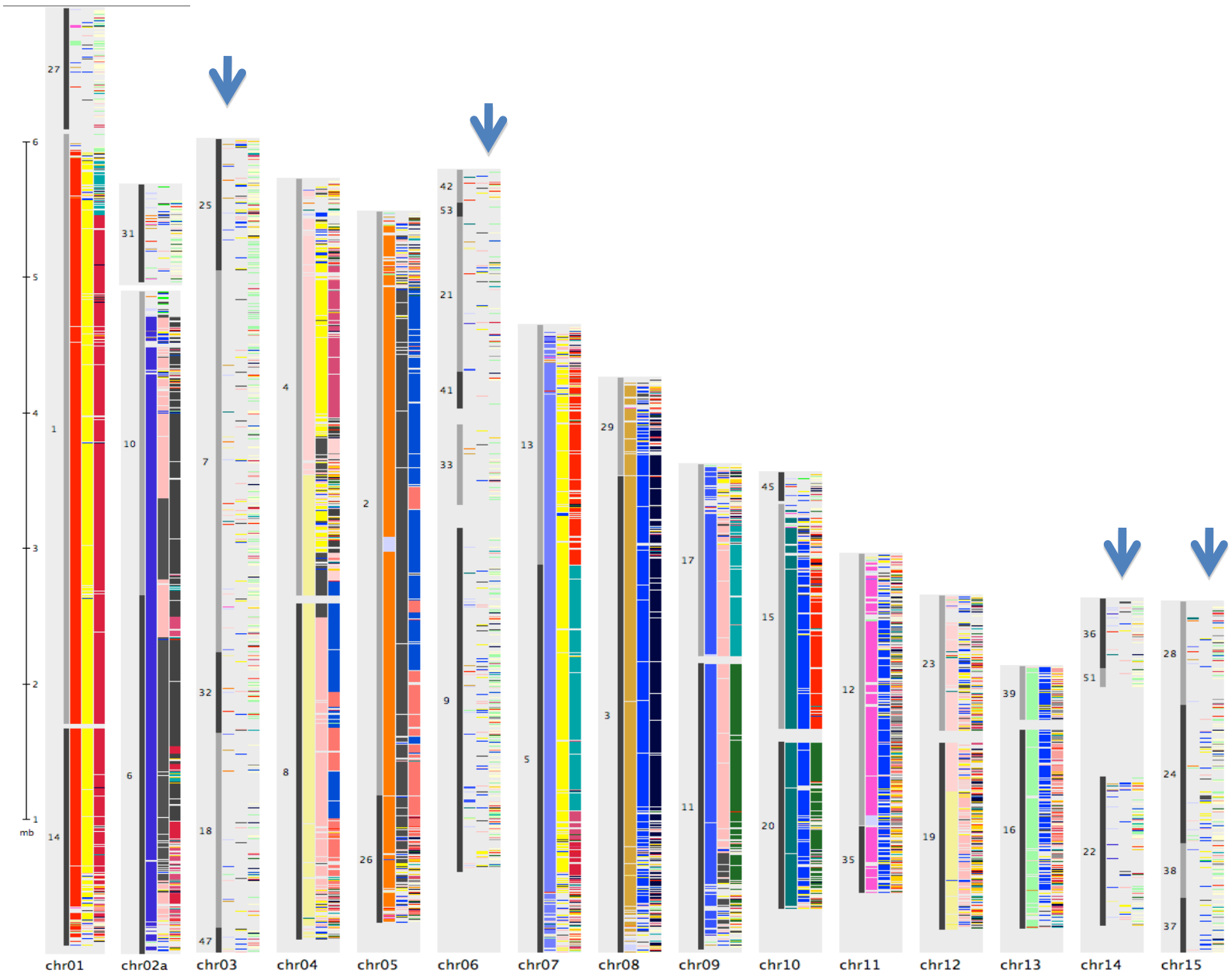

Figure S2. Syntnic alignments among four *Fusarium* genomes with Fol as the reference genome. The columns in each panel are *Fol* scaffolds mapped to the optical maps, the *Fv, Fg* and *Fs* genomes mapped to the *Fol* genomic sequences based on blastn alignments (cutoff 1e-10). The arrows point to the LS chromosomes in *Fol* that lack significant homologous sequences in all other three genomes.

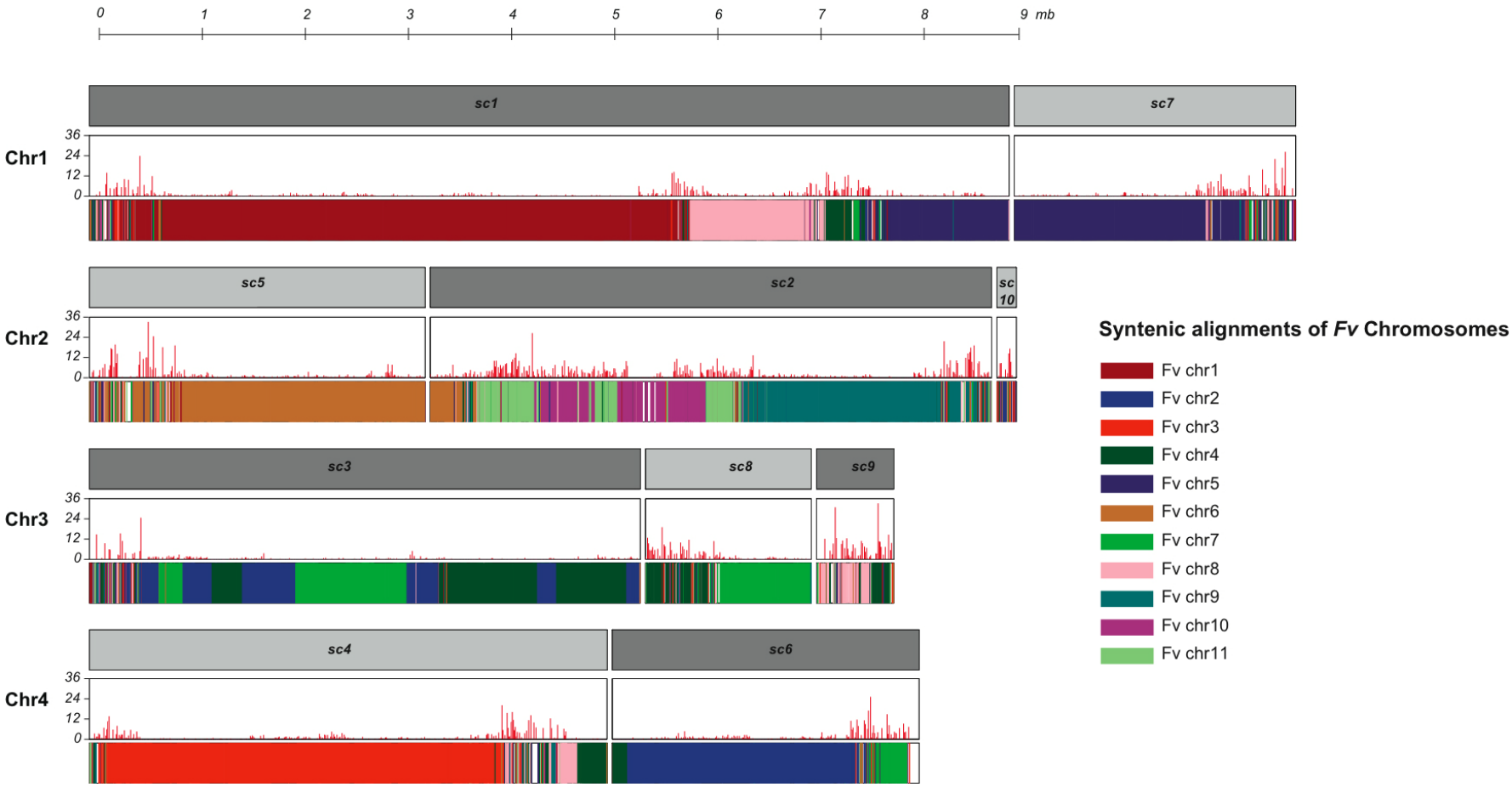

Figure S3. Whole genome alignments of *Fg* to *Fv.* The alignments display end to end synteny in large blocks with the exception of *Fv* chromosome ends and reveal multiple chromosome fusions in *Fg*. The previously described highly polymorphic and recombinogenic regions of *Fg* (*5*) correspond to *Fv* chromosome ends, including the implied interstitial fusion sites.

.

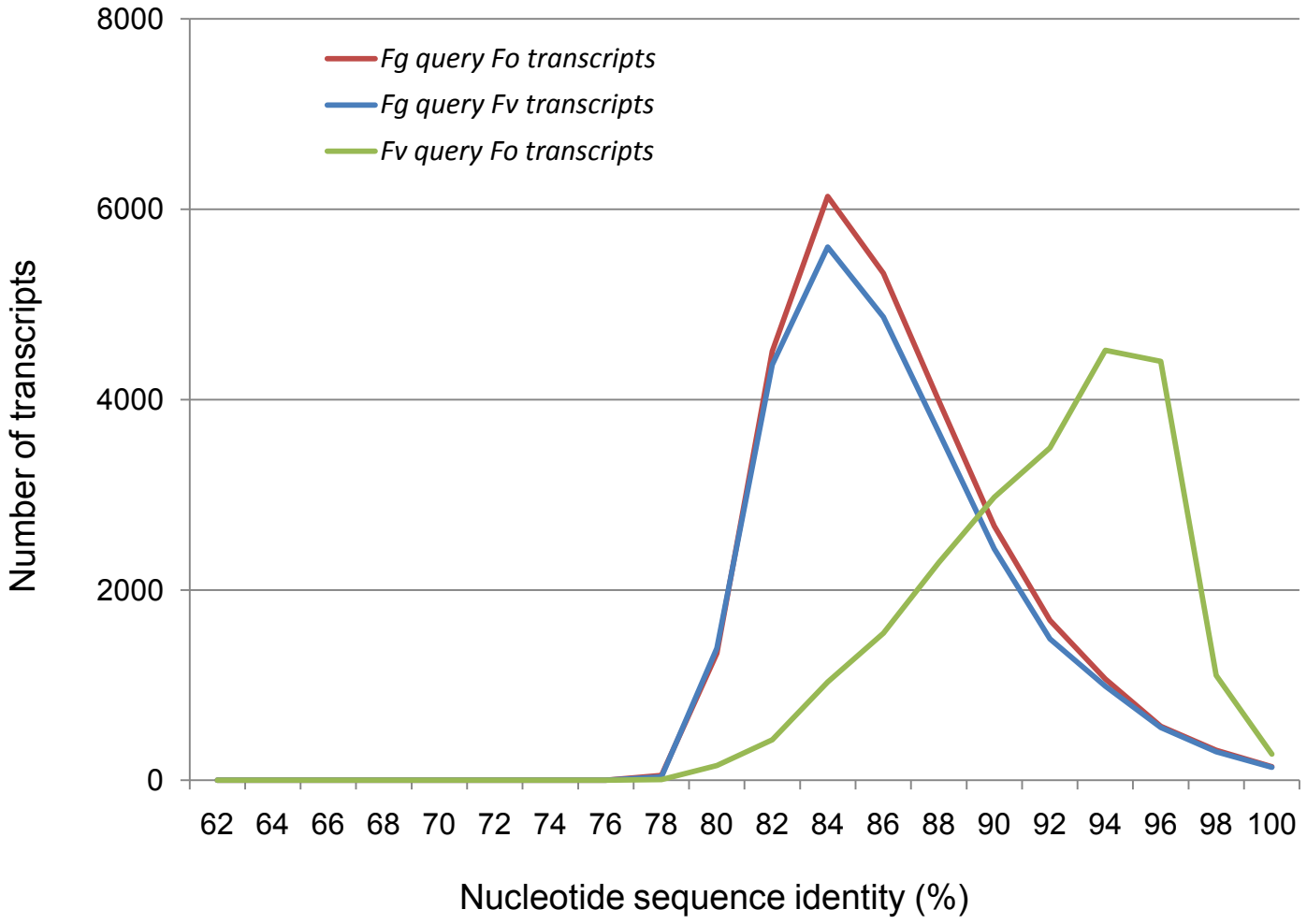

Figure S4. Pair-wise nucleotide sequence identity of coding sequences among three *Fusarium* transcripts. Based on blastn alignments of orthologous genes (cutoff 1e-10) using *Fg* (against *Fv* and *Fo*) and *Fv* (against *Fo*) transcripts as query sequences.

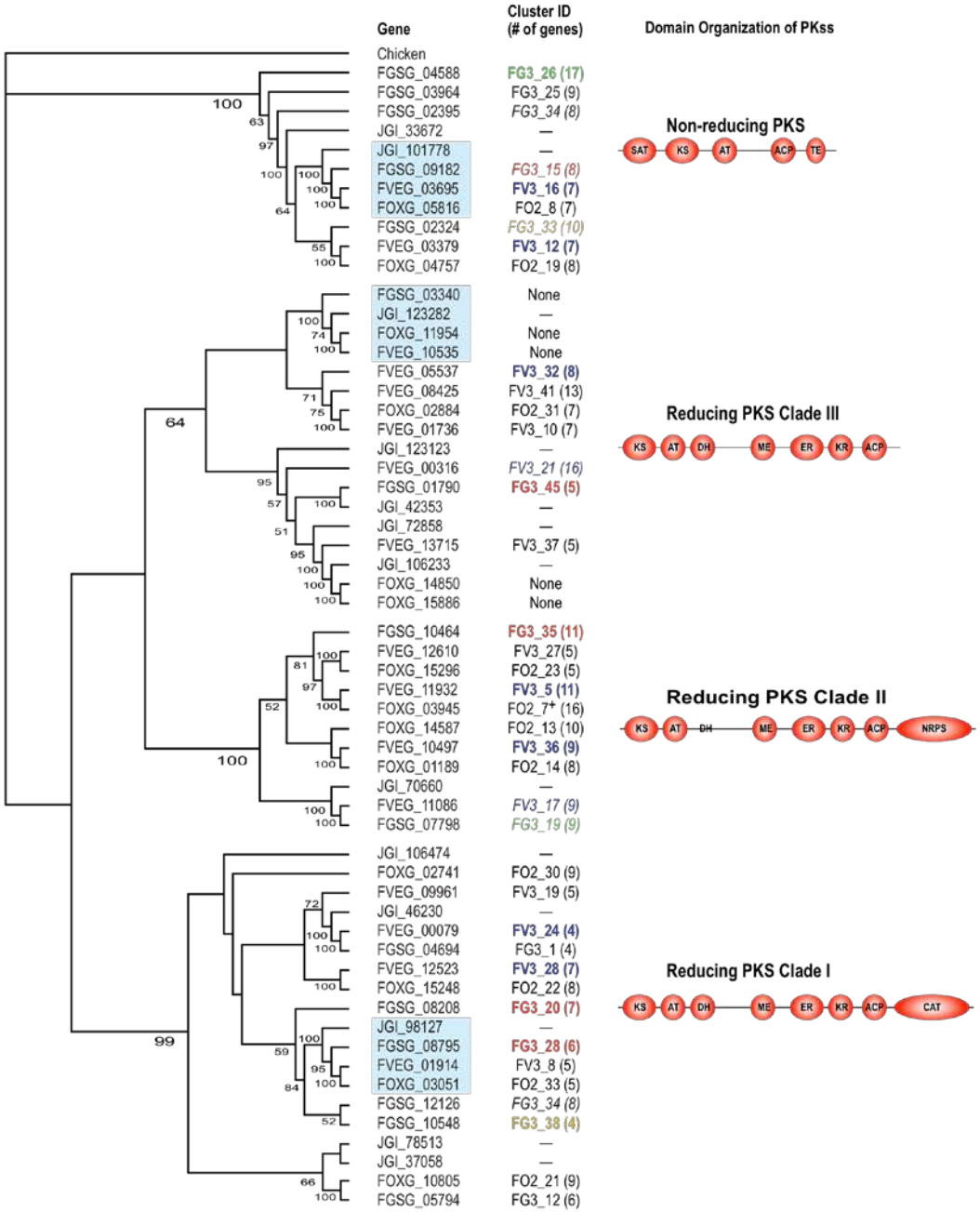

Figure S5. Cladogram representing a phylogenetic analysis of 58 *Fusarium* PKS KS (ketosynthase) and AT (acyl transferase) domains. ClustalW-aligned amino acid sequences were used to construct a gene genealogy, using parsimony in PAUP* 4.0b10. Bootstrap values were generated with 1000 replications to test for significance. The gene cluster identifier is listed next to the PKS gene that defines the cluster. The number of genes included in each cluster is listed in parentheses. The known SMB gene clusters are in italics. The clusters in bold letters are co-expressed as determined by microarray expression data [43]. The clusters in green are specifically expressed *in planta*. The clusters in red are specifically expressed during sexual development; and the clusters in yellow are expressed in culture. -- = cluster not tested; None = no cluster identified with the required four SMB related genes within a 20 kb window. ACP = acyl carrier protein domains; AT = acyl transferase; CAT = carnitine transferase; DH = dehydrogenase, ER = enoylreductase; KR = keto reductase; KS = ketosynthase; ME = methyltransferase; SAT = starter unit acetyltransferase, and TE = thioesterase.

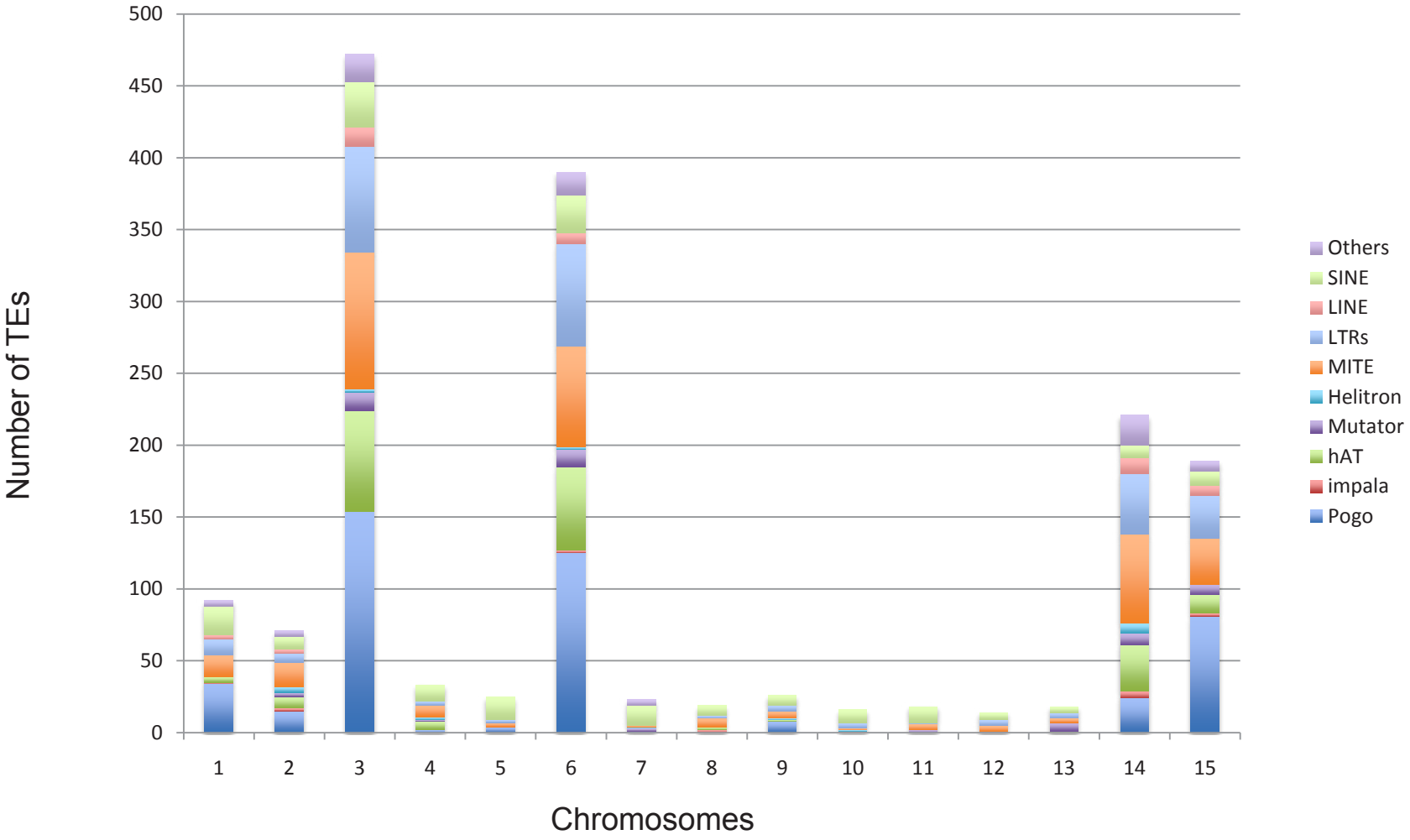

Figure S6. Distribution of TEs in *Fol* chromosomes. Full-length transposable elements were annotated using a combination of computational predictions based on BLAST analysis for transposase genes and manual inspection.

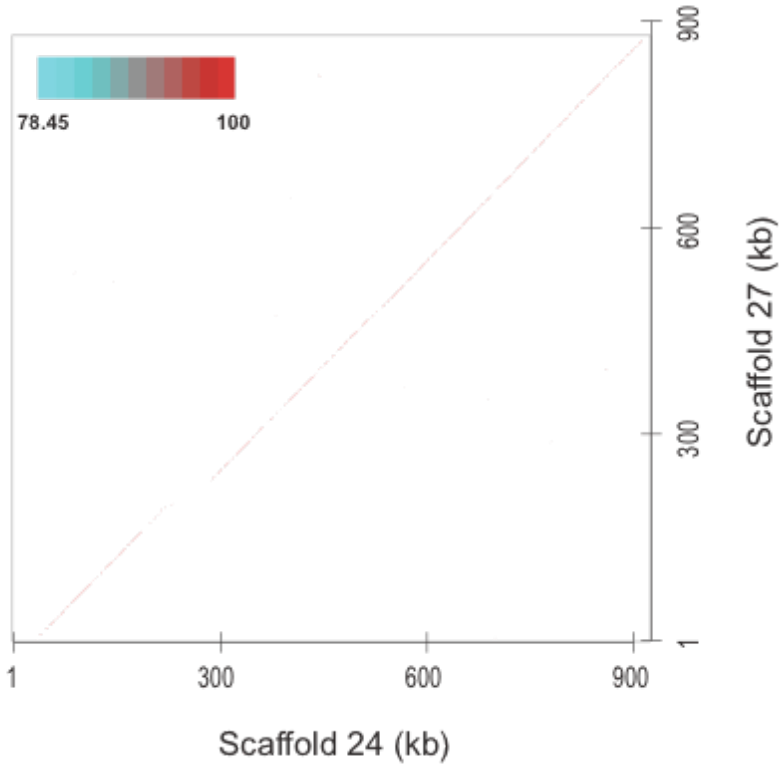

Figure S7. Dot-plot of inter-chromosomal segmental duplication as an example showing high sequence identity reflecting recent duplication, based on blastn (1e-20). The color legend indicates the level of sequence identity.

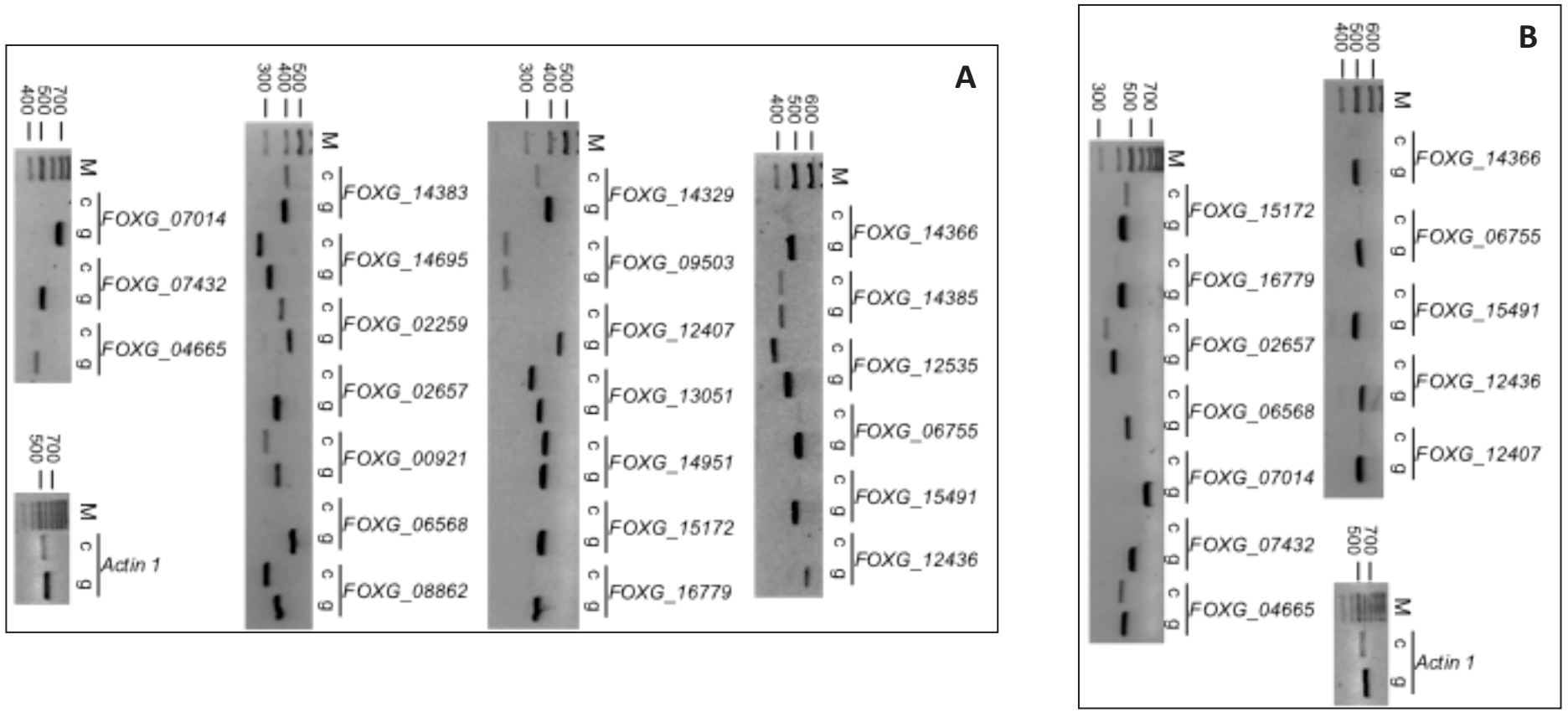

Figure S8 RT-PCR products showing the expression pattern of the indicated *F. oxysporum* CWDE genes (see Supplementary Information for gene annotations) during infection of tomato plants. cDNAs (c) were generated from total RNA isolated from roots of infected tomato plants three (A) or seven (B) days after inoculation, and used as a template for PCR amplification with gene specific primers (see Supplementary Table S13). Genomic DNA (g) was used as control for transcript size. Sizes of marker bands (M) are in bp. Predicted genes *FOXG_09503, FOXG_14383, FOXG_14385, FOXG_14951* and *FOXG_15172* lack introns (see Supplementary Table S16 for transcript sizes). Control cDNA obtained from uninoculated plants failed to produce any amplification bands (results not shown).

FOXG_04665 and FOXG_15172 are only expressed on day 7 (Part B). The very faint bands in FOXG_06755 and FOXG_15491 migrate at the same position as the gDNA, although the primers for these gene were designed to flank an intron. Therefore, the band corresponds to genomic rather than to cDNA.

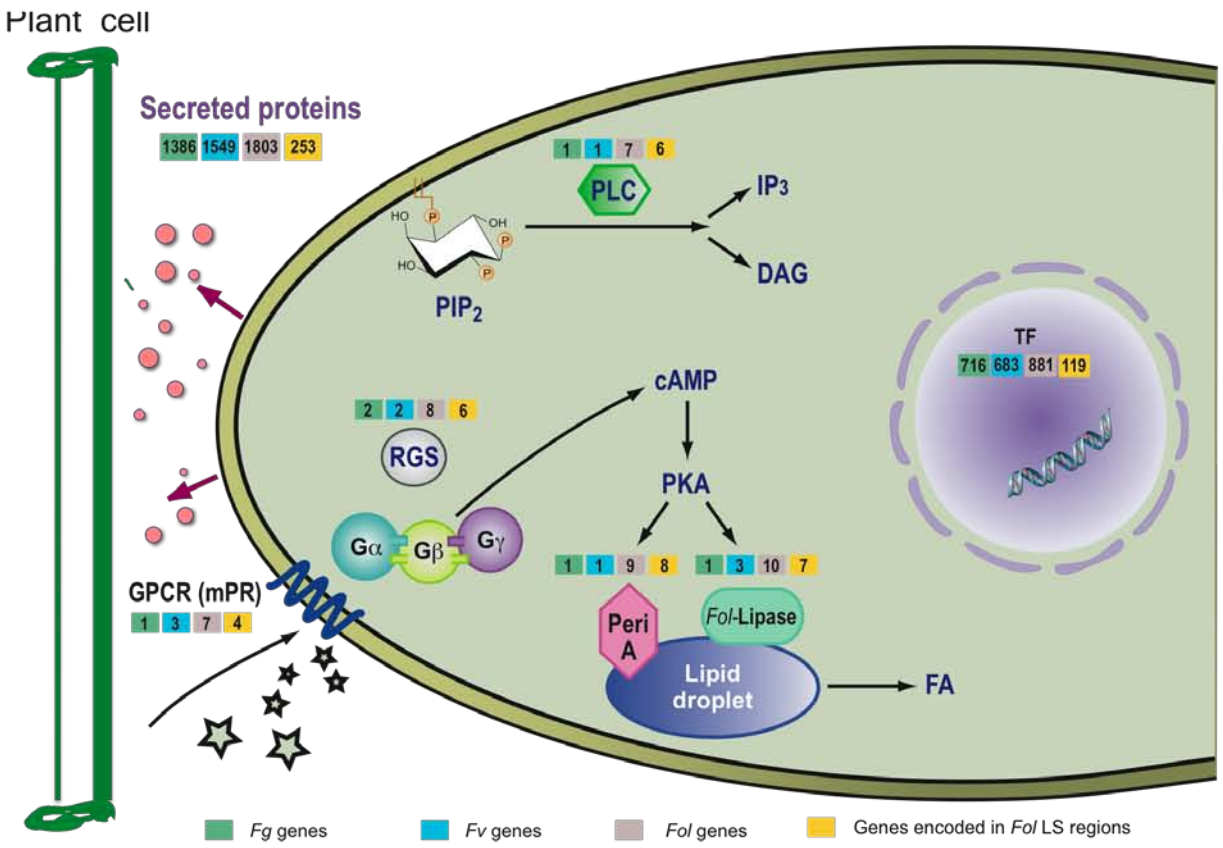

Figure S9. Proteins predicted from genes in the *Fol* LS regions are enriched for proteins predicted to be secreted (p<7e-5) and proteins of the lipid metabolism signalling pathways (p<7.6e-5).

The number of genes encoded on the *Fol* LS regions, shown in yellow, out of the total of 2448 genes encoded on LS chromosomes. In addition, genes identified from *Fg* [45], *Fv* (blue), and *Fol* (purple) in the same categories are shown. See Table S18 for details.

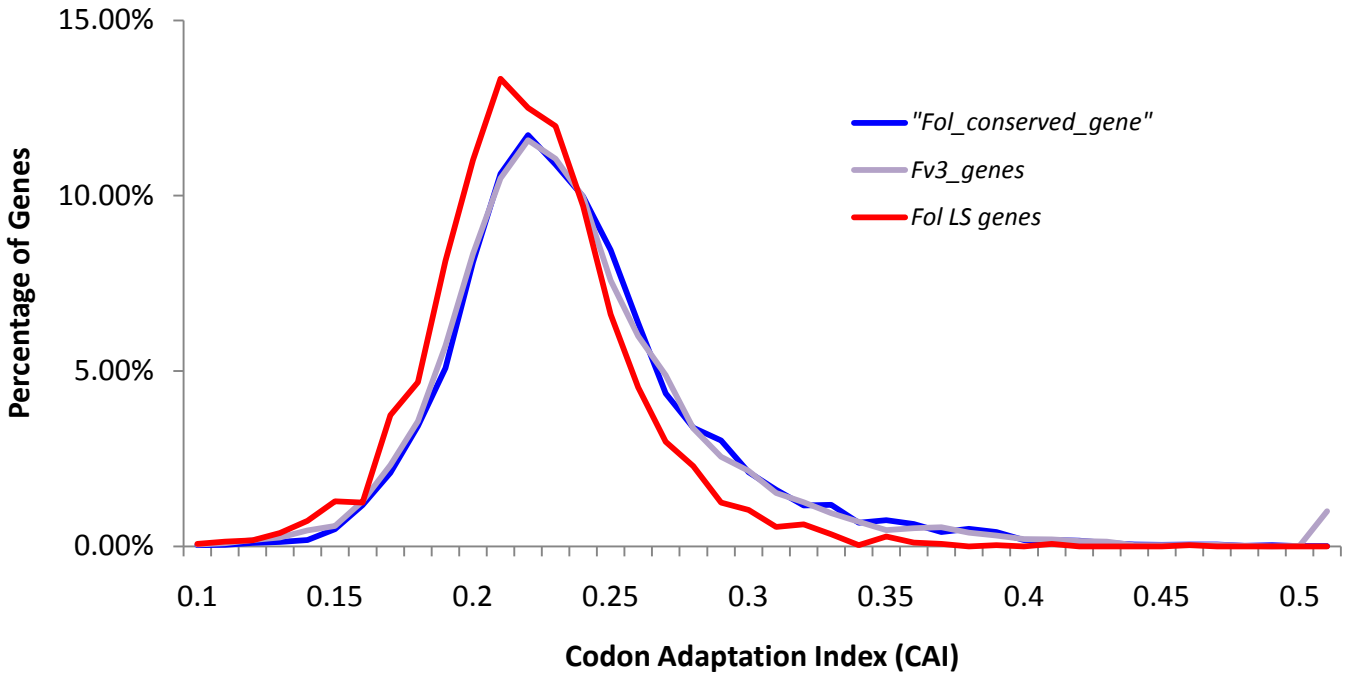

Figure S10. The Codon Adaptation Index (CAI) distribution of genes encoded *Fol* LS regions compared to genes encoded in *Fol* conserved region and the *Fv* genes. The CAI derived from the RSCU estimations is computed using the EMBOSS tool 'cai' (http://oryx.ulb.ac.be/embosshelp/cai.html).

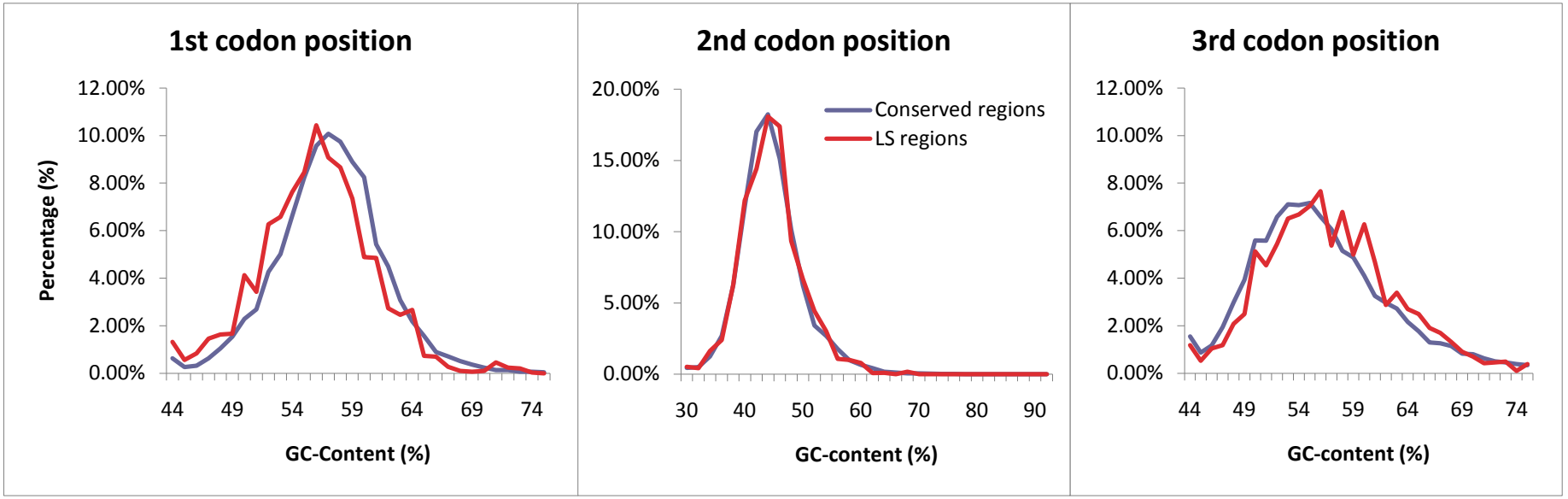

Figure S11. GC-content distribution of *Fol* genes encoded in the conserved versus LS regions.

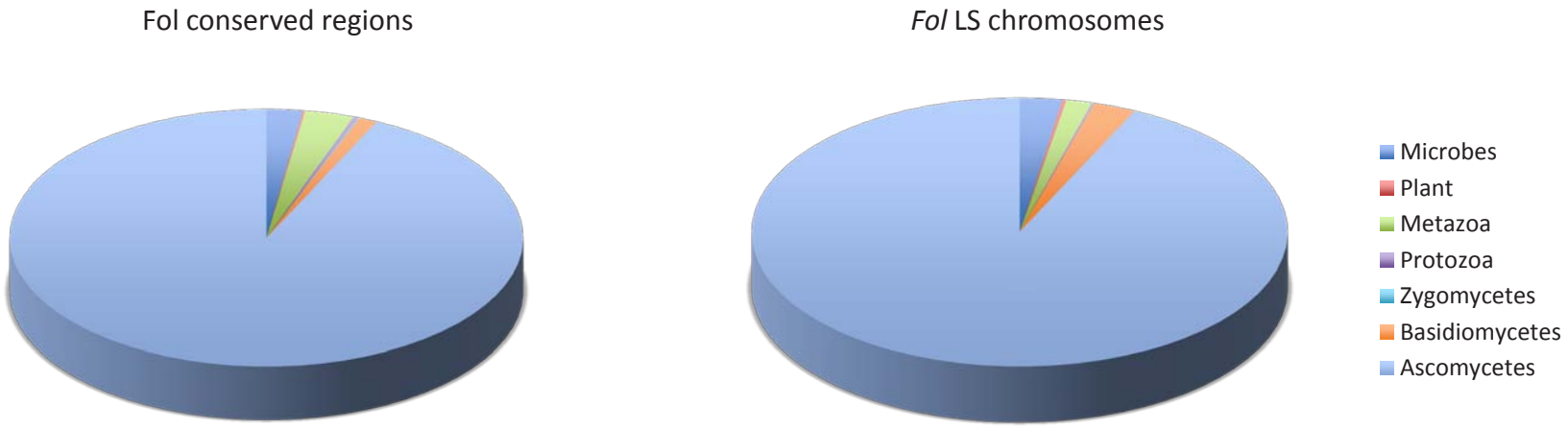

Figure S12. Homologous profile of *Fol* genes encoded in the conserved versus LS regions comparing to proteomes across different kingdoms.
The *Fol* proteins were searched using BLASTP (1e-20) against the NCBI metazoan, plant, microbial gene sets available at ftp://ftp.ncbi.nlm.nih.gov/gene/DATA/GENE_INFO (February 21, 2008 version) and the non-*Fusarium* fungal database including the protein sets from Ascomycete: two fungal genomes from each subphylum Sordariomycetes *(Magnaporthe grisea, Neurospora crassa),* subphylum Leotiomycetes*, (Botrytis cinerea, Sclerotinia sclerotiorum),* and Eurotiomycetes (*Aspergillus fumigatus, A. oryzae)*; Basidiomycete: *Ustilago maydis, Coprinus cinereus, Cryptococcus neoformans* serotype A; and a zygomycete fungal *Rhizopus oryzae* protein set.

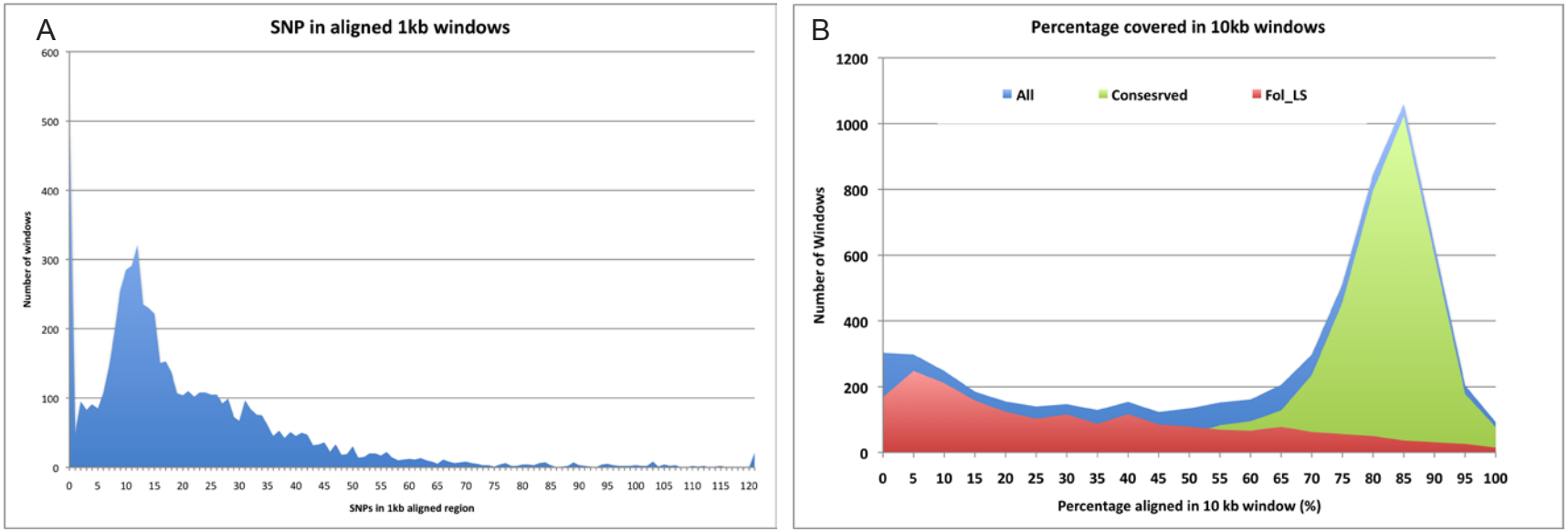

Figure S13. The SNP density (A) and the sequence coverage (B) of the reference assembly *Fol* aligned by Illumina reads from the strain FO5176.

A total of 26.7 million Illumina reads of 51 bases were aligned to the *Fol* assembly using MAQ (43). About 40% of the reference assembly is not present in the strain. A) . The overall SNP rate is less than 20 SNPs per 1 kb window. B). There are two peaks for the fraction of the reference genome that can be aligned by the Illumina reads. The major peak is centered on 85% (range from 60%-100%), and a smaller peak at the low end from 0-10%. These two peaks are clearly separated between the conserved and the *Fol*-LS regions.

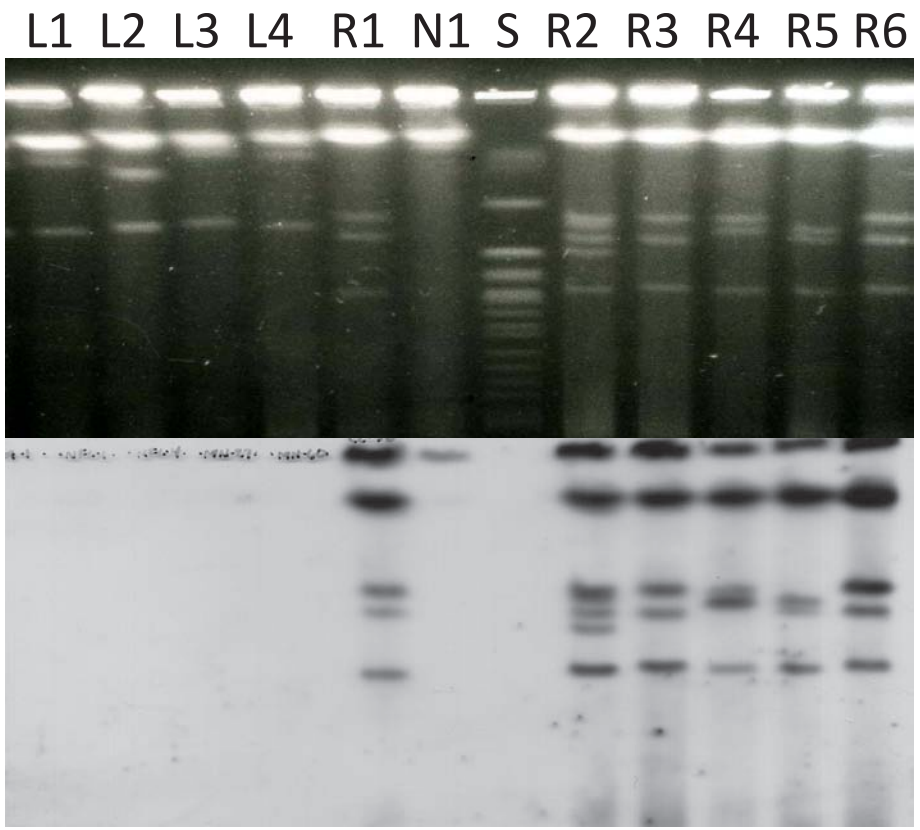

Figure S14. Karyotype variation and lineage- specific chromosomes among tomato-infecting strains of *F. oxysporum*.

(Above) Small chromosomes (< 2.3 Mb) separated from four strains of *F. oxysporum* f. sp. *lycopersici* (L1-L4), six strains of *F. oxysporum* f. sp. *radicis-lycopersici* (R1-R6), and one non-pathogenic strain of *F. oxysporum* from tomato (N1). Size standards are chromosomes from *Saccharomyces cereviseae* (S). (Below) Southern blot hybridized with a clone containing a lineage-specific repetitive sequence.

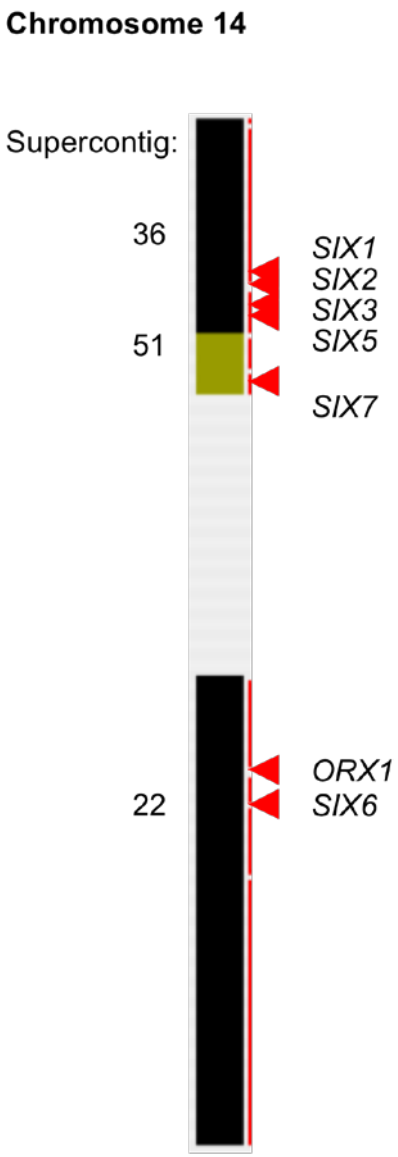

Figure S15. Schematic representation of the scaffolds attributed to chromosome 14 of Fol4287, showing the location of the *SIX (Secreted In Xylem)* genes discussed in this work and *ORX1 (OxidoReductase secreted in Xylem 1)*.

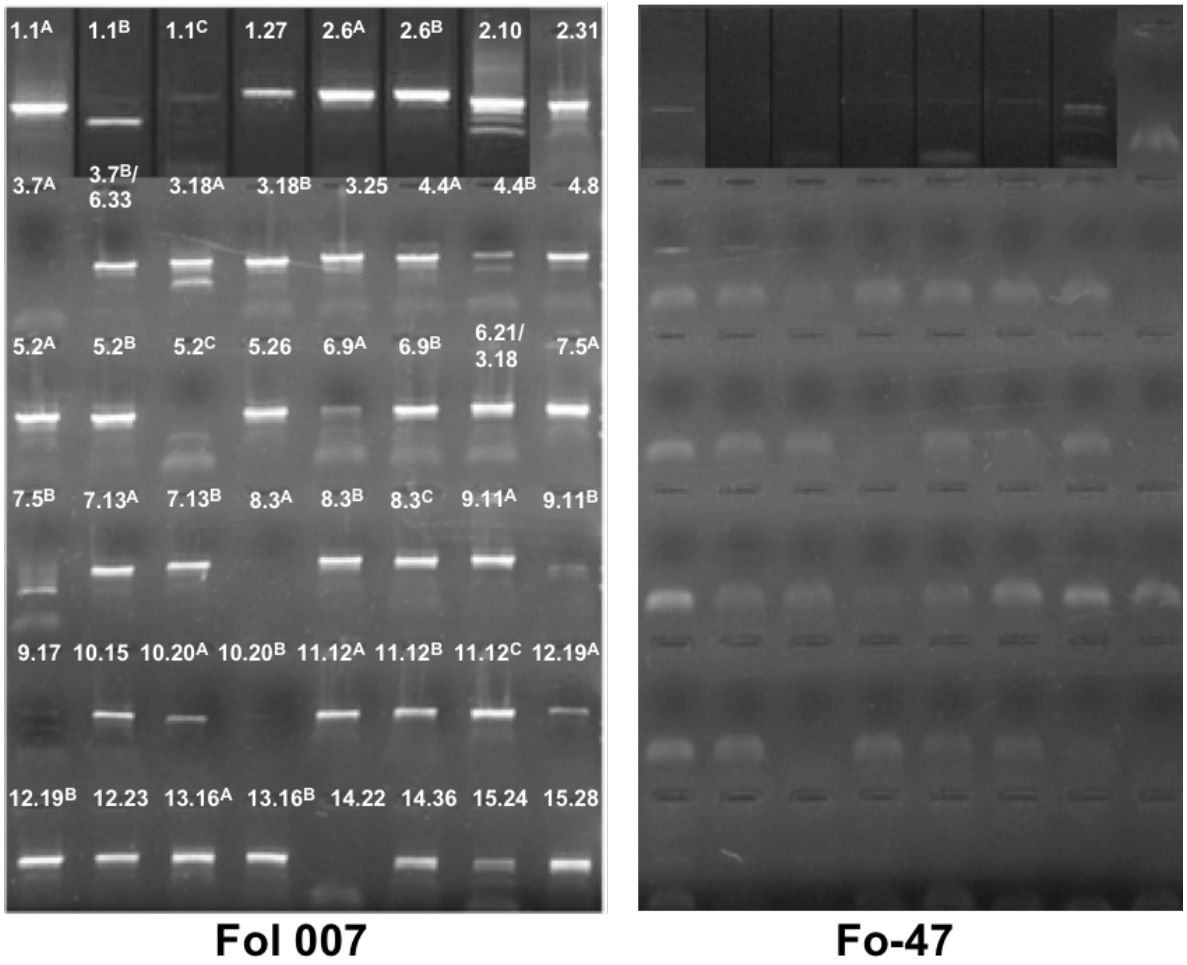

Figure S16. DNA markers based on Foxy insertions. The primer pairs were first tested for the Fol strain used in our study (Fol007), which is closely related to the sequenced isolate, and for Fo-47. Most primer pairs yielded a product with Fol007, but none with Fo-47. Numbers indicate the chromosome (number before the dot) and scaffold (number behind the dot) for which the primer was designed. Some primers are expected to anneal to more than one scaffold /chromosome, due to duplications in the Fol genome. Multiple Foxy insertions on one scaffold are indicated with letters (A, B, C etc.).

Chromosome-specific primers were designed approximately 500 bp upstream of 48 Foxy insertions in the Fol4287 genome sequence released by the Broad Institute (isolate Fol4287) (**Table S17**). Each was used with a reverse primer in Foxy.

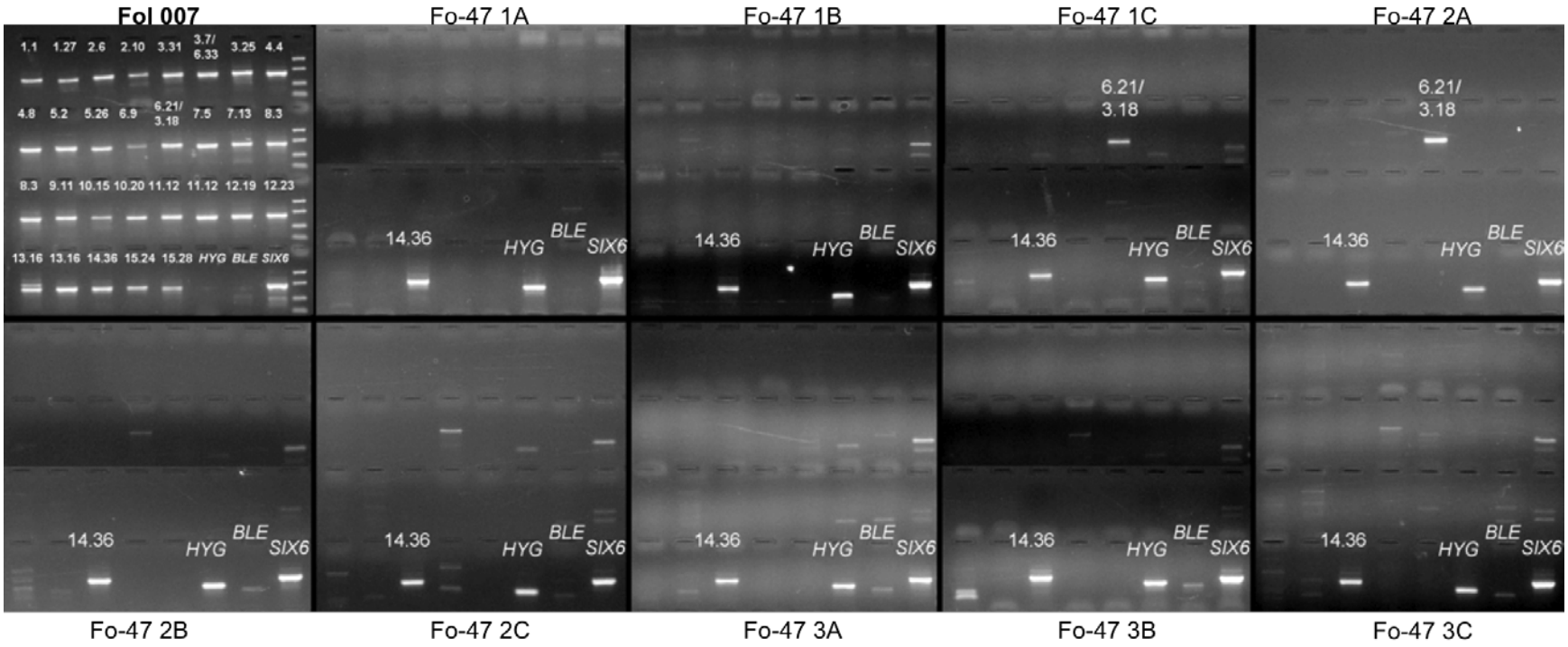

Figure S17. Twenty-nine positive primer pairs were selected for testing the double drug-resistantFo47^{+} strains for the presence of Fol007-derived chromosomes. All double drug-resistant strains were positive for the chromosome 14 marker (14.36, scaffold 36), and two strains (1C and 2A) were also positive for a chromosome 3/6 marker (6.21/3.18, scaffold 21 or 18). In some lanes a specific bands are visible (not labeled). All PCR products of the correct size are indicated with the chromosome/scaffold number.

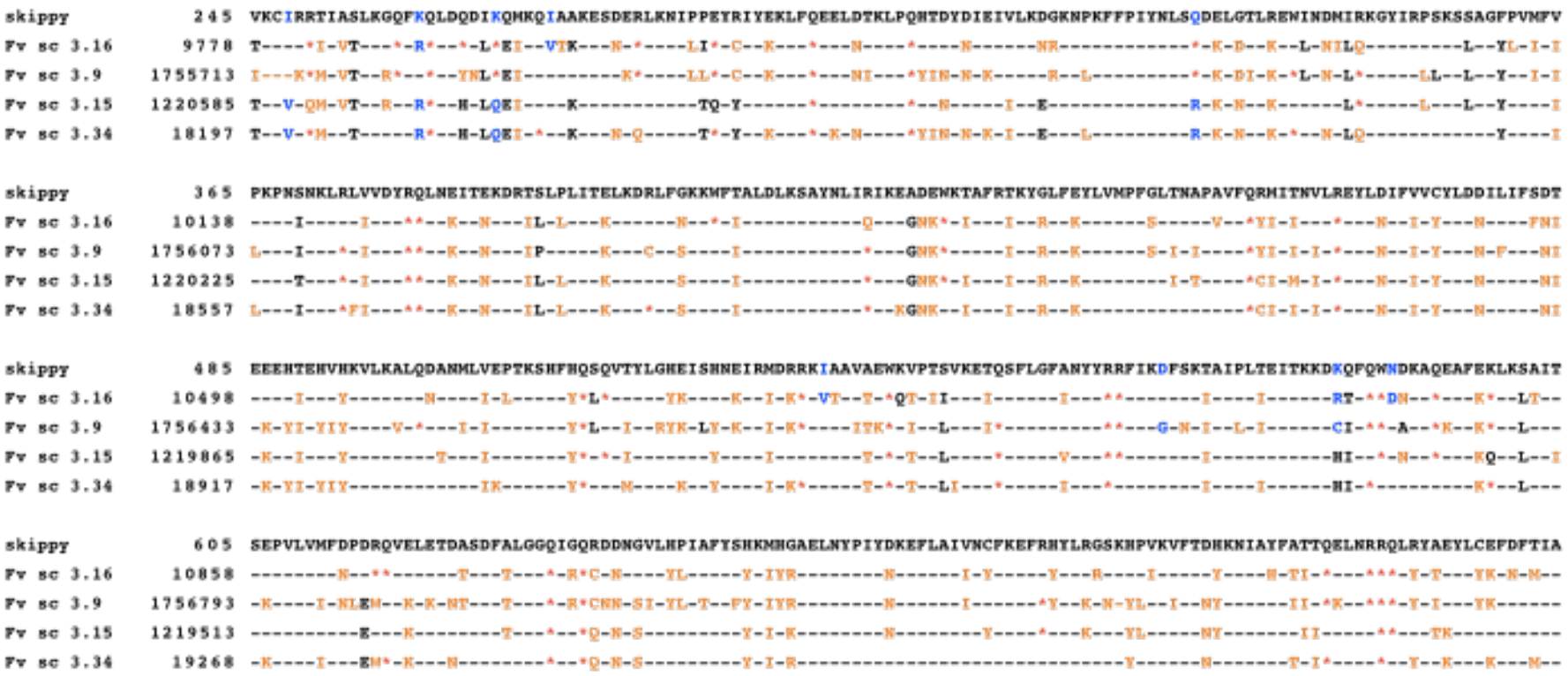

Figure S18. Alignment of four mutated skippy-like TEs from *Fv* to *Fol* Skippy suggests past occurrence of RIP in *Fv*.

Numerous nonsense codons (red) were introduced by RIP-type mutations. Missense mutations explainable by single C:G to T:A mutations from *Fol* Skippy to *Fv* SLRE are shown in orange. Missense mutations that can be explained by RIP of *Fv* SLRE to *Fol* Skippy are shown in blue. All other mutations as shown in black.

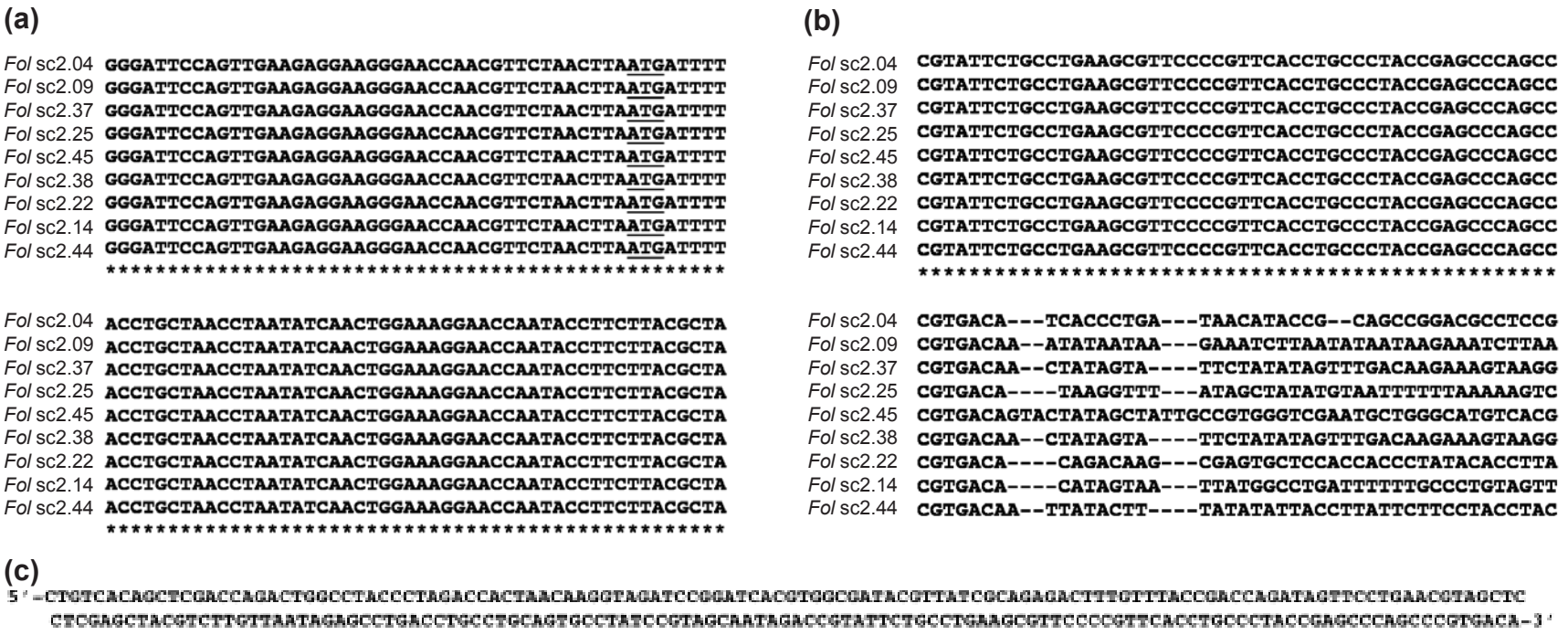

Figure S19. A novel non-RIPped skippy-like retroelement (SLRE) from *Fol*. A similar element had been previously identified as "Skippy" (51) and is related to "Maggy" transposons from *M. grisea* (81). Nine non-mutated *slre* TE coding regions were identified by tblastx searches of the *Fusarium* genome sequences with all previously identified fungal and non-fungal TEs. **(a) partial DNA sequence alignment of skippy elements reveals absence of RIP.** No RIP-type mutations were found across the nine full-length *skippy* elements. Part of the complete alignment is shown from nt 2523592 on the - strand of sc2.04 (the full-length element is 5688 bp long, from sc2.04 nt 2525324-2519636). The predicted Skippy ATG (nt 2523550) is underlined. The preceding *gag* ORF (not shown) is less conserved and characterized by numerous indels. **(b) The 3' boundary of skippy elements.** Top of panel shows 3' end of the completely conserved LTR (ends at CGTGACA in the bottom panel). **(c) Sequence of the identical LTRs.** The 239-bp long LTRs extend from 2519636-2519874 and 2525087-2525325.

(a)

```
Fv_01865    781 FKSRLELWQKH--NRQELPENIVIFRDGVSTGQFAQVLRTELPRIRIACNGKYPKNK---
Fol_03010   783 FKSRLELWQKH--NRQELPENIVIFRDGVSTGQFAQVLRTELPRIRIACNGKYPKNK---
Fg_08752    818 FKSRLALWQMH--NRQQLPENIVIFRDGVSEGQFAQVLQRELPRIRIACNAKYPKNK---
Nc_04730    870 FKTRLELWRSNPANNRSLPENILIFRDGVSEGQFQMVIKDELPLVRAACKLVYPAGK---
Pa_g2992    707 FKSRLRLWQKH--NGAKLPENILIYRDGVSEGQFNMVLTSELPHIRTACSQMYGK-Q---
Fol_12456   784 LKSRLGLWKTKG-KHAALPENILIYRDGVSEGQYDMVLSQELPQLRRACEQMYPAADTKK
Fol_14081   764 LKSRLGLWKTKG-KHAALPENILIYRDGVSEGQYDMVLSQELPQLRRACEQMYPAADTKK
Fol_16455   766 LKSRLGLWKTKG-KHAALPENILIYRDGVSEGQYDMVVSQELPQLRRACEQMYPAADTKK
Ac_01617    772 LKSRLSLWKTKG-KHTALPENILIYRDGVSEGQYDMVLSQELPQLRRACEQVYPTADTKK
Fv_00803    750 FGELIPRWRAN--HPSMVPKHLIYFRDGVGEGQFAYVLDQEVEEIKKYLSTVLPAGQ---
Fol_00711   750 FGELIPRWRAN--HQGLVPKHLIYFRDGVGEGQFAYVLDQEVEEIKKYLATILPGGQ---
Fg_00348    746 FAELLPQWRHN--HPGKIPAHLIYMRDGVSEGQFAHVLEQEVSEIKKFFGGSLPPDK---
Mg_11029    798 FGPLVERWCKT--MR-CAPEHVFYLRDGVSEGQFAHVMALEVAELRKVLNKVG--GN---
consensus   961 fksrl-lWr--------lPenivifRDGVseGQfa-Vl-qElp-irrac--vyp-g----

Fv_01865    836 -APKISIIVSVKRHQARFYPTS-EENAMEKNHNIQNGTVVDRGITEARYWDFYLTAHASI
Fol_03010   838 -APKISIIVSVKRHQTRFYPTS-EENAMEKNHNIQNGTVVDRGITEARYWDFFLTAHASI
Fg_08752    873 -PPRISLIVSVKRHQTRFYPTS-SE-SMTSKNNIENGTIVDRGVTQARYWDFFLTAHSSI
Nc_04730    927 -LPRITLIVSVKRHQTRFFPTD-PKHIHFKSKSPKEGTVVDRGVTNVRYWDFFLQAHASL
Pa_g2992    761 -QPRITLIVSVKRHQTRFYPTD-PQQTHFRSKSPKEGTVVDRGVTNVRYWDFFLQAHASL
Fol_12456   843 GLPRFTIIVCGKRHKTRFYPTT-EQDCDR-SGNTKPGTVVDRGVTEARNWDFFLQAHAAL
Fol_14081   823 GLPRFTIIVCGKRHKTRFYPTT-EQDCDR-SGNTKPGTVVDRGVTEARNWDFFLQAHAAL
Fol_16455   825 GLPRFTIIVCGKRHKTRFYPTT-EQDCDR-SGNTKPGTVVDRGVTEARNWDFFLQAHAAL
Ac_01617    831 GLPRFTIIICGKRHKTRFYPTT-EQDCDR-SGNTKPGTVVDRGVTEARSWDFFLQAHAAL
Fv_00803    805 -MPKFTVIVATKRHHIRFFPQ-----RGDKNGNPLPGTLVEREVTHPFMFDFYLNSHVAI
Fol_00711   805 -MPKFTVIVATKRHHIRFFPQ-----RGDKNGNPLPGTLVEREVTHPFMFDFYLNSHVAI
Fg_00348    801 -IPKMTVIIATKRHHVRFFPQ-----RGDKNGNPLPGTLVEKEVTHPFMFDFYLNSHVAI
Mg_11029    850 -NPKITVMVATKRHHIRFFPKPGDSSSGDRNGNALPGTVVERVVTHPFHYDFYLCSHVAI
consensus  1021 -lPritiiv--KRH-tRFyPts-e-----k-gn--pGTvVdrgvT-ar-wDFfLqaHaai

Fv_01865    894 KGTARPAHYTVLLDEIFQAKFK-------SEAANELEKFTHELCYLFGRATKAVSICPPA
Fol_03010   896 KGTARPAHYTVLLDEIFRSKFK-------SEAANELEKFTHELCYLFGRATKAVSVCPPA
Fg_08752    930 KGTARPAHYTVLLDEVFRAKYG-------AEAANELERYAHELCYLFGRATKAVSICPPA
Nc_04730    985 QGTARSAHYTVLVDEIFRADYG-------NKAADTLEQLTHDMCYLFGRATKAVSICPPA
Pa_g2992    819 QGTARPAHYTVLIDEIFRPSYG-------AQAANNLEQVTHDMCYLYGRATKAVSICPPA
Fol_12456   901 QGTARPCHYYIVHDEIFRQIYAKSIPVPFQSIADIVEDLTHNMCYLL---LENPGIPFPS
Fol_14081   881 QGTARPCHYYIVHDEIFRQIYAKSIPVPFQSIADIVEDLTHNMCYLL---LENPGIPFPS
Fol_16455   883 QGTARPCHYYIVHDEIFRQIYANSIPLPFLSIADIVEDLTHNMCYLFGRATKSVSLCPPA
Ac_01617    889 HGTARPCHYYIVHDAIFRQIYAKLIPSPFQNIADIVEDLTHNMCYLFGRATKAVSLCPPA
Fv_00803    859 QGTARPVHYHIILDEMN-------------MPVNDLQKMIYQQCYSYARSTTPVSLHPAV
Fol_00711   859 QGTARPVHYHVILDEMN-------------MPVNDLQKMIYQQCYSYARSTTPVSLHPAV
Fg_00348    855 QGTARPVHYSVILDEMA-------------MPVNDLQKMIYQQCYSYARSTTPVSLHPAV
Mg_11029    909 QGTARPTHYQVIHDEVG-------------YSPDELQKMLYQQCYQYARSTTPVSLHPAI
consensus  1081 qGTARp-HY-vllDeifr--yg----------aneleklthnlCYlfgratk-vsicppa
```

(b)

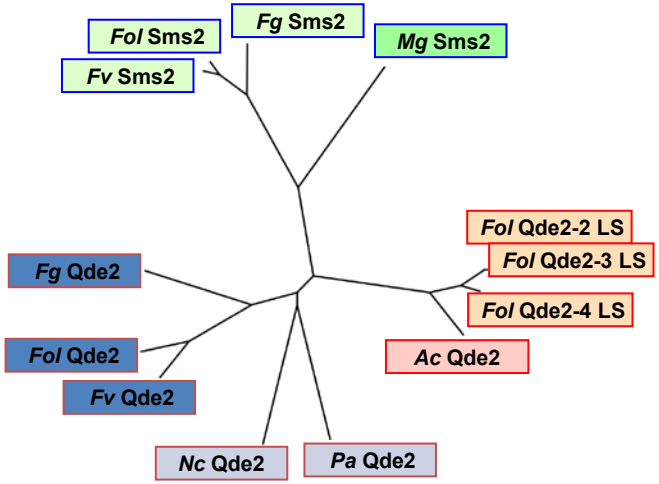

Figure S20. Three conserved *Fol* Qde2 proteins are localized on LS chromosomes. (a) Partial alignment of the QDE-2 PIWI domains. *Fusarium* Qde2 proteins cluster with the *Neurospora* and *Podospora* homologues (top five lines), as do *Fusarium* Sms2 homologues and *Magnaporthe* Sms2 (bottom four lines). Three additional *Fol* Qde2s are more related to *Ajellomyces* Qde2 than the *Fusarium* Qde2s (four center lines). Identical (blue), conserved (cyan) and similar (green) residues are colored; completely different residues are shown in grey. ClustalW was used with default settings. **(b) Relationship between *Fusarium* QDE-2 homologues.** A single ClustalW unrooted tree was constructed with the full-length predicted Qde2 and Sms2 proteins from *Fol*, *Fv*, and *Fg* (see Table 1) and their closest non-*Fusarium* matches among sequenced fungi *(Magnaporthe grisea* Sms2, MGG_11029; *Ajellomyces capsulatum* Qde2, HCAG_01617; *Podospora anserina* QDE-2, PODANSg2992; *Neurospora crassa* QDE-2, NCU04730).