

Michigan State University

Report to NCCC-212, November 2020

ANNOUNCEMENTS:

1. New collaborators needed! If you have access to galls from the blueberry stem gall wasp and interested in collaborating with us, please contact Pat Edger (edgerpat@msu.edu). Additional information regarding this insect pest can be found here: https://www.canr.msu.edu/news/blueberry_gall_wasp_management_guidelines

OBJECTIVE I – DEVELOP IMPROVED FRUIT GERMLASM THROUGH COOPERATIVE BREEDING

Blueberry: Molecular breeding efforts in blueberry have been slow due in large part to the lack of genetic information and genomic resources. Previous studies have developed genetic and genomic resources for a wild diploid species ($2n=2x=24$) of blueberry (Gupta et al., 2015). However, this draft genome has a large number of scaffolds (13,757 total; N50 of ~145kb), high percentage of gaps (~27.35%) in ~393.16Mb assembly, and most importantly, does not reflect the genome complexity of the economically important and tetraploid ($2n=4x=48$) highbush blueberry. Last year, we published a chromosome-scale genome for the tetraploid northern highbush blueberry (Colle et al., 2019). The haplotype-phased assembly consists of 48 chromosome-length pseudomolecules with 1,679,081,592 bases of assembled sequence, ~1.29% gaps, and an average of 32,140 protein coding genes per haplotype (128,559 total). Benchmarking Universal Single-Copy Orthologs (BUSCO) analysis was performed to assess the completeness of the assembly and quality of the genome annotation. The annotated gene set contains 1,413 out of 1,440 (98%) BUSCO genes. The availability of this draft genome will facilitate the discovery and analysis of genes encoding economically important traits and accelerate genome-guided breeding efforts. We leveraged this genome, combined with gene expression and metabolite data measured across fruit development, to identify candidate genes involved in the biosynthesis of important phytonutrients among other metabolites associated with superior fruit quality.

Genomic resources for blueberry are publicly available: [Dataset](#) and [Genome Browser](#).

Current work: As part of a USDA SCRI project (VacCAP; <https://www.vacciniumcap.org/>), we are currently assembling a pangenome for Northern Highbush (NHB), Southern Highbush (SHB), and cranberry (CB). For each crop, 12 genotypes that are highly representative of the pedigree of NHB, SHB and CB cultivars and that capture the greatest amount of genetic diversity were selected. Our preliminary data suggests that roughly 30% of genes are unique to individual blueberry cultivars, due to gene duplication and loss patterns across the genome. Pangenome analyses will be conducted to identify the variable gene content that contributes to fruit quality differences between various cultivars. A Vaccinium genotyping platform will be developed that will target both the core (80%) and dispensable (20%) portions of the pangenome for NHB, SHB and CB.

Strawberry: A high-quality reference genome for the octoploid to serve as a platform for identifying agriculturally important genes and applying genomic-enabled breeding approaches. The assembly of the octoploid strawberry genome, with an estimated genome size of 813.4 Mb, has been particularly challenging due to its high heterozygosity and ploidy level. We recently published a chromosome-scale genome for the cultivated octoploid strawberry (*Fragaria x ananassa*) (Edger et al., 2019). The total length of the final assembly is 805.5Mb distributed across 28 chromosome-level pseudomolecules, plus 408.2Mb of haplotype variants. We annotated 108,087 protein-coding genes along with 30,703 long non-coding RNA (lncRNA) genes. Gene annotation and genome assembly quality were evaluated using the Benchmarking Universal Single-Copy Orthologs (BUSCO) method. Most (99.17%) of the 1,440 core genes were identified in the annotation, supporting a high-quality genome assembly and annotation. Furthermore, we identified genes that encode various important traits, including metabolites associated with fruit quality and disease-resistance.

Genomic resources for octoploid strawberry are publicly available: [Dataset](#) and [Genome Browser](#).

We previously also published a near-complete genome of diploid woodland strawberry (*Fragaria vesca*) (Edger et al., 2018). This genome assembly (220.8 Mb total) has a contig N50 length of ~7.9 million base pairs (Mb), representing a ~300-fold improvement of the previous version. The vast majority (>99.8%) of the assembly was anchored to 7 pseudomolecules. We obtained ~24.96 Mb of sequence not present in the previous version of the *Fragaria vesca* genome and produced an improved annotation that includes 1496 new genes. Genomic analyses uncovered numerous, large-scale scaffolding errors present in each chromosome in the previously published version of the *F. vesca* genome. This new version of the genome has already been leveraged to identify resistance locus to Fusarium wilt in octoploid strawberry (Pincot et al., 2018).

Genomic resources for diploid *F. vesca* are publicly available: [Dataset](#) and [Genome Browser](#).

We also recently published a chromosome-scale genome for another diploid progenitor (*Fragaria iinumae*) of the cultivated strawberry (Edger et al., 2020).

Genomic resources for diploid *F. iinumae* are publicly available: [Dataset](#) and [Genome Browser](#).

Lastly, using these new genomic resources, two new genotyping arrays (850K and 50K) have been developed for strawberry – please see (Hardigan et al., 2020) for details.

Collaborative Projects with NCCC-212 Members

“VacciniumCAP: Leveraging genetic and genomic resources to enable development of blueberry and cranberry cultivars with improved fruit quality attributes” was funded by NIFA-SCRI from 2019-2024. The project involves multiple NCCC-212 committee members.

References:

1. Colle, M., C.P. Leisner, C.M. Wai, S. Ou, K.A. Bird, J. Wang, J.H. Wisecaver, A.E. Yocca, E.I. Alger, H. Tang, Z. Xiong, P. Callow, G. Ben-Zvi, A. Brodt, K. Baruch, T. Swale, L. Shiue, G. Song, K.L. Childs, A. Schillmiller, N. Vorsa, C.R. Buell, R. VanBuren, N. Jiang, and P.P. Edger. 2019. Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. *GigaScience* 8(3), giz012.
2. Edger, P.P., T.J. Poorten, R. VanBuren, M.A. Hardigan, M. Colle, M.R. McKain, R.D. Smith, S.J. Teresi, A.D.L. Nelson, C.M. Wai, E.I. Alger, K.A. Bird, A.E. Yocca, N. Pumplun, S. Ou, G. Ben-Zvi, A. Brodt, K. Baruch, T. Swale, L. Shiue, C.B. Acharya, G.S. Cole, J.P. Mower, K.L. Childs, N. Jiang, E. Lyons, M. Freeling, J.R. Puzey, and S.J. Knapp. 2019. Origin and evolution of the octoploid strawberry genome. *Nature Genetics* 51(3), 541.
3. Edger, P.P., R. VanBuren, M. Colle, T.J. Poorten, C.M. Wai, C.E. Niederhuth, E.I. Alger, S. Ou, C.B. Acharya, J. Wang, P. Callow, M.R. McKain, J. Shi, C. Collier, Z. Xiong, J.P. Mower, J.P. Slovin, T. Hytonen, N. Jiang, K.L. Childs, and S.J. Knapp. 2018. Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. *GigaScience* 7(2), gix124.
4. Patrick P Edger, Michael R McKain, Alan E Yocca, Steven J Knapp, Qin Qiao, Ticao Zhang. 2020. Reply to: Revisiting the origin of octoploid strawberry. *Nature Genetics* 52 (1), 5-7.
5. Pincot, D.D.A., T.J. Poorten, M.A. Hardigan, J.M. Harshman, C.B. Acharya, G.S. Cole, T.R. Gordon, M. Stueven, P.P. Edger, and S.J. Knapp. 2018. Genome-Wide Association Mapping Uncovers Fw1, a Dominant Gene Conferring Resistance to Fusarium Wilt in Strawberry. *G3: Genes, Genomes, Genetics*. 8(10), 1817.
6. Michael A Hardigan, Mitchell J Feldmann, Anne Lorant, Kevin A Bird, Randi Famula, Charlotte Acharya, Glenn Cole, Patrick P Edger, Steven J Knapp. 2020. Genome synteny has been conserved among the octoploid progenitors of cultivated strawberry over millions of years of evolution. *Frontiers in Plant Science* 10, 1789.