Graph pangenome of eggplant and (its) allied species

Luciana Gaccione, Marie Bolger, Laura Toppino, Ezio Portis, Sergio Lanteri, Giuseppe Leonardo Rotino, Maria Rosaria, Jaime Prohens, David Alonso, Giuseppe Aprea, et al.
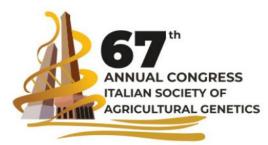
HAL Id: hal-04670168

https://hal.inrae.fr/hal-04670168v1

Submitted on 11 Aug 2024

67th

ANNUAL CONGRESS
ITALIAN SOCIETY OF
AGRICULTURAL GENETICS

Bologna, 10-13 September 2024

EXPANDING FRONTIERS IN CROP GENETICS

The Italian Society of Agricultural Genetics holds its LXVII Annual Congress in Bologna, September 10-13, 2024.

Join us to enjoy a broad and stimulating scientific programme, prestigious international invited speakers, and the historical downtown of Bologna.

Organised by

ALMA MATER STUDIORUM
UNIVERSITA DI BOLOGNA

crea
Consiglio per la ricerca in agricoltura
e l'analisi dell'economia agraria

Dipartimento di Scienze e Tecnologie Agro-Alimentari

Centro di Ricerca Cerealicoltura e Colture Industriali

Graph pangenome of eggplant and (its) allied species

Luciana Gaccione[1], Marie Bolger[2], Laura Toppino[3], Ezio Portis[1], Sergio Lanteri[1], Giuseppe Leonardo Rotino[3], Maria Rosaria Tassone[3], Jaime Prohens[4], David Alonso[4], Giuseppe Aprea[5], Paola Ferrante[5], Veronique Lefebvre[6], Emily Idahl[7], Maximilian Schmidt[8], Hatice Filiz Boyaci[9], Giovanni Giuliano[5], Björn Usadel[2,7], Lorenzo Barchi[1]

[1]DISAFA, Plant Genetics, University of Turin, 10095 Grugliasco (TO), Italy
[2]IBG-4, CEPLAS, Forschungzentrum Jülich, 52428 Jülich, Germany
[3]CREA, Research Centre for Genomics and Bioinformatics, 26836 Montanaso Lombardo (LO), Italy
[4]Universitat Politècnica de València, 46022 Valencia, Spain
[5]ENEA, Casaccia Res Ctr, 00123 Roma, Italy
[6]INRAE, GAFL, 84140 Montfavet, France
[7]Institute for Biological Data Sciences, Faculty for Mathematica and Natural Sciences, HHU, 40225 Düsseldorf, Germany
[8]Plant Breeding Department, Geisenheim University, 65366 Geisenheim, Germany
[9]Faculty of Agriculture, University of Recep Tayyip Erdogan, Rize, Türkiye

Recent advances in sequencing technologies and the reduction of their costs have accelerated the assembly of plant genomes. However single reference genomes do not fully capture the species diversity, leading to the development of the pangenome concept. A pangenome encompasses all DNA sequences available for a species or larger taxa, including single nucleotide polymorphisms (SNPs) and small/large structural variants (SVs).

To better characterize these variations, the development of graph-based pangenomes have emerged. They store sequence and SV information in nodes and paths, thereby expanding the linear reference genome's coordinate system to more effectively capture genetic diversity.

In this study, Oxford Nanopore long reads from 40 accessions, including 33 accessions of *Solanum melongena* among which the reference line '67/3', as well as 7 accessions of the allied species *S. insanum* and *S. incanum* were assembled using NextDenovo. The resulting assemblies were polished with Medaka and chromosome-scale assemblies were generated for six *S. melongena* and two accessions each of *S. insanum* and *S. incanum*, by incorporating chromosome conformation capture (Hi-C) data. The chromosomes of the other 30 accessions were reconstructed using RagTag, based on the reference line '67/3' sequence. A combination of Helixer (Deep Neural Networks *ab initio* gene prediction) and Braker-3 (both *ab initio* and evidence-based gene prediction) was applied to annotate each genome, resulting in a gene set ranging from 30,886 to 33,449, with a BUSCO completeness from 95.3 to 97.2%.
The obtained assemblies were used to build up two different graph-based pangenomes. The first one, which includes the 40 accessions, was built using the PanGenome Graph Builder (PGGB) pipeline. The second one, comprising only the 33 *S. melongena* accessions, was obtained with Minigraph-Cactus pangenome pipeline. The latter was then used for genotyping 311 *S. melongena* accessions from the EU G2P-SOL project core collection, which were sequenced with Illumina short read sequencing. Overall, 4,532 high-quality SVs were identified.

The accession panel was phenotyped, among others, for resistance to fungal wilts at two sites: Montanaso Lombardo (Italy), and Antalya (Türkiye), and associations between phenotype and genotype were identified using BLINK model within GAPIT R package, on multi-environment data. Here, we report the genome-wide association (GWA) results for resistance to *Fusarium oxysporum* f. sp. *melongenae* and *Verticillium dahliae*.

Our results highlight that the graph pangenome provides an efficient model for representing and analyzing multiple genomes and their variations. This offers a comprehensive resource for eggplant breeding programs and enhances the understanding of eggplant genetic architecture of traits, particularly in relation to fungal wilt resistance.