The use of cassavabase: a tool for cassava genomics at NRCRI, Nigeria.

Alex Ogbonna1,2, Chiedozie Egesi1*, Adeyemi Olojede1, Ezenwanyi Uba1, Nneka Okereke1, Emmanuel Okogieben1, Lukas Mueller2
1National Root Crops Research Institute (NRCRI), Umudike, Nigeria; 2Boyce Thompson Institute for Plant Research Ithaca, New York, U. S. A.
Corresponding author email address: aco46@cornell.edu, Registrant ID # 4593

Introduction

The Next Generation Cassava Breeding aims to significantly accelerate genetic improvement of cassava breeding and unlock the full potential of cassava, a staple crop central to food security and livelihoods across Africa (http://www.medicinacassava.org/about.html).

Cassavabase (www.cassavabase.org) was created to centralise information tracking, genotype and phenotypic data, and Genomic Selection prediction analyses.

Data is being collected on evaluation sheets and field notebooks which might be altered or misplaced pointing to the fact that evaluated data are not protected.

In addition, it takes longer time for the collected data to be entered into the computer, cleaned and prepared for analysis, giving room for error during data entering if proper care is not taken by the curator.

NRCRI uses Fieldbook application and Android tablet for phenotype collection in the field for it routine germplasm development.

The Android tablet is convenient because data files can easily be transferred; it gives a reasonable time for data collection and has little glare while being used in the field during data recording.

Cassavabase ultimately aims to achieve and facilitate untimely sharing of information, which means overcoming the traditional reluctance of researchers to share data.

Objectives

➢ To develop a data management strategy for massive germplasm development that will aid data accessing, sharing and retrieval at NRCRI, Umudike, Nigeria.

➢ To carry out data analysis using genomic prediction models on the cassavabase server.


A total of 500 pre-selected genotypes from NRCRI germplasm were evaluated at western form of NRCRI, Umudike, Nigeria. The genotypes were properly curated and uploaded into Cassavabase. The trial was created on Cassavabase, transferred to android tablet and setup with Fieldbook app for phenotyping. The use of the newly developed cassava traits ontology was fully implemented for this trial. Cassavabase genotypes selection pipeline was used to carry out genomic selection for this trial (population).

Use Case 1: Trial Creation.

Use Case 2: Data Collection

Use Case 3: Data Uploading

Use Case 4: Data Analysis

Genetic selection is the use of statistical methodology to predict the merit (GEBV), figure 2 of a genotype using genotypic information from it self and it’s relatives and phenotypic information from it relatives.

Cassavabase implements genetic selection algorithms for analysis of field trials and estimation of GEBVs.

Use Case 5: Data Analysis Output

Use Case 6: Breeding Outcome

➢ NB110070, NB110036 and NB110005 were ranked as the top producing lines from the analysis.

Cassava Trait Ontology

Use Case 7: Data Analysis Output

Conclusions

➢ The use of Cassavabase at NRCRI Cassava Breeding Program, has given hope to data accessing, sharing, retrieval and managing which brings about data integration and utilization.

➢ It has a data management component for phenotypic data generated by evaluation and testing.

➢ The solidS and genome browser component has a pipeline for performing genomic selection (GS) and running the genomes at any scale with dozens of aligned annotation tracks respectively.

➢ Cassavabase has given our breeding scheme a new book, ending the era of having trial information on field books and on the shelf of scientists. However, Cassavabase will ensure an open data policy for accelerating research on Cassava.

Acknowledgments

We thank National Root Crops Research Institute, Umudike, Nigeria and Boyce Thompson Institute, Ithaca, New York for their support. Funding for this project was provided by the Next Generation Cassava Project.