A contiguous de novo genome assembly of sugar beet EL10 (Beta vulgaris L.)

J Mitchell McGrath[1], Andrew Funk[2], Paul Galewski[2], Shujun Ou[2], Belinda Townsend[3], Karen Davenport[4], Hajnalka Daligault[4], Shannon Johnson[4], Joyce Lee[5], Alex Hastie[5], Aude Darracq[6], Glenda Willems[6], Steve Barnes[6], Ivan Liachko[7], Shawn Sullivan[7], Sergey Koren[8], Adam Phillippy[8], Jie Wang[9], Tiffany Liu[9], Jane Pulman[9], Kevin Childs[9], Shengqiang Shu[10], Anastasia Yocum[11], Damian Fermin[11], Effie Mutasa-Göttgens[12], Piergiorgio Stevanato[13], Kazunori Taguchi[14], Rachel Naegele[1], Kevin Dorn[15]. [1]USDA-ARS, 1066 Bogue St., Michigan State Univ., East Lansing, MI USA. [2]Plant Breeding, Genetics, and Biotechnology Program, Michigan State Univ., East Lansing, MI USA. [3]Department of Plant Sciences, Rothamsted Research, West Common, Harpenden, Hertfordshire AL5 2JQ, UK. [4]BioNano Genomics, 9640 Towne Centre Drive, San Diego, CA USA. [5]Los Alamos Nat'l Lab, Biosecurity and Public Health, Los Alamos, NM USA. [6]SESVANDERHAVE N.V., Industriepark Soldatenplein Zone 2 Nr 15, 3300 Tienen, Belgium. [7]Phase Genomics, 4000 Mason Road, Suite 225, Seattle, WA USA. [8]Genome Informatics Section, Computational and Statistical Genomics Branch, National Human Genome Research Institute, Bethesda, Maryland, USA. [9]Center for Genomics-Enabled Plant Science, Plant Biology Department, Michigan State. Univ., East Lansing, MI USA. [10]United States Department of Energy, Joint Genome Institute, Berkeley CA USA. [11]A2IDEA, 674 S. Wagner Rd., Ann Arbor, MI USA. [12]University of Hertfordshire, Hatfield, Hertfordshire AL10 9AB UK. [13]DAFNAE, University of Padova, Viale Università 16, 35020 Legnaro (PD), Italy. [14]Hokkaido Agricultural Research Center, National Agriculture and Food Research Organization, Shinsei Memuro, Hokkaido, 082-0081 Japan. [15]USDA-ARS, 1701 Centre Ave, Crops Research Laboratory, Fort Collins, CO USA.
Corresponding Author(s): J. Mitchell McGrath (mitchmcg@msu.edu) and Kevin Dorn (kevin.dorn@usda.gov)

ABSTRACT

A contiguous assembly of the inbred 'EL10' sugar beet (*Beta vulgaris* ssp. *vulgaris*) genome was constructed using PacBio long read sequencing, BioNano optical mapping, Hi-C scaffolding, and Illumina short read error correction. The EL10.1 assembly was 540 Mb, of which 96.7% was contained in nine chromosome-sized pseudomolecules with lengths from 52 to 65 Mb, and 31 contigs with a median size of 282 kb that remained unassembled. Gene annotation incorporating RNAseq data and curated sequences via the MAKER annotation pipeline generated 24,255 gene models. Results indicated that the EL10.1 genome assembly is a contiguous genome assembly highly congruent with the published sugar beet reference genome. Gross duplicate gene analyses of EL10.1 revealed little large-scale intra-genome duplication. Reduced gene copy number for well-annotated gene families relative to other core eudicots was observed, especially for transcription factors. Variation in genome size in *B. vulgaris* was investigated by flow cytometry among 50 individuals producing estimates from 633 to 875 Mb/1C. Read depth mapping with short-read whole genome sequences from other sugar beet germplasm suggested that relatively few regions of the sugar beet genome appeared associated with high-copy number variation.

RE-PRINT