ACWP2

Annual Report Year 2 (2020)

General Information			
Project name	African Cassava Whitefly Project Phase 2: Cassava-whitefly control for sustainable food security in Sub-Saharan Africa		
Sub-grantee organization name	The International Center for Tropical Agriculture (CIAT Colombia)		
Sub-grant ID#	B0540x9	BMGF program officer	Dr Christina Owen
Date sub-grant awarded		Project end date	October 31, 2022
Sub-grant amount (USD)	$800,000	Project duration	4 years
Report period - *from*	November 1, 2019	Report period - *to*	October 31, 2020
Sub-grant principal investigator	Luis Augusto Becerra Lopez-lavalle	Email address	L.A.Becerra@CGIAR.org
Mailing address	The International Center for Tropical Agriculture (CIAT Colombia). Km 17, Recta Cali–Palmira CP 763537 Apartado Aéreo 6713 Cali, Colombia	Phone number	+57 2 4450000

1. Progress and results

1. Progress narrative (2 pages maximum)

In paragraphs or bullet points, please provide the following information:

1. **General Progress:** Describe the general progress of the sub-grant, including where it is progressing as expected, where it is not, whether it is still on track to complete expected results, and if not, what proposed modifications are contemplated.

This document will report on the activities completed during the 6 months NCE of phase I and year 1 of phase II partially.

Aim 2: The rapid generation of cassava varieties possessing multiple whitefly and virus-disease resistances (to ensure sustainability over the medium term).

Activity 2.1 Pre-breeding crosses conducted in East Africa, between WFR Latin American and African cassava varieties. In partnership with national cassava-program breeders (Dr. Robert Kawuki in Uganda and Dr. Kiddo Mtunda in Tanzania, both of whom are involved in the NextGen project and Mr. Obed Mwenye in Malawi), we shall make all possible combinations of crosses between the WFR LA and African cassava varieties. The parental material will be genotyped molecularly to confirm their identities and barcoded **[CIAT-RHUL-UCR Milestone 2: November 2019]**.

PROGRESS (CIAT): DNA extracts from the African genotypes collection (5CP) held at RHUL were sent to CIAT for genotyping and check against collections maintained at NRI and cassava African breeders. Matching genotypes and identity confirmation is now completed and previous data corrected accordingly. **Activity completed.**

Activity 2.2 WFR improvement in African 5CP genotypes and landraces. The 5CP lines (Mkumba, NASE3 and NASE14 and their susceptible comparators, Sauti and Mkuranga) will provide the parental material for crosses directed towards improved WFR, associated with virus-disease resistances. The genetic/biochemical pedigree of the selected 5CP genotypes is poorly understood and they are likely to be hybrids or non “true-breeding”. By exploiting this heterogeneity to assist the accurate identification of WFR genes, we shall: (i) test the genetic, biochemical and phenotypic stability/inheritance of the 5CP material, (ii) deliver the potential enhancement of the WFR trait within the 5CP material and (iii) identify potential unwanted sources of susceptibility. Sequence polymorphisms and metabolites/molecular features will be identified to enable their use as quantitative trait markers in future breeding programs. This approach has been used previously for other traits, such as the identification and enhancement of pro-vitamin A content in cassava. It provides another key advantage, because it provides an opportunity to

exploit the pre-existing pyramid of QTLs necessary for the maintenance consumer or agronomic traits and other biotic resistances **[CIAT-RHUL-UCR Milestone 5: June 2019]**.

PROGRESS (CIAT): 5CP accessions (510 samples from Africa) were processed using a SNP-chip for the Nanofluidic Dynamic Arrays (SNPY-Array; Fluidigm®, USA) developed by CIAT that contains 96 single nucleotide polymorphism (SNP) to genotyping in cassava. The technique allowed simultaneously collecting both end-point and real-time data from a unique chip cell with 97% confidence. Genetic Duplicates Test: In our variety identification test, all samples that are genetic duplicates belong to the same group. In total, we found 40 genetic duplicates groups (GD-Groups) that represent 40 different genotypes. This set of duplicate samples contains 508 samples from Uganda, Tanzania and Malawi. Supporting data provided as supplementary "Plant Molecular Analysis 5CP". These 5CP genotyping results are complemented in the RHUL report. The results generated from this Activity 2.2 have been included in a scientific manuscript and submitted to PLoS One journal where is currently under review. **Activity completed.**

Activity 2.3. Advanced intercrossing activities to create pre-breeding cassava progenies homozygous for the WFR loci and possessing superior WFR to ECU72 (the original WFR donor). Cassava lines homozygous for WFR are needed to transfer superior WFR into regionally preferred, African-adapted cassava germplasm. The current project generated two advanced intercrosses (see Annex 1 SI2.3 for details of CM8996 and GM8586 and the intercrossing method). We shall continue to utilize this resource to delineate precisely and verify WFR gene regions. We shall then identify molecular markers within the cassava WFR genes because they shall be the most accurate, *i.e.* there is no opportunity for the markers to segregate from the causal genes. Advanced intercrosses will be made for three further generations at CIAT, using progeny identified to possess the best WFR **[CIAT-RHUL-UCR Milestone 7: ~~December 2019~~, revised and moved to December 2020]**. The advanced intercrossed cassava will then be used to: (i) identify, verify and refine QTLs for WFR and (ii) identify the potential epistasis that exists for WFR QTL (see Activity 2.4 below).

PROGRESS CIAT: We are developing twelve F2 "advanced intercross" populations for resistance to whitefly: The seeds of these offspring were planted in soil, multiplied in cuttings and are being phenotyped in the greenhouse.

F2 "advanced intercross" populations for resistance to whitefly:

# advanced intercross	Female	Male	Cross	# offspring
1	CM8996-581	CM8996-199	GM12200	193
2	CM8996-596	CM8996-199	GM12201	360
3	CM8996-199	CM8996-199	AM1588	241
4	CM8996-246	CM8996-199	GM12199	644
5	CM8996-758	CM8996-199	GM12202	626
6	CM8996-193	CM8996-199	GM12198	53
7	GM8586-198	GM8586-64	GM13464	291
8	GM8586-198	GM8586-103	GM13465	412
9	GM8586-198	GM8586-153	GM13466	93
10	GM8586-64	GM8586-64	AM1620	110
11	GM8586-49	GM8586-64	GM13463	144
12	GM8586-103	GM8586-103	AM1621	82

Activity completed.

Activity 2.4 WFR-phenotyping, metabolomic and transcriptomic analysis of segregating populations to elucidate QTL underlying whitefly resistance/tolerance. The previous activity (2.3) shall deliver "advanced intercrossed" cassava that by the F_3 should contain some progeny that breed true for WFR (the proportion should increase by the F_4) **[CIAT-RHUL-UCR Milestone 20: December 2021]**. The F_2 and F_3 segregating populations, therefore, will also be evaluated for WFR ("phenotyped" – see SI2.4 in Annex 1) and material sent to UCR and RHUL for transcriptome (eQTL) and metabolome analyses, respectively, to develop robust molecular and quantitative markers for WFR QTL. Four F_2 intercrosses will be made between WFR genotypes identified from the CM8996 and GM8586 populations and ~200 progeny from each intercross will be phenotyped for WFR (see Annex 1, SI2.4 Figure 1 for the workflow). Each plant will genotyped (see Annex SI2.4.2) using current molecular markers linked to putative *WFR* loci on Chr 2, 7 and 10 to verify the number and identity of the *WFR* loci present in each F_2 progeny and plants with all three WFR loci present will be studied. eQTL analyses will provide data about the associated presence or absence of WFR loci in Chr 5, 6, 11, 14 and 18 regions. The subset of WFR-marker positive plants and WFR-deplete (negative control) will be further characterized. Similar strategies for characterization of the plants in F_3 populations will also occur using the same strategy **[CIAT-RHUL-UCR Milestone 22: June 2022]**. True true-breeding genotypes will be identified from the F_4 populations, but due to the time need to make the crosses, this may occur after Yr. 4 of this proposal **[CIAT-RHUL-UCR Milestone 21: June 2022].**

PROGRESS (CIAT): Report will be delivered in 2021-2022.

Activity 2.4.1 Whitefly-infestation bioassays (WFR phenotyping). The current project developed an objective and automated method for WFR phenotyping of cassava (see Annex 1 SL2.4). This method will be used to assess plants from the intercrossed F_2 and F_3 progeny to determine their WFR or WFS. Both antibiosis (nymph death) or antixenosis (repellence) will be assessed.

Whitefly infestations will be performed at CIAT and paired-samples split for RNA and metabolite analyses. RNAs will be sent to UCR for RNA-seq (eQTLs), while tissue will be sent to RHUL for chemotyping. As we progress from the F_2 to the F_3 population, the proportion of superior WFR plants (relative to ECU72) is expected to increase.

PROGRESS CIAT: We developed a Phenotyping methodology for whitefly resistance in cassava, called "Nymphstar", which consists of a greenhouse Phenotyping method and an imageJ plugin for automated counting nymphs. For 4 years we phenotyping the progeny (F1) of ECU72 (WF-R) and COL2246 (WF-S) crosses (CM8996) and ECU72 (WF-R) and TMS60444 (WF-S) crosses (GM8586). We are in the process of Phenotyping of the twelve F2 progenies, obtained in activity 2.3. We have completed the phenotyping of 5 of these F2s (AM1588, GM12200, GM12201, GM12202, GM12199), of which we selected AM1588, as the best F2 whitefly resistant. The 15 top resistant and 15 top susceptible of this F2 were selected to be assessed. Whitefly infestations will be performed at CIAT and paired-samples split for RNA and metabolite analyzes. RNAs will be sent to UCR for RNA-seq (eQTLs), while tissue will be sent to RHUL for chemotyping. Supporting data provided as supplementary (word file).

Activity 50%.

Activity 2.4.2 Genotyping progeny possessing superior WFR. Initial screening of an advanced intercrossed F_2 population will be performed using a set of molecular markers that span the three putative WFR loci on Chr 2, 7, and 10 (CIAT). Phenotypes (Obj. 2.4.1) will be correlated with genotypes at the three WFR loci (see Annex 1 SI2.4.2). Analysis of the progeny harboring all three WFR regions, a subset of these regions and completely lacking these regions will determine the ability of these markers to identify the superior WFR seen in the F_2 progeny of the CM8996 and GM8586 crosses. Approximately 200 of the best progenies of the F_2 populations, as well as their WFR parents, will be chemotyped (RHUL, see activity 2.4.3) and genotyped by RNA-seq (UCR) and these activities will validate the putative WFR markers and QTL regions.

We shall use RAD sequencing to provide excellent coverage for subsets of the cassava genome of interest, to identify the best (most tightly linked) WFR molecular markers. RNA-seq analysis of transcriptomes (exome sequencing) of the WFR parents and progeny from the advanced intercrosses will also be conducted, as an important complement to RAD sequencing. The SNPs (eQTLs) generated from RNA-seq will span the entire genome. CIAT will send RNAs to UCR for library construction and analyses. We will use the eQTLs to fine map the current WFR loci on Chr 2, 7, and 10 and the ancillary QTLs on Chr 5, 6, 11, 14 and 18 will be able to be assessed and correlated with WFR (see Annex 1) **[CIAT-RHUL-UCR Milestone 10: June 2020].**

PROGRESS CIAT: The selected F2 AM1588:

-DNA was extracted from 196 individuals of this progeny and RAD-sequencing approach was used for sequencing made by BGI.

-SNPs analysis, mapping and QTL analysis in process.

Activity 50%.

Outputs: 3 manuscripts (1-3) generated from phase I activities have been submitted and accepted (see below) and at least 1 related to the phase II activities described above are in preparation. Dissemination to the public has been carried out through presentations at college open days.

1. A metabolomics characterisation of natural variation in the resistance of cassava to whitefly. Perez-Fons et al, 2019. **BMC Plant Biology, 19: 518**. https://doi.org/10.1186/s12870-019-2107-1
2. Genome-wide analyses of cassava Pathogenesis-related (PR) gene families reveal core transcriptome responses to whitefly infestation, salicylic acid and jasmonic acid. Irigoyen et al, 2020. **BMC Genomics, 21:93 2020. https://doi.org/10.1186/s12864-019-6443-1**
3. The metabotyping of an East African cassava diversity panel: a core collection for developing biotic stress tolerance in cassava. **Perez-Fons et al, 2020. PLoS ONE**, submitted 25th July 2020.
4. NYMPHSTAR: an accurate high-throughput quantitative method for whitefly (Aleurotrachelus socialis Bondar) resistance phenotyping in cassava. Adriana Bohórquez-Chaux , Luisa Fernanda Leiva-Sandoval, María Isabel Gómez-Jiménez, Fausto Rodríguez and Luis Augusto Becerra López-Lavalle. **In preparation.**

2. **Key milestone deviation:** If specific key milestones were not met according to the proposed plan, briefly discuss the reasons they deviated and the proposed corrective actions. In your discussion, reference the specific key milestones from the Results Framework.

3. **Course correction:** Whether or not you are on track to meet key milestones, with the benefit of current experience, are there any modifications you would propose to the activities, outcomes, outputs, or key milestones of this work.

Evaluation of whitefly resistance phenotype has proved to be challenging and has affected the process of identifying adequate WFR parents in the F1 family to produce F2 and F3 progenies homozygous for the WFR loci. Overall, no deviations from the original proposed plan are expected and delivery of key milestones described in activities 2.3 and 2.4 will be delayed between 6-12 months. The issue was previously reported and timelines concerning Activity 2.3 (milestone 7) have been moved accordingly.

4. **Plans for next reporting period:** To the extent that you have proposed modifications to any of the results identified in the Results Framework, identify why modification of these results will still lead to achievement of the sub-grant, and how your organization will be able to successfully achieve the results as modified.

All activities related to: Activity 2.4 WFR-phenotyping, metabolomic and transcriptomic analysis of segregating populations to elucidate QTL underlying whitefly resistance / tolerance. The previous activity (2.3) delivered "advanced intercrossed" cassava F2. The best F2 whitefly resistance (AM1588) was selected, the most resistant offspring are being used for crosses and advance to F3. Publication in peer-journals the results the mapping and analysis of QTLs of the F1s CM8996 and GM8586.

5. **Risks:** Are you aware of any significant risks or concerns that have not previously been identified, and that may affect your organization's ability to achieve the agreed-on results? If so, indicate how your organization is addressing those risks.

Low

6. **Sustainability:** If your organization intends for your goals to be sustained after the sub-grant period has ended, what actions have your organization and associated partners taken and what actions will you be taking to facilitate sustainability, and how will the project be continued?

Not expected

7. **Scalability:** If your organization intends for this sub-grant to increase in scale after the grant period has ended, what actions have your organization and partners taken and what actions will you be taking to facilitate that increase in scale?

Not expected

8. **Lessons Learned:** What lessons have you learned during the past year that will help you to achieve your intended results moving forward?

Communication and cooperation between theme partners has been essential for milestones completion.

Interaction with other projects linked to cassava cultivation in Africa, like NextGen, has proved to be an added value for understanding the multiple players in the cassava research community and the different challenges that farmers face for enhancing crop's yield and productivity.

1. **Realized impacts of COVID-19:** Please outline any impacts of COVID-19 on your team's working arrangements, technical progress, and financial standing.

The assistance of field workers and technicians to Campus and complying with COVID19 safety guidelines, i.e., keeping social distance has had a major impact in routine functioning of the research work. Despite having ample facilities and a lot of space, the number of people who attend the campus is limited; this has delayed the work a bit. We have made a great effort to maintain the Phenotyping activities, the planting of the F2 in the field for crosses and the laboratory activities.

2. Results framework

For each milestone relevant to the reporting period, please provide an estimate of percentage completion and completion date where appropriate.

Output Number	Output Description	Target Completion Date	Percentage completion	Actual completion date

CIAT-RHUL-UCR Output 1	Presentation of ACWP Phase II data and national/international meetings.	31 January 2019	100%	Speaker of presentation "QTLs Mapping in Cassava for Whitefly Resistance" in Plant and Animal Genome conference 2019. San Diego CA, 16th January, 2019.
CIAT-RHUL-UCR Output 2	Well-characterized parental germplasm for multi-locational/environmental field testing and use in genetic crossing programs.	1 November 2019	100%	June 2020
CIAT-RHUL-UCR Output 3	WF bioassays and quantification of WFR/WFS training of scientists in African National Programs (on going each year).	31 October 2022	Not applicable	-
CIAT-RHUL-UCR Output 4	Six-month report	31 May 2019	100%	Teams reported October 2019
CIAT-RHUL-UCR Output 5	List of F_2 parents for "advanced intercross" F_3 population.	1 June 2019	50%	June 2020
CIAT-RHUL-UCR Output 6	Six-month report	31 October 2019	100%	Teams reported December 2019
CIAT-RHUL-UCR Output 7	Production of multiple F_2 "advanced intercross" populations	Revised date of completion: 01 December 2020	100%	Delivery of samples to RHUL and UCR expected by December 2020.
CIAT-RHUL-UCR Output 8	Presentation of ACWP-Phase II data and national/international meetings.	31 January 2020	100%	Poster presentation "QTLs Mapping in Cassava for Whitefly Resistance" and "High throughput phenotyping of resistance against whiteflies (A. socialis Bondar) using image analysis" in X Encuentro Latinoamericano y del Caribe de Biotecnologia Agropecuaria, REDBIO, Montevideo, Uruguay, 10-15 Noviembre, 2019
CIAT-RHUL-UCR Output 9	Manuscript #1 submitted	1 June 2020	100%	Submitted to PLoS One, July 2020
CIAT-RHUL-UCR Output 10	Uploading of RAD-seq, RNA-seq and metabolite data to public data bases associated with Manuscript #1.	1 June 2020	100%	Metabolite data publicly available through open access publications (The Plant Journal, BMC Plant Biology, PLoS One)
CIAT-RHUL-UCR Output 11	Six-month report	31 May 2020	100%	Covid-impact mitigation plan submitted
CIAT-RHUL-UCR Output 12	Six-month report	31 October 2020	100%	Submitted
CIAT-RHUL-UCR Output 13	Manuscript #2 submitted	1 December 2020		
CIAT-RHUL-UCR Output 14	Uploading of RAD-seq, RNA-seq and metabolite data to public data bases associated with Manuscript #2.	1 December 2020		

CIAT-RHUL-UCR Output 15	Presentation of ACWP Phase II data and national/international meetings.	31 January 2021		
CIAT-RHUL-UCR Output 16	Six-month report	31 May 2021		
CIAT-RHUL-UCR Output 17	Manuscript #3 submitted	1 December 2021		
CIAT-RHUL-UCR Output 18	Uploading of RAD-seq, RNA-seq and metabolite data to public data bases associated with Manuscript #3.	1 December 2021		
CIAT-RHUL-UCR Output 19	Six-month report	31 October 2021		
CIAT-RHUL-UCR Output 20	Production of an F_3 "advanced intercross" population.	1 December 2021		
CIAT-RHUL-UCR Output 21	Well-characterized F_3 generation germplasm for use as parents to produce F_4 generation True-breeding germplasm that will be used transfer the multigenic WFR to regional WFS and WFR African varieties.	1 June 2022		
CIAT-RHUL-UCR Output 22	List of F_3 parents that could be used for future "advanced intercross" F_4 population.	1 June 2022		
CIAT-RHUL-UCR Output 23	Presentation of ACWP-Phase II data and national/international meetings.	31 January 2022		
CIAT-RHUL-UCR Output 24	Manuscript #4 submitted	1 October 2022		
CIAT-RHUL-UCR Output 25	Uploading of RAD-seq, RNA-seq and metabolite data to public data bases associated with Manuscript #4.	1 October 2022		

2. Budget progress report

Budget progress narrative (2-page maximum)

In paragraphs or bullet points, please provide the following information:

1. **General budget progress:** Describe the general progress of meeting budget expectations; including where the sub-grant is progressing as expected, where it is not, whether the sub-grant is still on track to be completed within the proposed budget, and if not, what proposed modifications are contemplated.

RHUL is spending at a slower rate because of completing phase I – especially the issues surrounding the genotyping and metabotyping of the 5CP population.

2. **Budget variances:** For variances that exceed 10 percent in either direction in the Total Cost category (i.e. Total Personnel, Total Supplies, Total Equipment), please describe these clearly.

At this stage there are no variance but there could be changes when the large populations arrive as extra personal resources will be required to complete the tasks.

3. **Budget plans for next reporting period:** Explain any significant reforecasting, any impact that the reforecasting will have on the total budget, and how your organization will be able to successfully perform within the re-forecasted budget.

It is envisaged that we will still be spending at a slower rate because of the restrictions imposed by the COVID-19 situation.

4. **Budget or Financial Risks:** Are you aware of any significant risks or concerns that have not previously been identified, and that may affect your organization's ability to perform this sub-grant within the designated budget? If so, how is your organization addressing those risks?

No budget risks are envisaged but we will require an extension to counter balance the issues that have arisen from COVID-19 restrictions.

Plant Molecular Characterization – Cassava Samples from Africa

Luis Augusto Becerra - Tatiana Ovalle – María Alejandra Bedoya

Plant material

CIAT's cassava molecular laboratory received in total 543 cassava leave samples from three African countries. In summary, Tanzania (144 samples), Uganda (380 samples) and Malawi (19 samples) with the goal of carrying out a molecular identification of these cassava varieties (Figure 1); for more details see attached file "Africa Sample List".

CIAT added the suffix's _TAN (samples from Tanzania), _UGA (samples from Uganda) and _MAL (samples from Malawi); in order to identify them during the variety identification analysis. The samples from Uganda were labeled according to the name that came in each of the bags containing the leaf tissue and that corresponds to the QR code. The suffix "MARI" of the Tanzanian samples was removed from the labels and only the genotype name and the number assigned in the original list was retained. Samples that presented fungi were not taken into account for the DNA extraction process or for molecular analysis, these samples were discarded. Similarly, three samples were discarded in this study because the label in the bag no correspond to QR code.

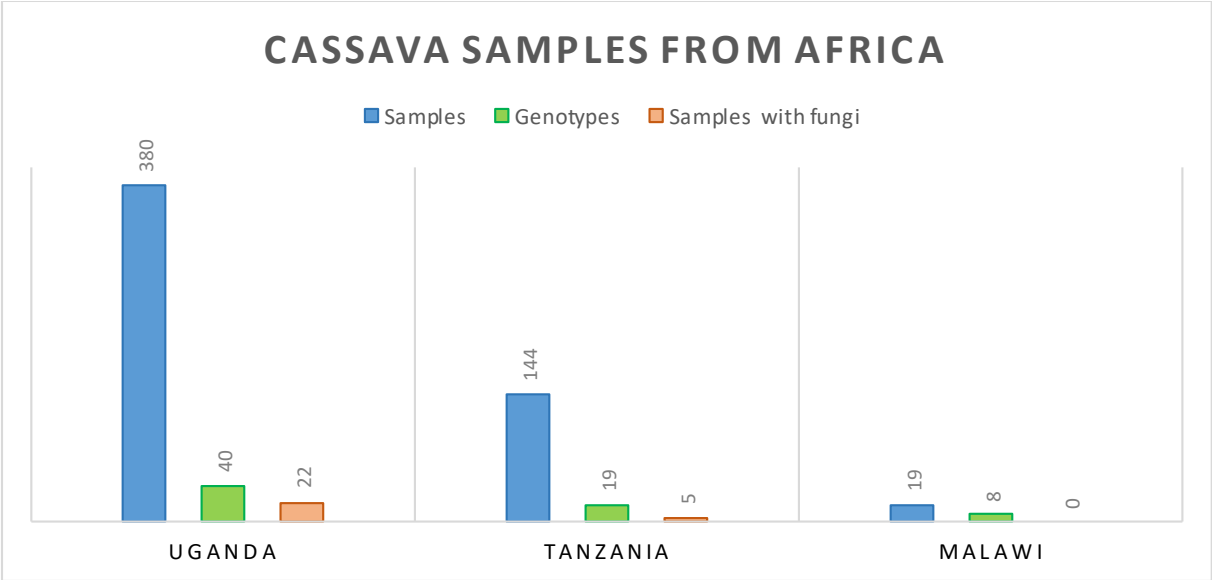

Figure. 1 Cassava leaf samples from Africa that arrive to CIAT's lab.

Methodology

DNA material

DNA was extracted using CTAB-based DNA extraction protocol described by Doyle (1991[1]) with minor modifications made by CIAT.

DNA Quality

All samples of DNA was evaluated in agarose gel 1% with propose to verify the quality (See Figure 2). DNA quantification was carried out by absorbance reading using the Sinergy quantifier.

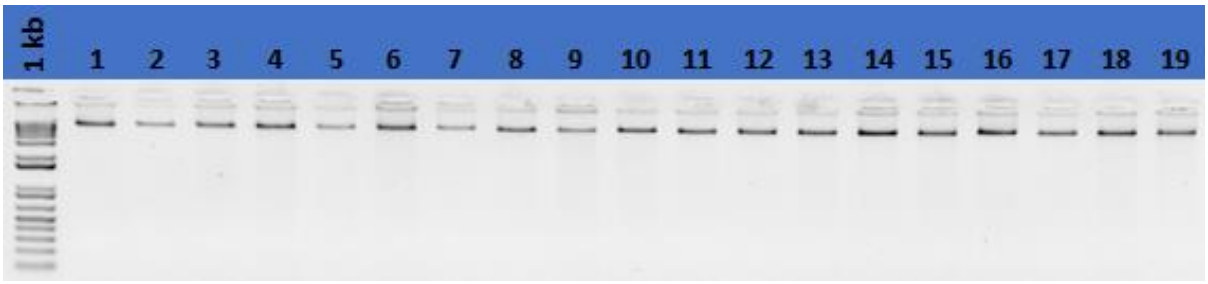

Figure. 2 DNA Quality extracted from cassava leaves.

SNPs Genotyping

The samples were processed using a SNP-chip for the Nanofluidic Dynamic Arrays (SNPY-Array; Fluidigm®, USA) developed by CIAT that contains 96 single nucleotide polymorphism (SNP) to genotyping in cassava. The technique allowed simultaneously collecting both end-point and real-time data from a unique chip cell with 97% confidence. DNA was then genotyped using a SNP-chip for the Nanofluidic Dynamic Arrays (SNPY-Array; Fluidigm®, USA)

Data Analysis

Data generated were integrated into the Galaxy Platform (http://galaxyproject.org/) and duplicate test was done in Eclipse SDK v. 4.2.2 using the NGSEP platform. Samples with a percentage of differences lower than 3% were considered genetic duplicates*. To carry out with varietal identification, we used a reference set of cassava from Asian, Africa, Latin American landraces and improve materials; in total 2500 samples were included in this analysis.

Results

510 samples from Africa were genotyping using 96 SNPs (Fig.3). Three samples (TZ130_7.3_TAN, ALADU ALADU_9.7_TAN, EYOPE_0285_UGA) were removed from the analysis, because they presented lost data greater than 30%. In total 510 samples were run in the analysis.

---

[1] Doyle, J. 1991. DNA Protocols for Plants. In Molecular Techniques in Taxonomy. Edited by Godfrey M. Hewitt, Andrew W. B. Johnston, and J. Peter W. Young. Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 283-293.

*In few cases, we showed samples in the same genetic duplicate group with differences up to 4.8%

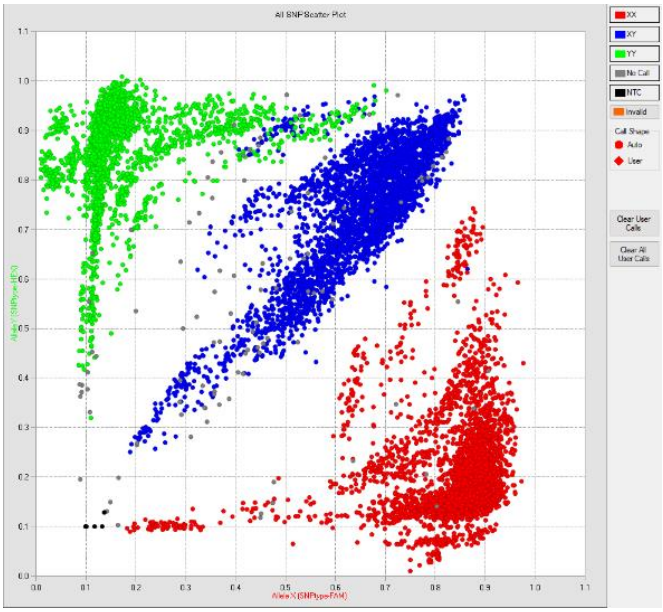

Fig.3 Cassava genotyping using 96 SNPS.

Genetic Duplicates Test

In our variety identification test, all samples that are genetic duplicates belong to the same group. In total, we found 40 genetic duplicates groups (GD-Groups) that represent 40 different genotypes. This set of duplicate samples contains 508 samples from Uganda, Tanzania and Malawi. Additionally, we detected two genotypes SAUTI_MALP-1 and RASTA19.3_TAN that are unique in this study and do not belong to any of the genetic duplicates groups defined in this analysis; these two samples are should be treated as two unique and unknown cassava genotypes (for more details see attached file "DuplicateTestAfrica_CIAT").

Uganda, Tanzania and Malawi have the same genotype **PER608** that correspond to sample **PER608** kept at CIAT. However, in Malawi the nomenclature is PERU608 (Table 1). **MKUMBA** is the same genotype in the three countries, but in Malawi is called TZ 03_MKUMBA. Additionally, Uganda has a material duplicate to **MKUMBA** called **PWANI.**

NASE18 from Uganda is genetic duplicate with **NASE14** from Tanzania, however, one sample labeled NASE18_0680_UGA belong to other genetic duplicate group. Additionally, **NASE18** has genetic duplicate with **72TME14** from Uganda and **TZ 130** from Malawi

Table 1. Group of genetic duplicate of the samples corresponding to PER608.

GD-Group	Sample	SNPs	%TotalDiff	Duplicates in the reference set
3	PER608_0001_UGA	91	0	PER608_CIAT
3	PER608_0002_UGA	91	0	NjuleRed_NRI
3	PER608_0003_UGA	91	0	
3	PER608_0004_UGA	90	0	
3	PER608_0005_UGA	90	0	
3	PER608_0006_UGA	91	0	
3	PER608_0007_UGA	88	0	
3	PER608_0008_UGA	91	0	
3	PER608_0009_UGA	91	0	
3	PER608_0010_UGA	91	0	
3	PER608_2.1-TAN	89	1.1	
3	PER608_2.2-TAN	89	1.1	
3	PER608_2.3-TAN	89	1.1	
3	PER608_2.4-TAN	89	1.1	
3	PER608_2.5-TAN	89	1.1	
3	PER608_2.6-TAN	89	1.1	
3	PER608_2.7-TAN	89	1.1	
3	PERU608_MAL P-1	91	0	
3	PERU608_MAL P-2	91	0	

Both Uganda and Tanzania showed the same genotype and the correct nomenclature to **ALADU ALADU** (Table 2), **COL2246**, **MAGANA**, **NJULERED**, and **TONGOLO**. In the case of **FUMBACHAI** from Uganda is the same genotype that **OFUMBA CHAI** from Tanzania. **NAM130** and **NARO _CASS1** from Uganda are genetic duplicates with **TZ 130** from Tanzania. **KBH2006/26** from Uganda is genetic duplicate with **MKURANGA** from Tanzania.

Table 2. Group of genetic duplicates of the samples corresponding to ALADU ALADU.

GD-Group	Sample	SNPs	%TotalDiff	Duplicates in the reference set
10	ALADU_ALADU_0701_UGA	94	0	Alady
10	ALADU_ALADU_0702_UGA	93	0	Alady_NRI
10	ALADU_ALADU_0711_UGA	94	0	
10	ALADU_ALADU_0712_UGA	94	0	
10	ALADU_ALADU_0713_UGA	94	0	
10	ALADU_ALADU_0714_UGA	94	0	
10	ALADU_ALADU_0715_UGA	94	0	
10	ALADU_ALADU_0716_UGA	94	0	
10	ALADU_ALADU_0717_UGA	94	0	
10	ALADU_ALADU_0718_UGA	94	0	
10	ALADU ALADU_9.1-TAN	92	0	
10	ALADU ALADU_9.2-TAN	92	0	
10	ALADU ALADU_9.3-TAN	92	0	
10	ALADU ALADU_9.4-TAN	92	0	
10	ALADU ALADU_9.5-TAN	92	0	
10	ALADU ALADU_9.6-TAN	92	0	
10	ALADU ALADU_9.8-TAN	92	0	

PER368 is genetic duplicate with some samples of **PER415** (GD-Group 6). This is explained because in the CIAT collection these two samples are genetic duplicates. However, others samples of PER415 are duplicates with **PER 317** from Uganda and Tanzania (GD-Group 8) see Table 3.

Table 3. Groups of genetic duplicates of the samples corresponding to PER317, PER368, PER415.

GD-Group	Sample	SNPs	%TotalDiff	Duplicates in the reference set
6	PER368_0101_UGA	94	0	NG11_NRI
6	PER368_0103_UGA	94	0	PER368_CIAT
6	PER368_0112_UGA	94	0	PER415_CIAT
6	PER368_0113_UGA	93	0	
6	PER368_0114_UGA	94	0	
6	PER368_0115_UGA	94	0	
6	PER368_0116_UGA	94	0	
6	PER368_0117_UGA	94	0	
6	PER368_0118_UGA	94	0	
6	PER368_0119_UGA	94	0	
6	PER368_1.1-TAN	92	0	
6	PER368_1.2-TAN	92	0	
6	PER368_1.3-TAN	92	0	
6	PER368_1.4-TAN	92	0	
6	PER368_1.6-TAN	92	0	
6	PER415_0062_UGA	93	0	
6	PER415_0064_UGA	94	0	
6	PER415_0065_UGA	94	0	
6	PER415_0066_UGA	93	0	
6	PER415_0067_UGA	94	0	
6	PER415_0068_UGA	93	0	
6	PER415_0069_UGA	94	0	
6	PER415_0070_UGA	94	0	
6	PER415_3.1-TAN	92	0	
6	PER415_3.2-TAN	92	0	
6	PER415_3.3-TAN	92	0	
6	PER415_3.4-TAN	92	0	
8	PER317_0082_UGA	91	0	Magana_NRI
8	PER317_4.1-TAN	89	1.1	PER608_NRI
8	PER317_4.2-TAN	89	1.1	PER273_CIAT
8	PER317_4.3-TAN	89	1.1	
8	PER317_4.4-TAN	88	0.0	
8	PER317_4.5-TAN	89	1.1	
8	PER317_4.6-TAN	89	1.1	
8	PER317_4.7-TAN	89	1.1	
8	PER317_4.8-TAN	89	1.1	
8	PER415_0072_UGA	91	0.0	
8	PER415_0074_UGA	91	0.0	
8	PER415_3.5-TAN	89	1.1	
8	PER415_3.6-TAN	89	1.1	
8	PER415_3.7-TAN	89	1.1	
8	PER415_3.8-TAN	89	1.1	

Both Uganda and Malawi showed the same genotype and the correct nomenclature to **MBUNDUMALI**, and ECU72 (Table 4).

Table 4. Group of genetic duplicates of the samples corresponding to ECU72.

GD-Group	Sample	SNPs	%TotalDiff	Duplicates in the reference set
1	ECU72_0021_UGA	94	0	ECU72_NRI_RH
1	ECU72_0022_UGA	94	0	Ecu72_NRI
1	ECU72_0023_UGA	92	0	ECU72_CIAT
1	ECU72_0024_UGA	93	0	
1	ECU72_0025_UGA	94	0	
1	ECU72_0026_UGA	94	0	
1	ECU72_0027_UGA	94	0	
1	ECU72_0028_UGA	94	0	
1	ECU72_0029_UGA	94	0	
1	ECU72_MAL P-1	94	0	
1	ECU72_MAL P-2	94	0	

Uganda showed the same genetic duplicate group for each genotype and their respective samples with the correct nomenclature to NGA11 (Table 5), ALBERT, EYOPE, F10-30-R2, F19, KBH2002/066, KBH200214, KIROBA, MKOMBOZI, NASE3, OEKHUMULELELA, TMEB1786, TZ_130 and YISAZO.

Table 5. Group of genetic duplicates of the samples corresponding to NGA11.

GD-Group	Sample	SNPs	%TotalDiff	Duplicates in the reference set
2	NGA11_0143_UGA	91	0	TMS60444_RH
2	NGA11_0144_UGA	91	0	ECU72_RH
2	NGA11_0145_UGA	91	0	PER273_NRI
2	NGA11_0146_UGA	91	0	NGA11_CIAT
2	NGA11_0147_UGA	91	0	
2	NGA11_0148_UGA	91	0	
2	NGA11_0149_UGA	91	0	
2	NGA11_0150_UGA	91	0	
2	NGA11_0153_UGA	91	0	
2	NGA11_0154_UGA	84	2.4	

However, Uganda present genetic duplicates. **LM1/2008** is duplicate with **TARIJIKA**, **CHO05/203** is duplicate with **SAGONJA-R** and one sample called **SAUTI_0183** is duplicate with all samples called **SHIBE.** Similarly, one sample called **PER385_0138** is duplicate with all samples labeled **SAUTI**.

Tanzania showed the same genetic duplicate group for each genotype and their respective samples with the correct nomenclature to NDOPE (Table 6), RASTA, NYAMKAGILE and LWAKITANGAZA.

Table 6. Group of genetic duplicates of the samples corresponding to NDOPE.

GD-Group	Sample	SNPs	%TotalDiff	Duplicates in the reference set
30	NDOPE 16.1-TAN	96	0	
30	NDOPE 16.2-TAN	96	0	
30	NDOPE 16.3-TAN	96	0	
30	NDOPE 16.4-TAN	96	0	
30	NDOPE 16.5-TAN	94	0	
30	NDOPE 16.6-TAN	96	0	
30	NDOPE 16.7-TAN	96	0	
30	NDOPE 16.8-TAN	96	0	

Malawi showed the same genetic duplicate group for each genotype and their respective samples with the correct nomenclature to KACHAMBA (Table 7) and SAGONJA (this genotype no correspond to SAGONJA from Uganda).

Table 7. Group of genetic duplicates of the samples corresponding to KACHAMBA.

GD-Group	Sample	SNPs	%TotalDiff	Duplicates in the reference set
9	KACHAMBA_MAL P-1	93	0	
9	KACHAMBA_MAL P-2	93	0	
9	KACHAMBA_MAL P-3	93	0	

Conclusion

Generally, in this genetic comparison among samples submitted by the 3 countries showed that are misidentification based on the labels provided. Please refer to the summary table for specific recommendations.

SUMMARY TABLE: Genotyping of breeding lines and Landraces found in Uganda, Tanzania and Malawi

		UGANDA		TANZANIA	MALAWI		CONCLUSION
G01	ECU72	ECU72_(0021-0029)			ECU72_(P1 & P2)		Shared by Uganda and Malawi as ECU72 (not submitted by Tanzania) - Correct classification
G03	PER608	PER608_(0001-0010)		PER608_(2.1-2.7)	PER608_(P1 & P2)		Shared by Uganda, Tanzania and Malawi as PER608 - Correct classification
G06	PER368/PER415	PER368_101, 103, (0112-0119)	PER415_ 0062 & (0064 - 0070)	PER368_(1.1-1.6)	PER415_(3.1-3.4)		Shared by Uganda, Tanzania and Malawi as PER368 or PER415 (PER368 and PER415 are duplicates in the CIAT collection)
G13	COL2246	COL2246 (0041-0049)		COL2246 (5.1-5.6)			Shared by Uganda and Tazania as COL2246 - Correct classification
G02	NGA11/TMS60444	NGA11_(0143-0150), 153 & 154					Shared by Uganda as NGA11 - Correct clasification
G08	PER273/BOL3	PER317_0082	PER415_ 0072 & 0074	PER317_(4.1-4.8)			Shared by Uganda and Tanzania as PER317 or PER415- Incorrect classification (these correspond to PER273 or BOL3 which are duplicates in the CIAT collection)
G04	MEX54/MEX96	PER335_(0122-0129), 134 & 135					Shared only submitted by Uganda as PER335- Incorrect classification (Correct name is MEX54 or MEX96)
G29	UG1	NASE 3_(0801-0805),(0807-0810)					Shared by Uganda as NASE3 - Correct clasification
G39	UG3	TZ_130_0586,0588,0589					Shared by Uganda as TZ_130 - Correct clasification
G05	UG4/UG5/UG6	NASE18_(0661-0669) & 0671	72TME14 (0561, 0567 & 0568)	NASE14_(6.1-6.8)	TZ130_(P1, P2 & P3)		Shared by Uganda as NASE18 (UG4), 72TME14 (UG5), Tanzania as NASE14 (UG6) and Malawi as TZ130 (UG3). UG4, UG5 & UG6 are duplicates between Uganda and Tanzania; hence a common name for NASE18, 72TME14 and NASE14 needs to be agreed. Incorrect Clasification in Malawi
G22	KE1/KE3	LM1/2008/363_(0302-0309), 0313, 0314	TAJIRIKA_(0383-0390), 0394, 0395				Shared by Uganda - LM1/2008/363 (K1) and TAJIRIKA (K3) are duplicats; name consolidation is required - Correct clasification
G16	KE2	F19_(0343-0349), 0353, 0354					Shared by Uganda - Correct clasification
G35	KE4	SAUTI_0183	SHIBE_(0402-0408)				Shared by Uganda - Cleaely SHIBE is the dominant group (K4) with one genotype labeled as SAUTI_0183 belongs to the SHIBE materials - Correct clasification;
G15	KE5	F10-30-R2_(0323-0329) & (0333-0335)					Shared by Uganda as F10-30-R2 - Correct clasification
G37	KE6	TMEB1786_0425, (0432-0439)					Shared by Uganda as TMEB1786- Correct clasification
G26	KE7	MKOMBOZI_(0363-0370), 0373, 0374					Shared by Uganda as MKOMBOZI- Correct clasification
G18	TZ1	KBH2002/066_(0483-0485),0488, (0493-0495)					Shared by Uganda as KBH2002/066- Correct clasification
G27	TZ2/TZ3	MKUMBA_(0761-0770)	PWANI_(0442-0449), 0452	MKUMBA_(13.1-13.7)	TZ 03_MKUMBA (P1, P2 & P3)		Shared by Uganda, Tanzania and Malawi as **MKUMBA**; but in Uganada is a duplicate with PWANI so this two needs to be consolidated
G20	TZ5	KBH2006/26_(0502-0508), 0513, 0514		MKURANGA_15.1, 15.2, (15.4-15.8)			Shared by Uganda as KBH2006/26 (TZ5) and Tanzania as MKURANGA (Not a CP5); MKURANGA is TZ5 (KBH2006/26) should be relabeled accordingly
G11	TZ6	ALBERT_(0463-0473)					Shared by Uganda as ALBERT- Correct clasification
G21	TZ7	Kiroba_(0542-0549), 0553					Shared by Uganda as KIROBA- Correct clasification
G40	MAL1	YISAZO_(0243-0249), 0253, 0254					Shared by Uganda as YISAZO- Correct clasification
G25	MAL2	MBUNDUMALI_(0163-0170), 0173, 0174			MBUNDUMALI_MAL P-1	MBUNDUMALI_MW06_MAL P-2	Shared by Uganda, Tanzania and Malawi as MBUNDUMALI- Correct clasification; note that in Malwi MBUNDUMALI_MAL P-1 and MW06_MAL P-2 are duplicated genotypes
G36	MAL3	SAUTI_(0184-0189),0193, 0194	PER335_0138				Shared by Uganda as SAUTI- Correct clasification; however, one sample labelled PER335_0138 belongs to this group; hence needs to be relabelled
G12	MAL4	CHO05/203_0221, (0223-0229), 0233 & 0234	SAGONJA-R_0203, 0205, (0207-0209) & (0213-0216)				Shared by Uganda as CHO05/203 (MAL4) and SAGONJA-R; both are duplicated - Correct clasification. Both Sample needs to be consolidated
G07	MAL5				SAGONJA_(P1, P2 & P3)		Shared by Malawi as SAGONJA- Correct clasification
G33	MOZ1	OEKHUMULELELA_(0263-0270), 0274, 0275					Shared by Uganda as OEKHUMULELELA- Correct clasification
G14	MOZ2	EYOPE_0283,0283,0286,0287,0289, (0295-0297)					Shared by Uganda as EYOPE- Correct clasification
G09	NOT CP5				KACHAMBA_(P1, P2 & P3)		
G10	NOT CP5	ALADU_ALADU_0701, 0702, (0711-0718)		ALADU ALADU_(9.1- 9.8)			
G17	NOT CP5	FUMBACHAI_(0741-0748), 0751, 0752		OFUMBA CHAI_(8.1-8.8)			
G19	NOT CP5	KBH200214_0521,0524,0525,0528,0529,0533					
G23	NOT CP5			LWAKITANGAZA_(18.1-18.8)			
G24	NOT CP5	MAGANA_0681,0682,0691,0692, (0696-0699)	NASE18_0680	MAGANA_11.1,11.2,(11.4-11.8)			NASE18_0680 belong to this group and not to UG4/UG5/UG6
G28	NOT CP5	NAM_130_(0651-0659)	NARO-CASS1_(0781-0790)	TZ130_(14.1-14.8) & TZ130_7.2, (7.5-7.8)			
G30	NOT CP5			NDOPE _(16.1-16.8)			
G31	NOT CP5	NJULE_RED_(0601-0608), 0611, 0612		NJULERED_(12.1-12.8)			
G32	NOT CP5			NYAMKAGILE_(17.1-17.8)			
G34	NOT CP5			RASTA_(19.1-19.8)			
G38	NOT CP5	TONGOLO_0722, 0723, (0731-0738)		TONGOLO_(10.1-10.6)			

Supplementary for Activities 2.3 and 2.4.1

B0540x9 CIAT: ACWP2 Year 2 Annual Report

Luis Augusto Becerra Lopez-Lavalle, Adriana Bohorquez-Chaux

1-General progress

-CIAT-RHUL-UCR Output 2: **Well-characterized parental germplasm for multi-locational/environmental field testing and use in genetic crossing programs.**

Target completion date: 1 November 2019. Percentage completion: 100%

Twenty (20) F1 parental lines more resistant to whiteflies, from the CM8996 (ECU72 x COL2246) and GM8586 (ECU72 xTMS60444) segregation families, well-characterized and phenotyping in greenhouses and field in CIAT. Used for genetic crossing programs.

-CIAT-RHUL-UCR Output 7: **Production of multiple F_2 "advanced intercross" populations.**

Target completion date: 1 June 2020. Percentage completion: 100%.

We are developing twelve F_2 "advanced intercross" populations for resistance to whitefly:

# advanced intercross	Female	Male	Cross	# offspring
1	CM8996-581	CM8996-199	GM12200	193
2	CM8996-596	CM8996-199	GM12201	360
3	CM8996-199	CM8996-199	AM1588	241
4	CM8996-246	CM8996-199	GM12199	644
5	CM8996-758	CM8996-199	GM12202	626
6	CM8996-193	CM8996-199	GM12198	53
7	GM8586-198	GM8586-64	GM13464	291
8	GM8586-198	GM8586-103	GM13465	412
9	GM8586-198	GM8586-153	GM13466	93
10	GM8586-64	GM8586-64	AM1620	110
11	GM8586-49	GM8586-64	GM13463	144
12	GM8586-103	GM8586-103	AM1621	82

The seeds of these offspring were planted in soil, multiplied in cuttings and are being phenotyped in the greenhouse.

F2 advanced intercross Phenotyping:

Statistical analysis of five F2 "advanced intercross": Statistical analysis was done using SAS software 9.3 for Linux with PROC GLIMMIX procedure. The effect of clones on the number of nymphs in the choice experiments performed in the years 2019, 2020, was tested using a generalized and mixed linear model adjusted to a negative binomial distribution to establish mean differences between treatments (Figures 1-4).

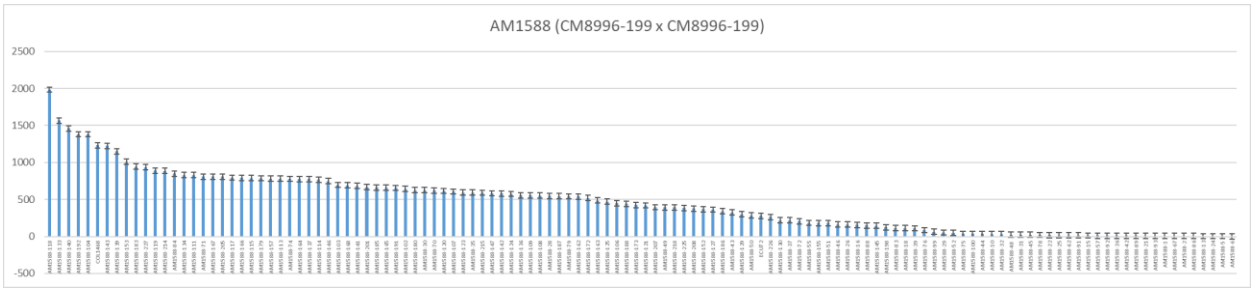

Figure 1: Adjusted Means of nymphs in the F2 autocross AM1588.

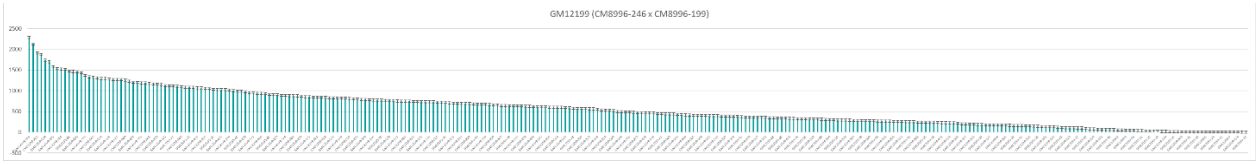

Figure 2: Adjusted Means of nymphs in the F2 advanced intercross GM12199.

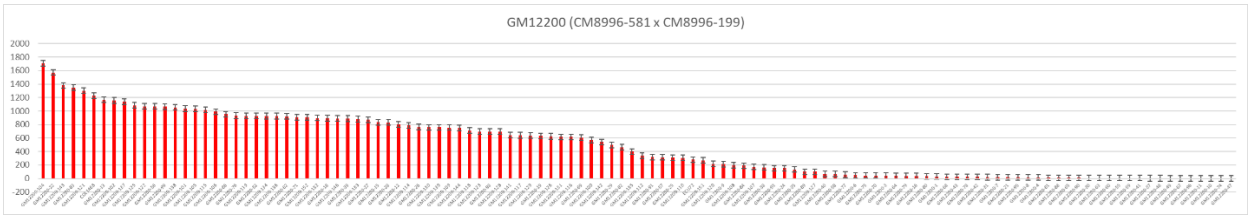

Figure 3: Adjusted Means of nymphs in the F2 advanced intercross GM12200.

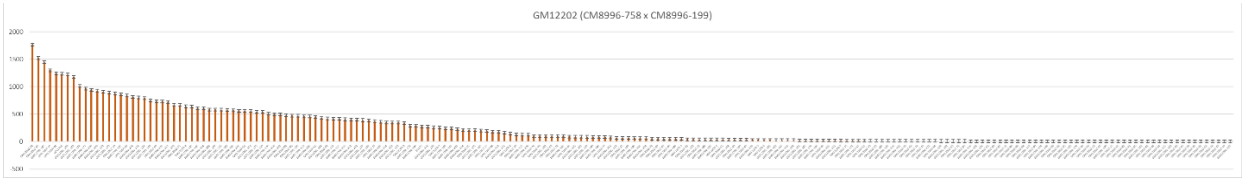

Figure 3: Adjusted Means of nymphs in the F2 advanced intercross GM12202.