

Figure S1. Corrected *B. lactucae* kinship clustering heatmap. Heatmap used to infer kinship based on SF5 derived markers ≥ 61 bp. This figure differs from Figure 4c because high identity clusters of isolates have each been reduced to single representative isolates.

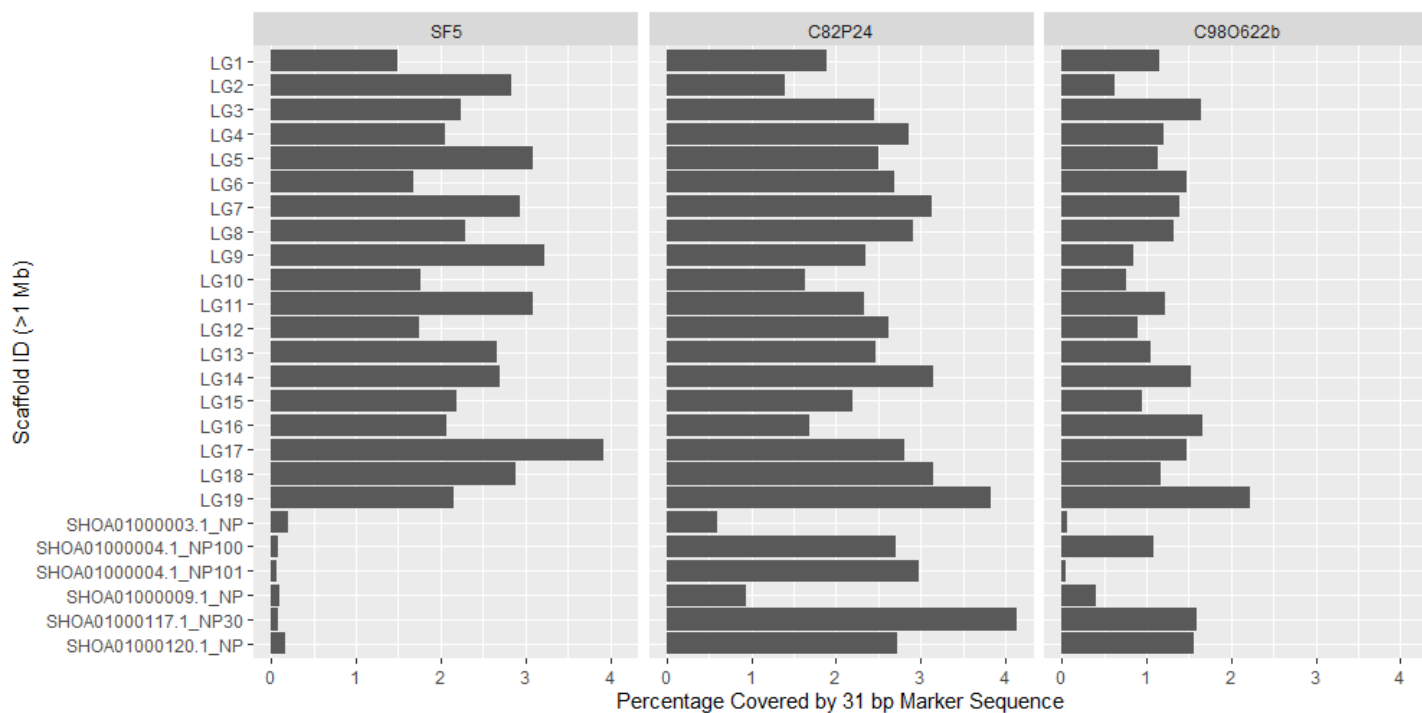


Figure S2. Percent of bases covered by *B. lactucae* pseudo-test cross markers derived from each parent. Markers derived from fragments ≥ 61 bp for SF5 and C82P24 or ≈ 61 bp for C98O622b were mapped to the genetically oriented genome assembly. Percentage coverage of scaffolds larger than 1 Mb was plotted. Very few pseudo-test cross markers derived from SF5 aligned to the unplaced scaffolds (prefixed SHOA) compared to genetically oriented scaffolds (prefixed LG) that did. This pattern did not hold for the other parents; some unplaced scaffolds had at least as much coverage as placed scaffolds. It is therefore likely that these scaffolds represent regions of high homozygosity in SF5 but not in other isolates. Therefore, generation of genetic maps of other isolates of *B. lactucae* will further refine the genome assembly.

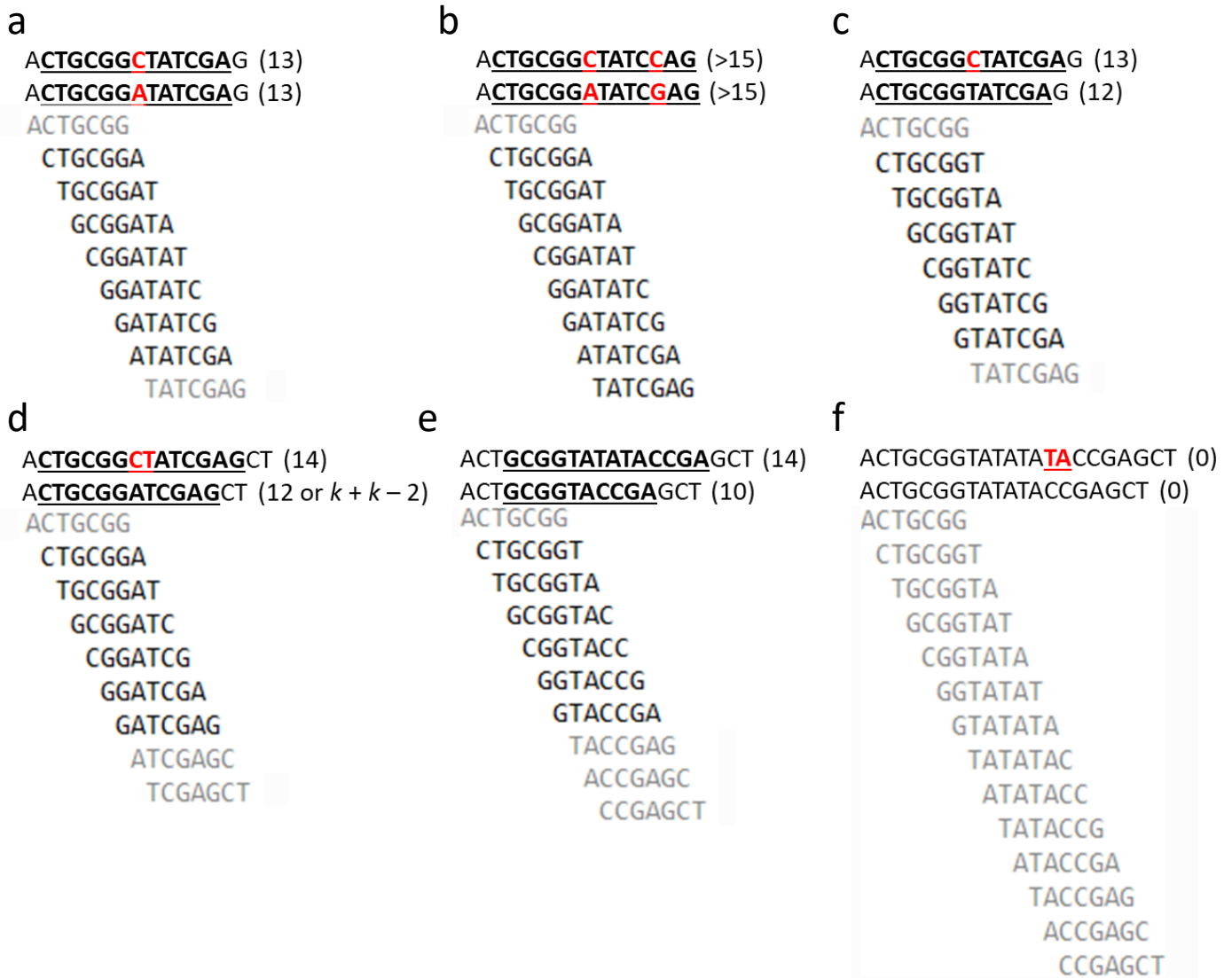


Figure S3. Associating variants with different fragment sizes. AFLAP searches for k -mers segregating uniquely from either parent. To do this efficiently, k -mers are assembled and filtered based on the size of the fragment produced. A single nucleotide variant will produce fragments of $k + k - 1$. For purposes of this illustration $k = 7$, rather than 31 as used in our analysis. Unshaded 7-mers below the sequence are found to be unique to the second sequence. Grey 7-mers are found in both sequences. a) A SNP will result in the retention of 7-mers containing the SNP, 7-mers common to both sequences are filtered out. Using 7-mers the resulting fragment = 13 = $k + k - 1$. b) If two variants are close to one another then larger fragments are assembled because the 7-mers retained contain multiple variants. In this example if no new variants were downstream then the six bases before the C/A variant and six bases after the C/G variant would be retained to assemble a fragment greater than $k + k - 1$. c) Indels may result in fragments equaling less than $k + k - 1$. A single nucleotide deletion results in the six base pairs either side being assembled into a unique 12 bp fragment. 7-mers which contain the nucleotide would still assemble to 13 bp. d) When a dinucleotide deletion is observed a 12 bp fragment is still obtained, 7-mers with the nucleotide now assembled into 14 bp fragments. e) When microsatellite variation is considered much smaller fragments than $k + k - 1$ may be assembled as repetitive sequence leads to the detection of common k -mers. f) If the TA repeat was repeated four times and only one copy deleted, then a 7-mer survey would not detect any variants.

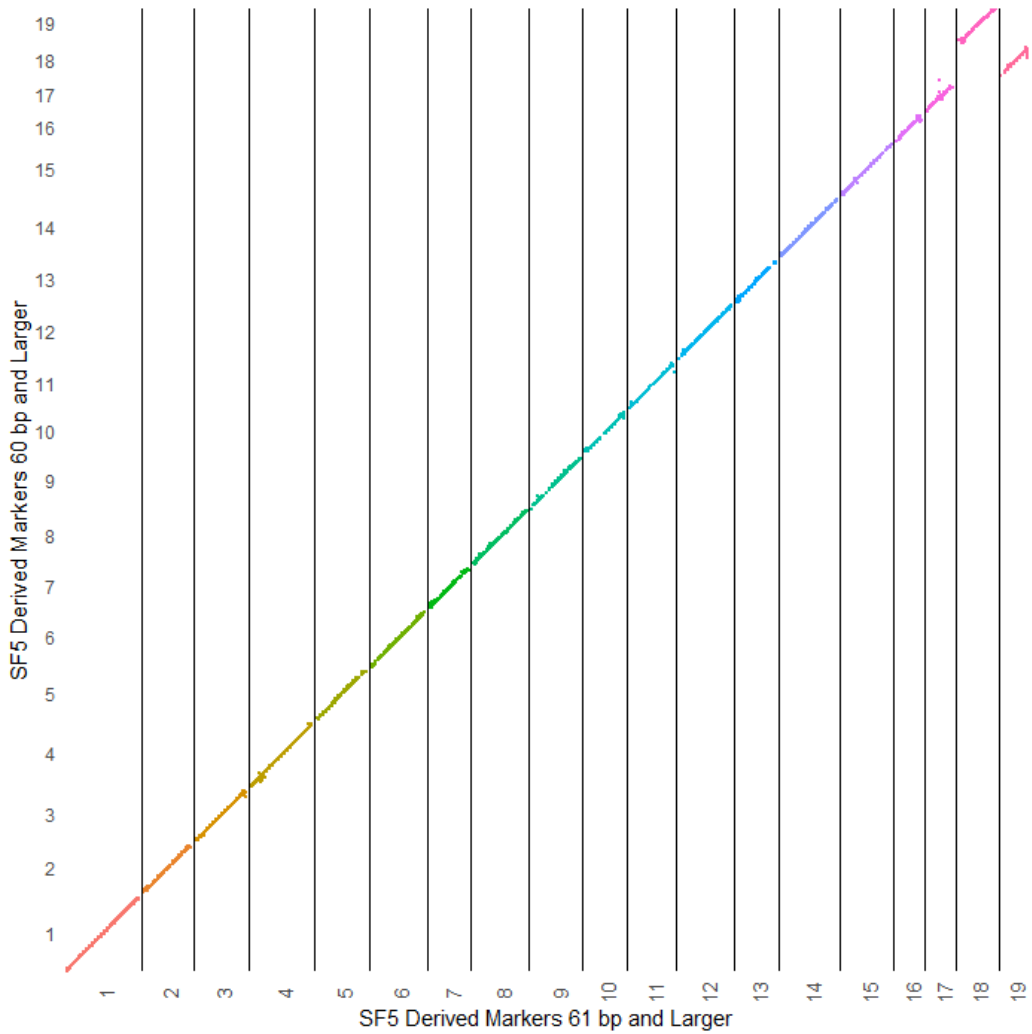


Figure S4. Addition of markers derived from 60 bp fragments to the SF5 genetic map. The addition of 4,070 markers derived from 60 bp fragments had little effect on the map as the shared marker order was not altered and the same 19 linkage groups were produced. The switched ordering of linkage group 18 and 19 in the plot is due to more markers derived from 60 bp fragments being assigned to linkage group 19 (1,838 to 1,923) than 18 (1,858 to 1,919).