Ultra-high density genetic maps based on whole genome shotgun sequencing provide a useful scaffold for the assembly of large and highly repetitive plant genomes. We have constructed an ultra-high density genetic map for sunflower, which is the target on an ongoing sequencing effort. We sequenced Helianthus annuus cultivars RH4280 and RH4281, which are the parental lines of our core recombinant inbred line (RIL) mapping population, to an average depth of 10x. Alignment to the contiguous sequences of the draft sunflower genome assembly revealed over two million high-quality SNPs fixed between the two parental cultivar lines. Ninety-six eighth-generation RILs were sequenced to a mean depth of 1.0x. In the RILs, genetic regions were called as descended from one or the other parent based on the presence of at least ten SNP calls at informative sites. Perfectly correlated regions not exhibiting significant segregation distortion were binned for use as genetic markers. A genetic map of seventeen linkage groups, corresponding to the seventeen H. annuus chromosomes was constructed de novo using a Minimum Spanning Tree algorithm as implemented in the MSTmap software. This map allowed us to confirm the composition of genomic scaffolds, assign them to linkage groups, and place them in linear order.

Ultra-high density genetic maps were based on the minimum sum of recombination events (MSTmap) between their segregation patterns and divides them into linkage groups if the sum is significantly different than observed across all markers. Markers on each linkage group are then ordered using a recursive minimum spanning tree algorithm. Kosambi’s mapping function was used to calculate the map distance between adjacent pairs of markers ordered by MSTmap.

Contigs containing segregating SNPs were compared to the template map in forward and reverse order and the best match was stored for each direction. A contig was placed with an upper distance of the best forward match and a lower distance of the reverse match if both were found on the same linkage group. 243,048 contigs were placed within 5 cM and 261,999 contigs were placed within 9 recombinations.

The probe sequences from a previous illumina map were matched by BLAST to the contigs in the sunflower assembly. The centimorgan positions of the positions on the two maps were compared (Fig. S). The two maps agreed very well in terms of synteny and ordering even though they were completely independently constructed. The micro and macro orders seem very well conserved, considering the illumina map was in part constructed on a different mapping population. The chromosomes from the sequence based maps can now be named and oriented relative to the previous literature.

Single copy illumina mapped probes Blasted to sequence contigs and plotted by map position on illumina map vs. map position on sequence map. 4093 Contigs mapped, 6064 single copy illumina probes mapped to 6298 contigs. Of the 6298 contigs 4093 had been mapped based on sequence position. 90.4% of the hits are in 32 syntenic blocks. The roughly 10% non-syntenic hits can be explained by blast picking the second best hit if the true homologue is fragmented into several contigs, or if the sequence is multi copy.