BITS :: Call for Abstracts 2023 - Oral communication

Type	Oral communication
Session	Genomics, transcriptomics, epigenomics and epitranscriptomics
Title	Reference genome assembly and annotation of an African reference cassava genome for identification of epigenetics variations
All Authors	Michael Landi (1,2), Trushar Shah (2), Adnan Niazi (1), Livia Stavolone (2,4), Laurent Falquent (3), Erik Bongcam-Rudloff (1), Andreas Gisel (2,5)

Affiliation

(1) Swedish University of Agricultural Sciences,
(2) International Institute of Tropical Agriculture
(3) University of Fribourg
(4) Institute of Sustainable Plant Protection, CNR
(5) Institute of Biomedical Technologies, CNR

Motivation

Cassava (Manihot esculenta) is a crucial tropical root crop that provides a significant dietary energy source for over 500 million people in tropical and subtropical regions. As these areas' population increases rapidly, cassava’s importance as a food source is becoming even more critical. Further, several studies have attributed cassava to growing under hostile and dry conditions, making it an essential crop for regions where other crops struggle to strive, an important crop, especially in the light of climate change. However, yield improvement in farmer’s fields rises slowly due to biotic and abiotic stress, limits of classical breeding, and the fact that the crop is clonally propagated and probably accumulates a negatively influencing load of DNA methylations. Up to now production increase of cassava in many regions is due to the increase of the cultivated area, often deforesting essential sources of local biodiversity. Biotechnological technologies offer solutions to the mentioned limitations, but despite its importance, the cassava genome still needs to fully deciphered due to its high heterozygosity and intricate genetic makeup. Recent breakthroughs in high-fidelity (HiFi) sequencing technology and advanced assembly techniques have enabled us to generate a haplotype-resolved genome of an important African cassava cultivar.

Methods

We utilized PacBio HiFi reads to sequence the TMEB117 cassava cultivar, a popular farmer cultivar for its high yields, excellent root quality, and desirable taste. After assembling the reads, we accessed the genome's quality, completeness, and consensus accuracy. We performed gene and repeat annotation and gene family analysis to understand the genome’s organization comprehensively. In addition, we conducted a comparative study of TMEB117 with TME204 and AM560-2 v8.1, two other high-quality genome assemblies of cassava.

Results

We achieved a highly accurate, almost complete haplotype-resolved genome of the African TMEB117 cassava cultivar. The two haplotype genomes attained a base-level accuracy of > QV 64, > N50 35Mbps, and a BUSCO completeness of 98.9%, making it the most complete and accurate haplotype-resolve cassava genome to date. Repeats mask more than 60% of the genome. We predicted over 45K gene models. Our comparative genome analysis of three African cultivars, TMEB117, TME204, and AM560-2 v8.1, revealed a large insertion in our proposed assembly of 10Mbps in chromosome 12 and large inversions in chromosomes 4, 12, 15, and 18. Within the data, we also detected reads representing an almost full genome of the fungus Alternaria alternata, although the plants were healthy and showed no symptoms.
The phased and annotated homologous chromosomes provide a comprehensive perspective of cassava’s heterozygous genome organization with improved accuracy and completeness at a haplotype-resolved level. These chromosome pairs are anticipated to be a valuable resource for cassava breeders and essential research for functional analysis to characterize important agronomical traits molecularly. Moreover, TME117 is highly susceptible to African Cassava Mosaic Virus (ACMV) infections compared to other cassava varieties. It is an ideal reference to study virus resistance mechanisms, including epigenetic variations such as differentially methylated sites and smallRNA expressions. A critical study will also investigate the large insertions and inversions in both haplotypes as a stride toward resolving unexplored regions of the cassava genome.

Info

-

filename	-

Figure

-

Availability	-

Corresponding Author

Name, Surname	Michael, Landi
Email	m.landi@cgiar.org

Submitted on	21.04.2023

Società Italiana di Bioinformatica
C.F. / P.IVA 97319460586
E-mail bits@bioinformatics.it

Sede legale Viale G. Mazzini, 114/B - 00195 Roma
Website bioinformatics.it