## SUPPLEMENTARY FIGURES

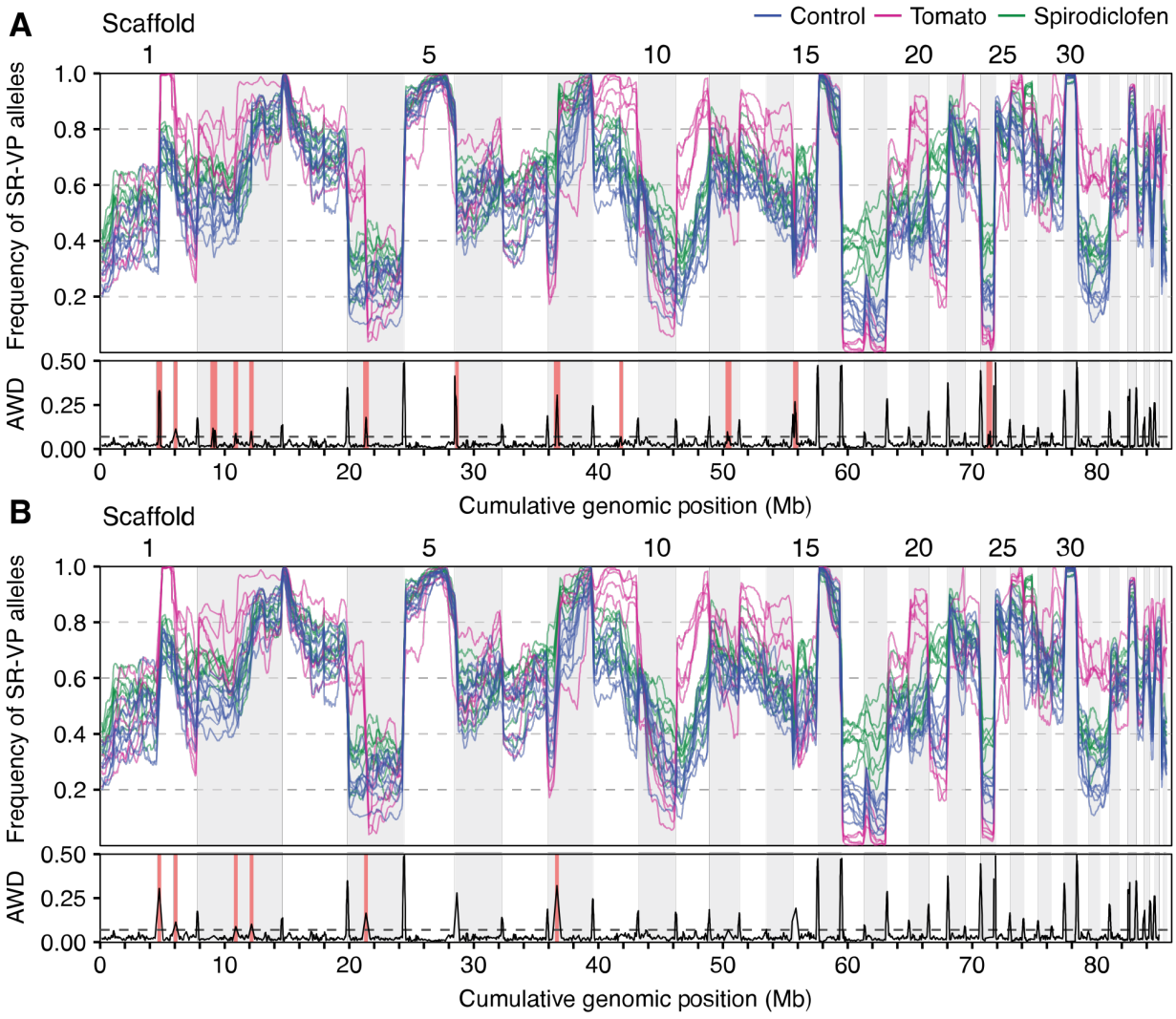

**Figure S1** Allele frequency data from experimental populations identifies misassemblies in the draft *T. urticae* Sanger assembly. (A) Allele frequencies in populations by control/treatment (legend, top right) visualized genome-wide (lines indicate frequencies of SR-VP alleles). A concatenation of the largest 44 Sanger scaffolds is shown (ordering is by decreasing length with alternating shading). At bottom, the average window distance (AWD) metric is plotted as calculated from the overlying allele frequency data. As expected, AWD values peak at (artificial) junctions between scaffold concatenations, establishing a *de facto* expectation for the range of AWD values consistent with errors in the Sanger assembly; nearly all of these peaks have AWD values >0.10. Subject to a conservative cutoff value based on two standard deviations of AWD values (0.07, dashed line), vertical red bars indicate sites of elevated AWD values that were investigated as potential misassemblies. (B) Allele frequencies and AWD values as shown in panel A, except that windows causing elevated AWD values as shown in panel A (red bars) were removed for AWD calculations (see Materials and Methods for the rationale). Briefly, at sites of high AWD values internal to Sanger scaffolds (such as might occur at a stretch of errant variant calls, or a small transposition, but not at a full-fledged misassembly), if the offending region is removed and the AWD trace is recalculated with the closest remaining windows, the resulting AWD values should return to representative genome-wide values (alternatively, they would remain high if a major misassembly is present). After applying this filter, only six potential locations of misassemblies remained (although given the window sizes used in the analysis, see Materials and Methods, we had little power to detect misassemblies internal to small Sanger scaffolds). Of the six remaining AWD peaks (red shading), ones on Sanger scaffolds 1, 4 and 8 were obvious misassemblies with very large AWD values, and the first peak on Sanger scaffold 2 was also identified as a misassembly based on data from short-read *de novo* assemblies of *T. urticae* strains (Table S1), and the report of Bryon *et al.* 2017. That the two remaining peaks were errors in the Sanger assembly was not supported by Illumina *de novo* assemblies, or by allele frequency data from independent studies (Figure S4 and Figure S5, and see Materials and Methods). Note that allele frequency and AWD values in panel B correspond to those in Figure 4A.

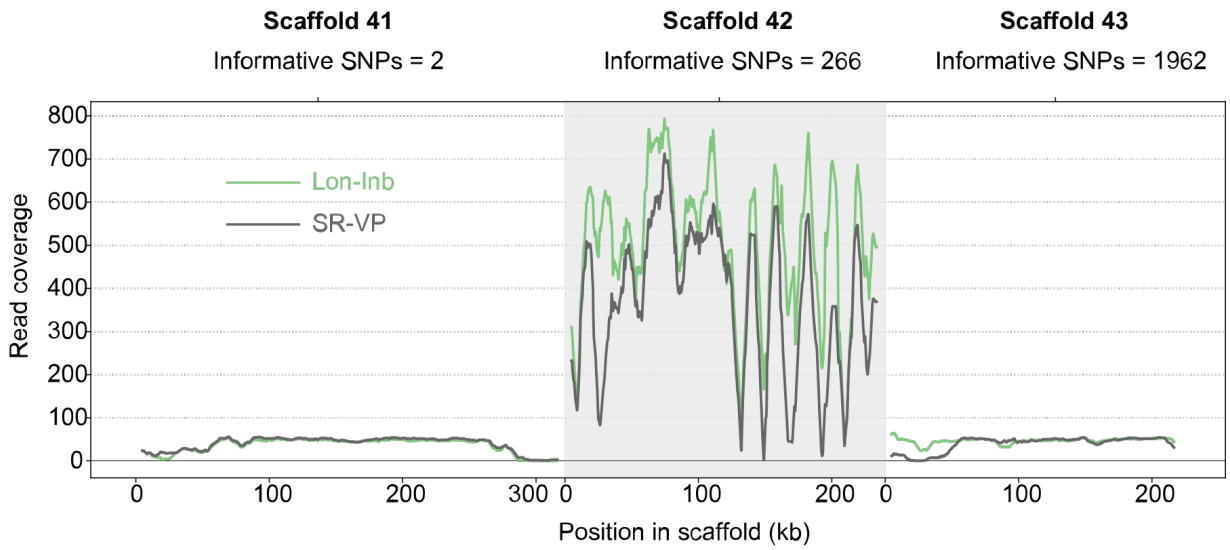


**Figure S2** Read coverage for inbred strains Lon-Inb and SR-VP reveals that Sanger scaffold 42 is highly copy variable. Coverage in short reads is shown for Sanger scaffolds 41-43 (indicated by alternating white and gray shading) for strains Lon-Inb (green) and SR-VP (dark gray). The number of high-quality, informative SNPs (average for experimental populations, see Materials and Methods) on each scaffold is given at top. Note the drops or variation in coverage among or between strains at the ends of Sanger scaffold 41, and at the beginning of Sanger scaffold 43. This pattern, which was also observed on other Sanger scaffolds, motivated our use of offsets from the ends of scaffolds when using the AWD approach to consolidate pseudochromosomes (see Materials and Methods). Read coverage was calculated with sliding windows of 10 kb with 500 bp offsets.

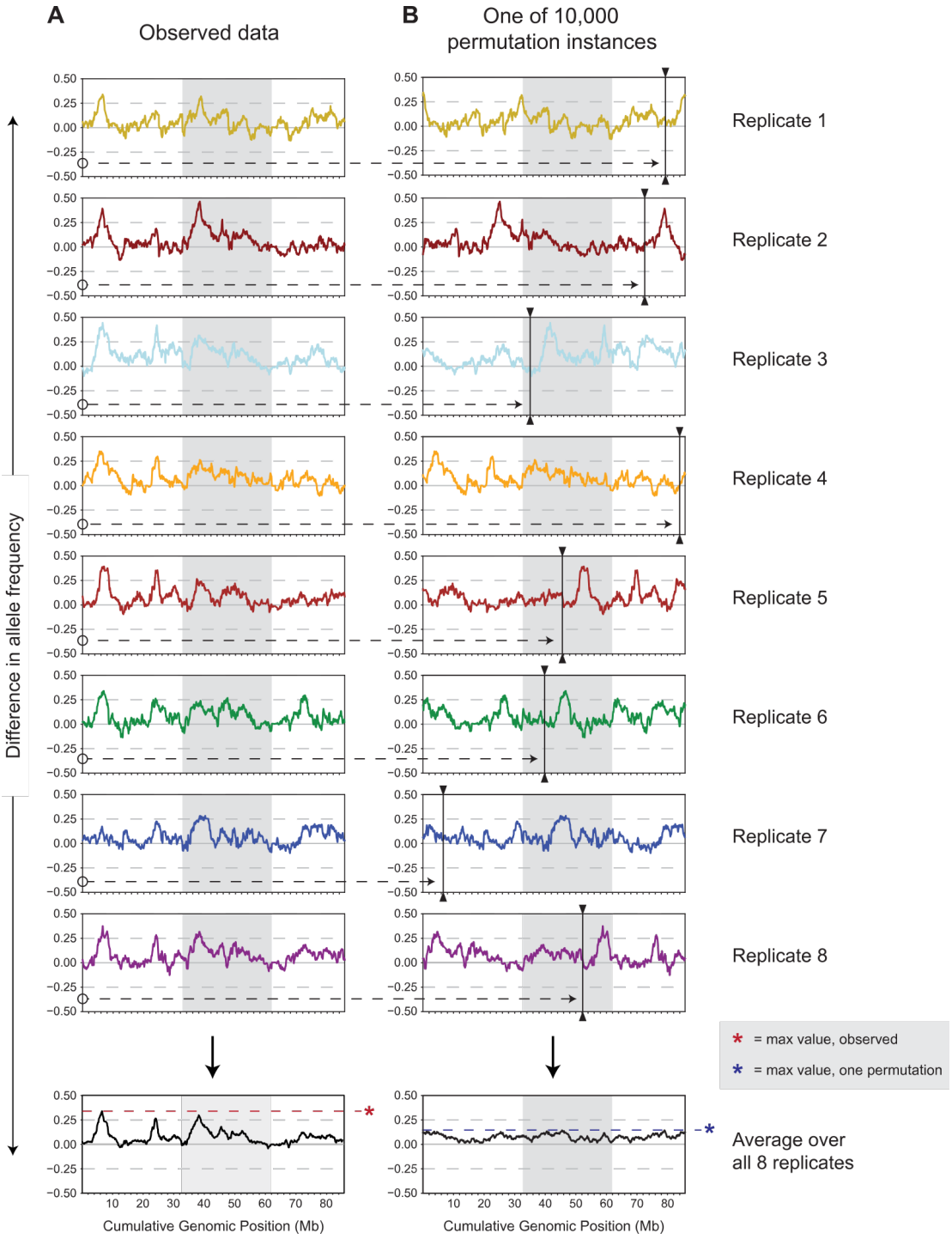
A
Observed data
B
One of 10,000
permutation instances
Replicate 1
Replicate 2
Replicate 3
Replicate 4
Replicate 5
Replicate 6
Replicate 7
Replicate 8
Difference in allele frequency
0.50
0.25
0.00
−0.25
−0.50
10 20 30 40 50 60 70 80
Cumulative Genomic Position (Mb)
* = max value, observed
* = max value, one permutation
Average over
all 8 replicates

**Figure S3** Illustration of permutation-based detection of QTL from replicated, paired selected and non-selected populations. In the left column, allele frequency changes between selected and non-selected samples are shown genome-wide for eight replicates. Specifically, the *T. urticae* genome is shown (3 pseudochromosome assembly, Figure 5; pseudochromosomes are indicated by alternate shading), and the eight replicates reflect the respective experimental data from eight paired spirodiclofen/control populations (see Figure 6A, and the Materials and Methods, for more information). Across replicates, consistent and large deviations in allele frequencies toward the SR-VP parent (positive deflections from zero as plotted) are apparent. As revealed from averaging across all eight replicates (bottom left), three peaks are pronounced (these peaks correspond to spiro-QTLs 1-3 in Figure 6B and Figure S6). To establish that the peaks do not result from chance deviations at the same genomic locations in multiple replicates, we performed permutations as indicated in the right column (a single instance is shown). Briefly, the start of the allele frequency trace for each replicate was shifted to a random location (see dashed black arrows from the left to right columns, and vertical solid black lines with terminal triangles). When shifted "off-genome" towards the end of this chromosome order, the overhanging trace segments were added back to the beginning of the chromosome order. Therefore, no information was lost (it is as if the traces were made circular by connecting the end of the third chromosome with the beginning of the first chromosome, and the start location was spun to a random location). As for the observed data (left column), for each permutation an average across all eight shifted traces was calculated (bottom right). The maximal deviation (dashed blue line and asterisk) is less than that calculated with the true genomic location data (left column, bottom, dashed red line and asterisk). By performing $10^4$ permutations, the distribution of maximum values expected by chance deviations of traces among replicates was established. The resulting distribution was used to assess significant deviations from zero that reflect responses to selection (see Materials and Methods for details).

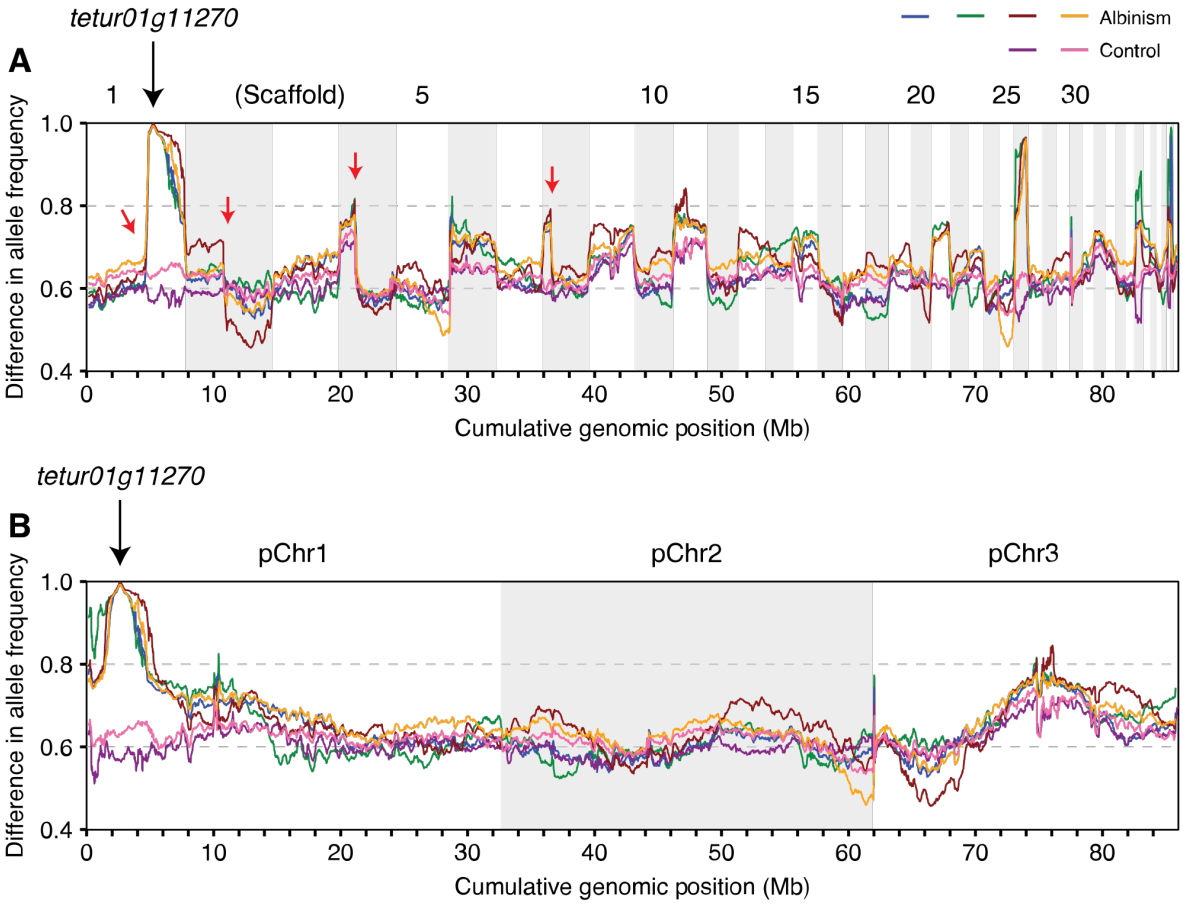


**Figure S4** Reanalysis of population allele frequency data from Bryon *et al.* (2017) on the three pseudochromosome assembly. (A and B) Allele frequency data for six populations (four selected for albinism and two unselected control populations as indicated in the legend, top right) used to identify the locus encoding phytoene desaturase (*tetur01g11270*) as causal for monogenic, recessive albinism in *T. urticae* in an earlier study (Bryon *et al.*, 2017). The reanalysis and display of this data on both the concatenated Sanger scaffolds (A) and the three pseudochromosomes (B) are as described in the legend for Figure 4. The locations of putative misassemblies in Sanger scaffolds as observed by Bryon *et al.* (2017), as well as in the current study (Figure 4), are indicated by red arrows in panel A.

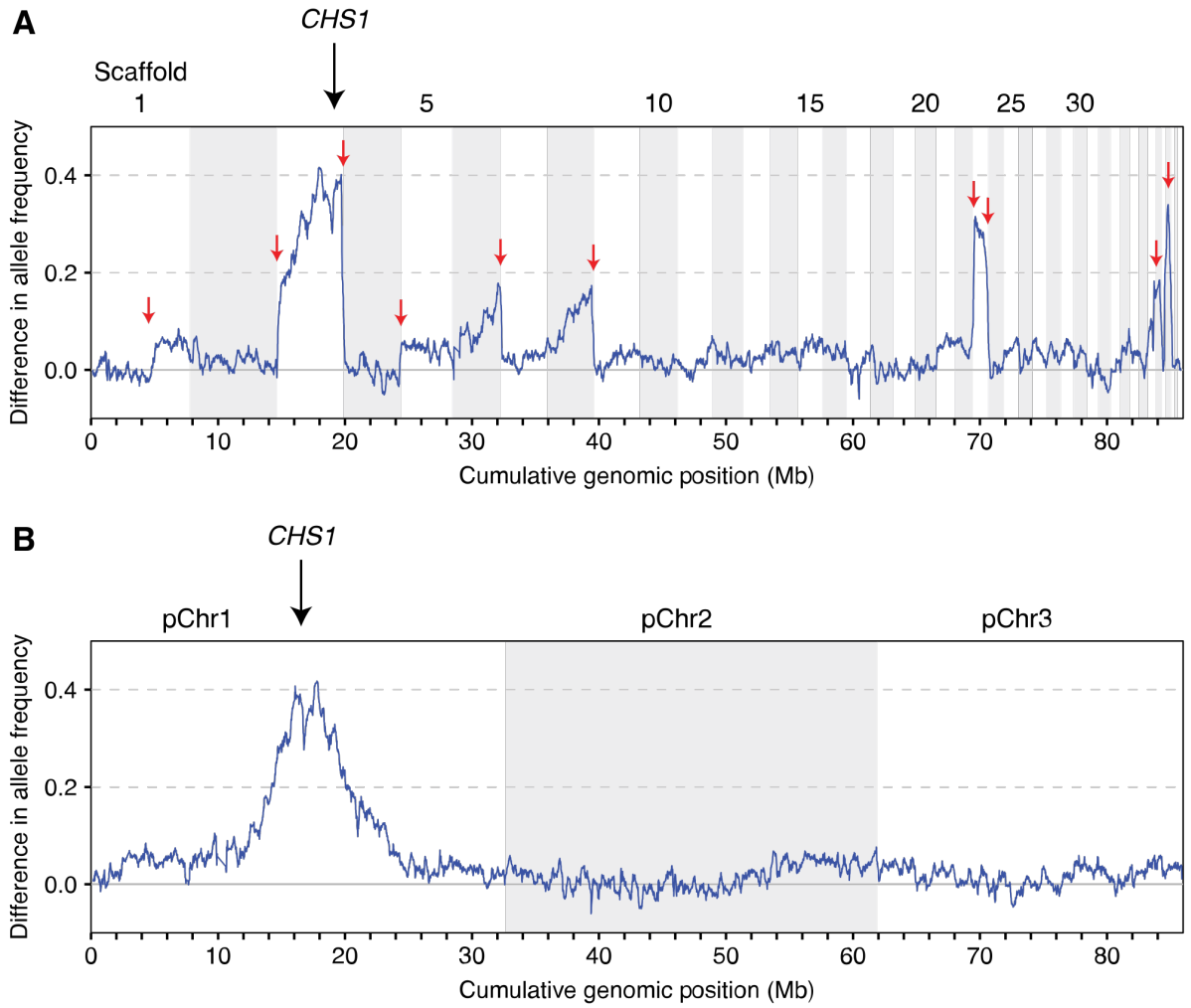


**Figure S5** Reanalysis of population allele frequency data from Van Leeuwen *et al.* (2012). The blue lines indicate the difference in allele frequency toward the etoxazole-resistant parent in an etoxazole-selected population relative to its unselected control population (the cross was from a susceptible *T. urticae* strain and a strain with monogenic, recessive resistance to etoxazole). The analysis was performed with the first 44 Sanger scaffolds (A) as well as with the three pseudochromosomes (B). Compared to data generated for the current study, as well as that generated by Bryon *et al.* (2017), the read data from the Van Leeuwen *et al.* (2012) study was lower-fold coverage, and the read lengths were shorter. Hence, the estimates of allele frequencies are noisier. Nevertheless, discontinuities in allele frequencies apparent in panel A (red arrows) disappear in panel B. The location of *chitin synthase 1* (*CHS1*), which harbors a target-site resistance variant contributed by the resistant strain, is indicated.

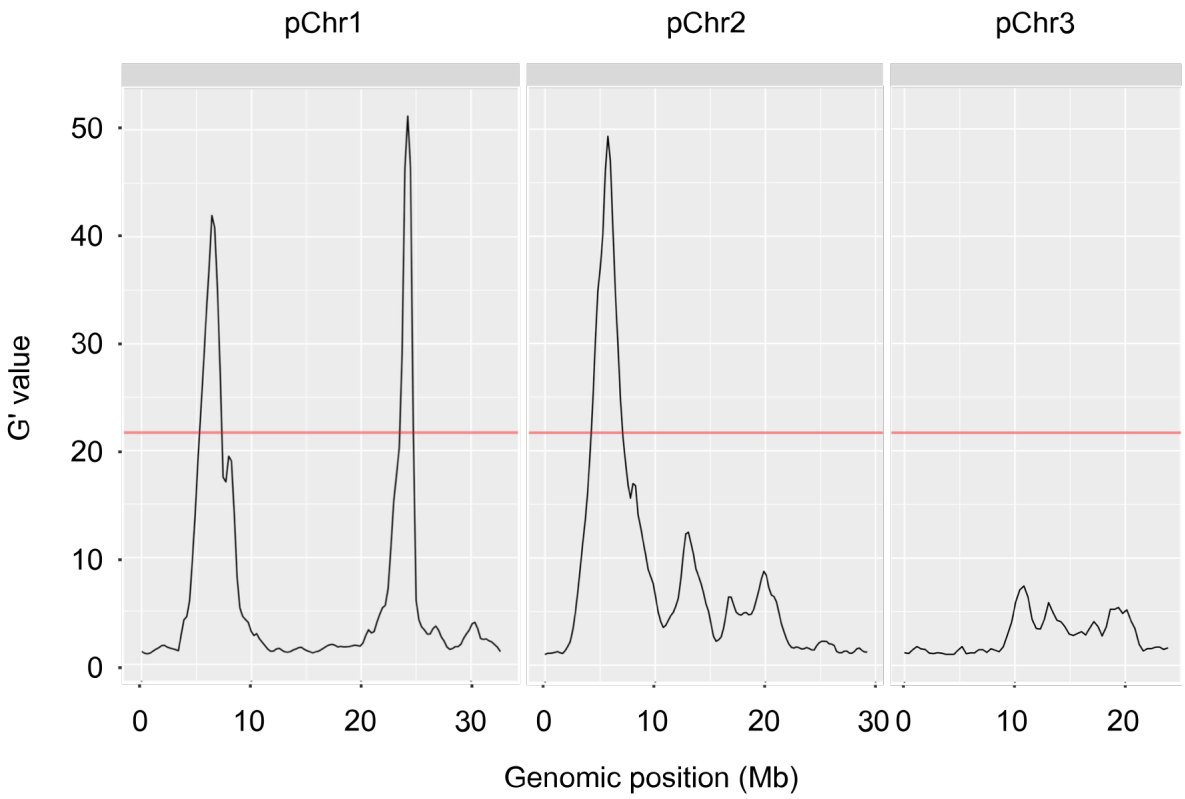


**Figure S6** Three QTL were detected in response to selection by spirodiclofen with QTLseqr (G' method). The black line indicates tricube-smoothed G' values, and the red line indicates a genome-wide FDR of 5%. The three peaks correspond to spiro-QTL 1-3 as also detected by a permutation approach (Figure 6B). Window sizes of 500 kb were used for the analysis.

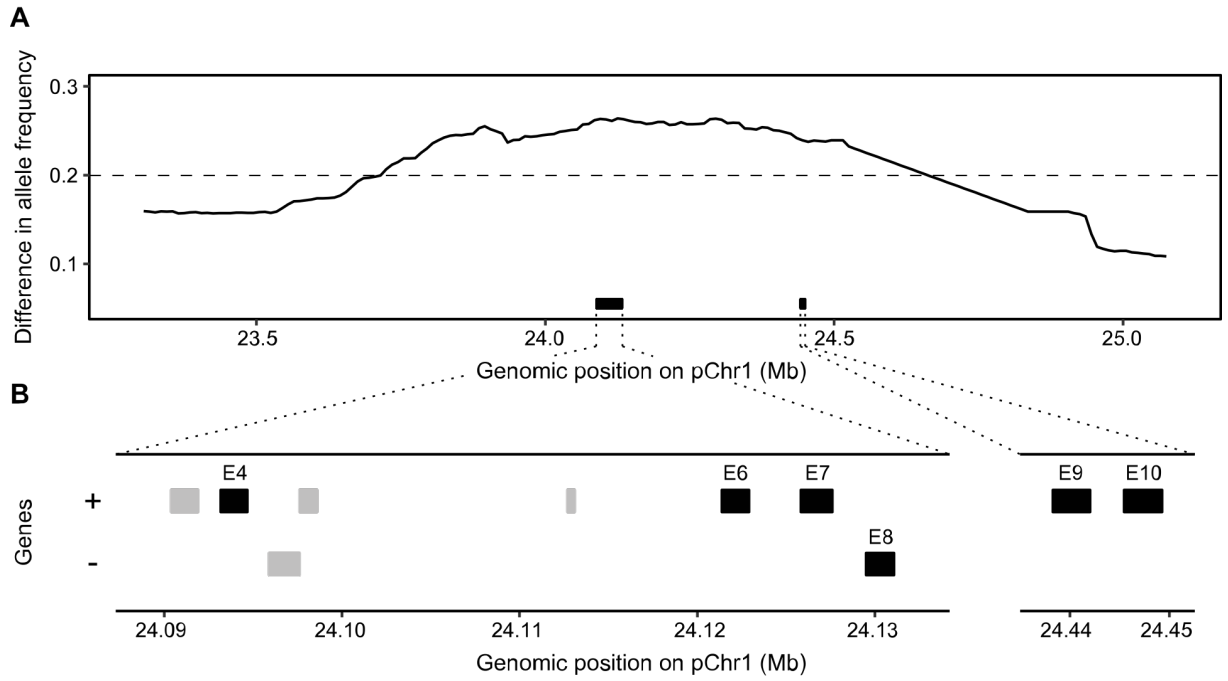


**Figure S7** The spiro-QTL 2 peak covers two clusters of *CYP* genes. The mean change in SR-VP allele frequency across the paired spirodiclofen/control replicates exceeds the 5% FDR threshold for QTL detection across a genomic region of ~960 kb at spiro-QTL 2 (A; see also Figure 6B). (B) This genomic interval harbors two clusters of *CYP* genes, including four pseudogenized/fragmented *CYP* genes (shown in gray) and six intact *CYP* genes of the CYP392 family (*CYP392E4, CYP392E6, CYP392E7, CYP392E8, CYP392E9,* and *CYP392E10*) (shown in black). The + and – signs denote forward and reverse gene orientations.

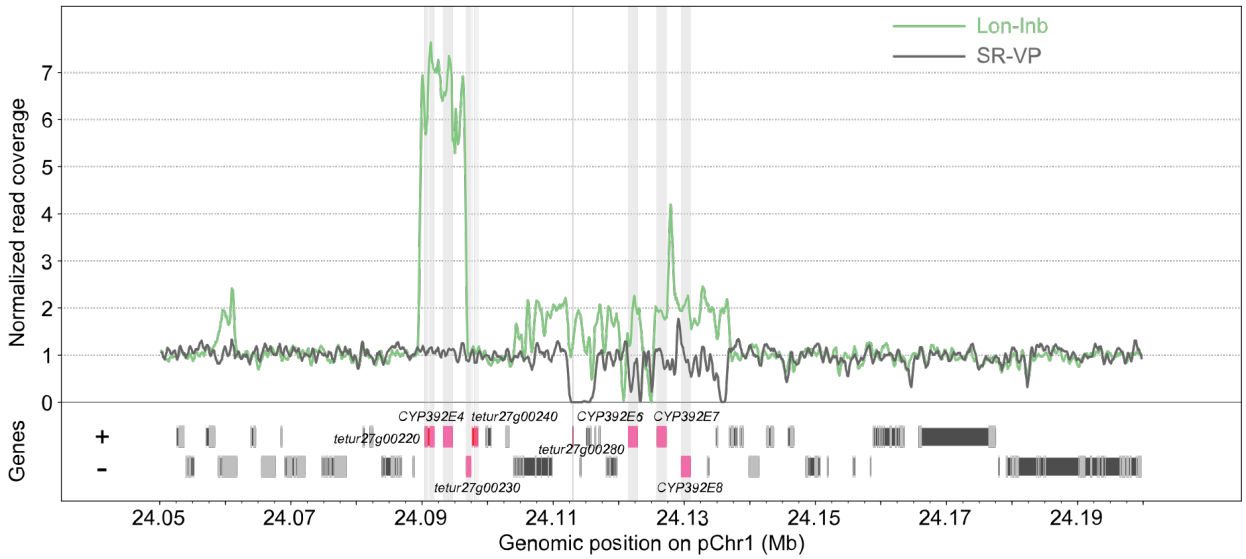


**Figure S8** Copy variation in *CYPs* between strains Lon-Inb and SR-VP at the peak region of spirodiclofen selection response at spiro-QTL 2. *CYP* genes are shown in pink (compare to Figure 6D and the respective legend) with light gray vertical bars denoting *CYP* gene model coordinates in the coverage plot shown at the top (coverage in Illumina short-read data of Lon-Inb, green; coverage of SR-VP, dark gray). Read coverage depth was normalized to the rest of pChr1. A coverage value of ~1.0 is expected for single copy regions, as is observed for Lon-Inb throughout most of the region (excluding *CYPs*), as well as for SR-VP (including for *CYP* sequences).

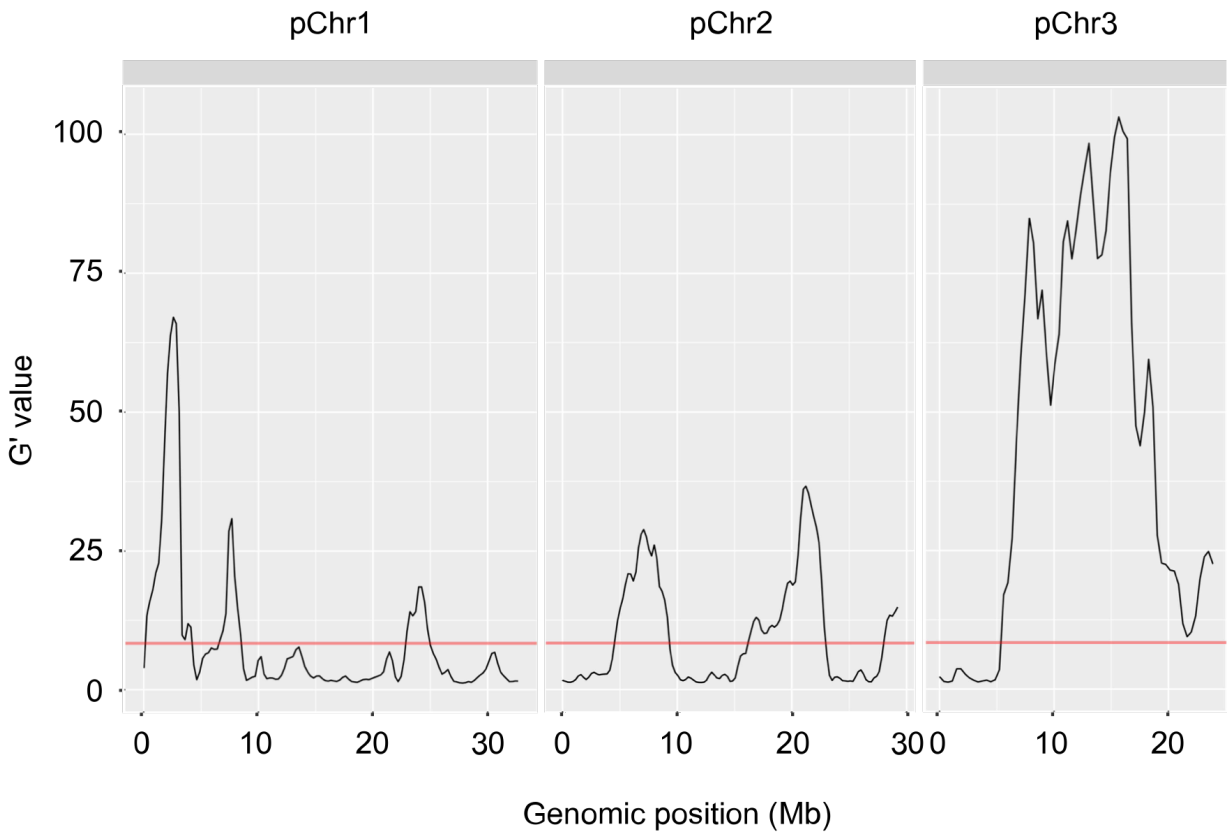


**Figure S9** QTL detected for response to selection on tomato with QTLseqr (G' method). The black line indicates tricube-smoothed G' values, and the red line indicates a genome-wide FDR of 5%. Window sizes of 500 kb were used for the analysis.

## Supplementary References


Bryon A., A. H. Kurlovs, W. Dermauw, R. Greenhalgh, M. Riga, *et al.*, 2017 Disruption of a horizontally transferred phytoene desaturase abolishes carotenoid accumulation and diapause in *Tetranychus urticae*. Proceedings of the National Academy of Sciences 114: E5871–E5880. https://doi.org/10.1073/pnas.1706865114

Ngoc P. C. T., R. Greenhalgh, W. Dermauw, S. Rombauts, S. Bajda, *et al.*, 2016 Complex Evolutionary Dynamics of Massively Expanded Chemosensory Receptor Families in an Extreme Generalist Chelicerate Herbivore. Genome Biology and Evolution 8: 3323–3339. https://doi.org/10.1093/gbe/evw249

Van Leeuwen T., P. Demaeght, E. J. Osborne, W. Dermauw, S. Gohlke, *et al.*, 2012 Population bulk segregant mapping uncovers resistance mutations and the mode of action of a chitin synthesis inhibitor in arthropods. Proceedings of the National Academy of Sciences 109: 4407–4412. https://doi.org/10.1073/pnas.1200068109

Van Leeuwen T., and W. Dermauw, 2016 The molecular evolution of xenobiotic metabolism and resistance in chelicerate mites. Annual Review of Entomology 61: 475–498. https://doi.org/10.1146/annurev-ento-010715-023907