

Genome Assembly of the Diploid Self Compatible Potato Cultivar Solyntus

Authors: Richard Finkers, Ate van der Burg, Menno ter Maat, Natascha van Lieshout, Martijn van Kaauwen, Danny Esselink, Fernanda A. de F. Guedes, Pim Lindhout, Michiel de Vries Richard Visser



Background

The potato is one of the most important food crops worldwide. Improvements to its traits can therefore have a major impact. Reading the genome structure of the potato is extremely tricky, however, as a regular potato consists of four genomes, which makes it difficult to determine the position of the genes. The recent research applied a diploid real potato plant with only one genome, a so-called homozygote, which makes it easier to read and compare the DNA base sequence. This line, Solyntus, was produced as part of Solynta's hybrid potato breeding programme.

Objective

- Develop a *de novo* of the diploid breeding line Solyntus

Results

Solyntus genome assembly

Smartdenovo was used to assembly approx. 100x Oxford Nanopore genomics reads to a genome assembly. Heterozygous regions were flattened using the purge haplotigs pipeline. The assembly was further polished by a combination of polishing tools such as racon (ONT reads) and ntEdit (Illumina reads) and scaffolded to pseudo molecules using RaGOO. This resulted in v1.0 of the assembly. Critical evaluation and basecalling ONT reads with an improved basecalling algorithm, further improved assembly statistics (version solynta_03_ph; Table 1).

Statistics without reference	v1.0	solynta_03_ph
# contigs	185	174
# contigs (>= 0 bp)	185	174
# contigs (>= 1000 bp)	185	174
# contigs (>= 5000 bp)	185	174
# contigs (>= 10000 bp)	185	174
# contigs (>= 25000 bp)	185	174
# contigs (>= 50000 bp)	184	173
Largest contig	27 013 815	23 312 788
Total length	716 118 559	732 026 319
Total length (>= 0 bp)	716 118 559	732 026 319
Total length (>= 1000 bp)	716 118 559	732 026 319
Total length (>= 5000 bp)	716 118 559	732 026 319
Total length (>= 10000 bp)	716 118 559	732 026 319
Total length (>= 25000 bp)	716 118 559	732 026 319
Total length (>= 50000 bp)	716 068 940	731 984 064
N50	7 550 499	8 060 818
N75	4 023 563	3 722 267
L50	27	26
L75	60	60
GC (%)	34.58	34.89
Mismatches		
# N's	0	0
# N's per 100 kbp	0	0

Table 1. Genome assembly statistics of the releases Solyntus assembly (v1.0) and the intermediate version currently in the assembly pipeline (solynta_03_ph).



Figure 1. Diploid potato breeding program

Genome completeness

Genome completeness has been assessed using BUSCO using the Solanaceae reference set. BUSCO completeness was 92,7% (88.9% single copy) while 2,9% of the BUSCO genes were fragmented.

Annotation

Gene Model Mapper (GeMoMa) was used to develop an initial annotation. Sequences of protein coding genes of Potato, Tomato and Pepper were used to infer protein coding genes in Solyntus; annotating approx. 75K genes.

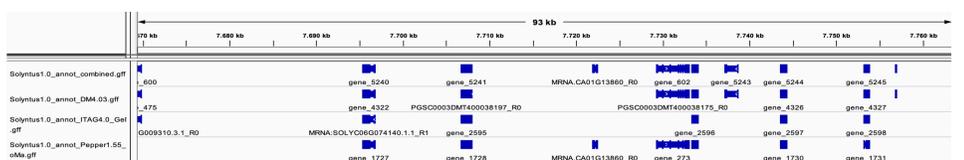


Figure 2. IGV track showing the individual inferred protein coding genes and a combined track. Most genes were annotated from all three genomes, though genome specific inference was found as well.

Availability

The genome sequence is available for download at <https://www.plantbreeding.wur.nl/Solyntus>. The plant material is available for purposes under MTA and can be requested by contacting: Michiel de Vries (michiel.devries@solynta.com)

Future Work

- Further scaffold the Solyntus genome
- *Ab initio* annotate the Solyntus genome sequence
- Further develop the genome resources for Solyntus at <https://www.plantbreeding.wur.nl/Solyntus>

Acknowledgements

The project is a collaboration between Solynta and Wageningen UR Plant Breeding. The research was partly financed by PepsiCo, contributing to a more sustainable food chain. SolyntusTM is a trademark that results from Solynta's hybrid potato breeding programme.