

Draft Assembly of the Sunflower Genome

Loren Rieseberg¹, Nolan Kane¹, Chris Grassa¹, Navdeep Gill¹, John Bowers², Zhao Lai³, Helene Berges⁴, Jerome Gouzy⁵, Eleni Bachlava⁶, Matthew King¹, Nicolas Langlade⁵, John Burke², Patrick Vincourt⁵, and Steve Knapp⁶

¹Botany Department, University of British Columbia, Vancouver, BC V6R 2A5, Canada, lriesebe@mail.ubc.ca

²Department of Plant Biology, University of Georgia, Athens, GA 30602 USA, jmburke@uga.edu

³Department of Biology, Indiana University, Bloomington, IN 47405, USA, zlai@cgb.indiana.edu

⁴French Plant Genomic Resource Center, INRA - CNRGV, BP 52627, 31326 Castanet Tolosan, France, Helene.Berges@toulouse.inra.fr

⁵Laboratoire des Interactions Plantes Micro-organismes, BP 52627, 31326 Castanet Tolosan, France, patrick.vincourt@toulouse.inra.fr

⁶Monsanto, Co., 2790 Gorman Lane, Woodland, CA 95695, USA, steve.knapp@monsanto.com

ABSTRACT

- Here we report on a draft sequence of the 3.6 Gb genome of cultivated sunflower (*Helianthus annuus*), the most economically important crop in the Compositae family.
- The genome was put together using a combination of 24x 454 whole-genome shotgun (WGS) sequencing and 80x Illumina WGS, along with ultra-high density genetic and physical maps that serve as scaffolds for the linear assembly of WGS sequences.
- Our genetic map, which was generated by sequencing 96 recombinant inbred lines to 1x depth, includes 2.6 million SNPs and covers 1361 cM.
- Likewise, our sequence-based physical map covers 97% of the sunflower genome in circa 3,300 contigs, with unique sequence tags every 3-6 kb.
- Assemblies using ABySS, CABOG, Newbler and SOAPdenovo will be compared, and new methods of dealing with highly repetitive sequences will be discussed.

Keywords – genetic map - genome assembly - physical map – sunflower genome - whole genome shotgun sequencing