

New Genomic Tools for Sweetpotato Improvement (GT4SP) are Coming On-Line

The sweetpotato breeding and genomics community of practice can now boast of new genomics and advanced modern breeding tools. The GT4SP's current polyploid analytic toolbox includes a pipeline for hexaploid genotyping by sequencing, software for assigning marker genotypes¹, establishing chromosome-scale linkage phase among marker alleles, constructing haplotypes, generating hexaploidy linkage maps, quantitative trait locus (QTL) analyses, simulating polyploid populations and, of course, a reference genome that can be used in all sweetpotato genomic studies even performing genome-wide association studies.



Fig. 1 Breeders learn how to use a tablet with the fieldbookApp for electronic data capture (credit F. Njunge).

What is the problem?

Sweetpotato is grown mainly in small plots by women and is often characterized as a poor person's food crop. Orange-fleshed sweetpotato delivering nutritional benefits to pregnant women and under five children has been introduced in sub-Saharan Africa by CIP and its partners. One of the principal challenges faced by breeding sweetpotato, is the obligate out-crossing and polyploidy (6 genomes) nature of the crop making segregation of useful traits unpredictable by classical genetics. Currently, there is an

explosion of interest in sweetpotato and polyploid genomics. However, the critical issue of how to interpret the molecular data remains, starting with the assignment of marker dosage, a.k.a. "genotype calling". In addition, data collection from breeding trials has used paper-based excel forms followed by double data entry into a relational database. Although well established, this method is time-consuming and error-prone. New data collection tools and processing have also been one of GT4SP applied objectives in close collaboration with the Sweetpotato Action for Security and Health in Africa (SASHA) project.

What are we doing about it?

Investigating the genetics and genomics of sweetpotato is providing us new approaches to sweetpotato improvement that will reduce the breeding cycle and increase predictability. GT4SP has developed tools and made them available to the research community. These tools are intended for experimental populations (rather than natural populations), facilitating genomics-assisted sweetpotato improvement. We focus on three main areas: (1) polyploid genotyping; (2) genetic and physical mapping; and (3) genomics-assisted

1. **Polyploidy** is the heritable condition of possessing more than two complete sets of chromosomes. Sweetpotato is a **hexaploid**, containing six homologous sets of chromosomes. **Genotyping** is the process of determining differences in the genetic make-up (**genotype**) of an individual by examining the individual's DNA sequence using biological assays and comparing it to another individual's sequence or a reference sequence. Assigning marker genotypes is estimating the dosage of marker alleles in the heterozygous condition. Genes come in pairs, called alleles, and each pair is located in a specific position (or locus) on a chromosome. If the two alleles at a locus are identical to each other, they are homozygous; if they are different from one another, they are **heterozygous**. A **haplotype** is a group of alleles in an organism that are inherited together from a single parent.



Key Partners

North Carolina State University (NCSSU) (Lead Institution); International Potato Center (CIP); Biosciences for Eastern and Central Africa (BeCA); Michigan State University (MSU); Boyce Thompson Institute (BTI) at Cornell University; University of Queensland (UQ); National Crops Resources Research Institute (NaCRRI), Uganda; Council for Scientific and Industrial Research, Crops Research Institute (CSIR-CRI), Ghana.

Fig. 2 The Sweetpotato Genomics Resource Website (Michigan State University)



breeding by quantitative trait analysis, genome wide association studies and genomic selection. These tools can also help elucidate the mode of inheritance (disomic, polysomic or a mixture of both) or reveal whether double reduction and multivalent chromosomal pairing occur. Much work remains to keep pace with the rapid developments in genomic technologies in other crops. However, the GT4SP polyploid analytical box offers the promise of understanding the sweetpotato genome at a level which hitherto has been elusive. Electronic data capture (EDC) tablets have been developed to enter, review and analyze data in real-time and to implement online data validation checks to assure data quality more effectively at the point of entry. The new data management tools (highly interactive data analysis platform; HIDAP and SweetPotatoBase) utilize the fieldbookApp (<https://sweetpotatobase.org/>) and are now in use in SSA with continuous training through webinars and workshops to the breeding community (Fig. 1).

What have we achieved so far?

The GT4SP project has already developed new genetic and genomic resources. Two reference genomes for sweetpotato, *Ipomoea trifida* and *I. triloba* genomes which are close relatives of the ancestors of sweetpotato *I. batatas*, and their associated genome browsers (<http://sweetpotato.plantbiology.msu.edu/index.shtml>) (Fig. 2). These new resources have enabled us to anchor maps, develop more informative single nucleotide polymorphic (SNP) markers, identify candidate genes associated with important traits such as biotic resistance traits such as resistance to sweetpotato virus disease, and nutritional traits like beta-carotene, starch and sugar content. The *Ipomoea e-PCR tool* is the newest tool on the reference genome that enables one to perform an in-silico PCR simulation. The program works by searching

for primer alignments in a reference genome that produce a product within the parameters selected. Gene annotation is being layered on the genome sequence browser. Soon, the global sweetpotato community will begin to utilize this important resource and progress more quickly in breeding.

To facilitate genomics-assisted breeding, we have developed GBSpoly, a robust SNP platform for genotyping our genetic mapping and breeding populations. This resource is coupled to linkage mapping and QTL analysis methods specially designed for polyploid crops such as sweetpotato. To be able to better serve the breeding community of the SASHA project, the work flows in HIDAP to SweetPotatoBase have been made more efficient (<https://sweetpotatobase.org/>).

Our genomics research is generating new tools, procedures, and dimensions to resolve important difficulties to make efficient progress in breeding sweetpotato. Genome sequences, gene annotations, SNP markers, genetic maps, interactive data analysis platforms, statistics for prediction models (SuperMassa) are already available to the breeder community. Through the capacity development unit, we are training the next generation of breeders on their use. Already, we have applied them to: 1) assessment of genetic diversity and selection of heterotic groups for creating new variations that will help in developing superior lines with a broadened genetic base; 2) design of markers from functional genes annotated on the reference genome; and 3) genome-wide association studies for beta carotene, weevil, and continuous bulking of sweetpotato. Future research will improve our understanding of the genetic base of the key sweetpotato traits to generate more quickly new orange-fleshed sweetpotatoes more resilient to environmental stresses without losing their nutritional and organoleptic qualities.

CONTACT

Dr. Craig Yencho (NCSU), craig_yencho@ncsu.edu