Cassava Genetics Laboratory Network: Platforms to support Cassava breeding and research

Adriana Bohorquez-Chaux, Janneth P. Gutierrez, Maria Isabel Gomez, Carmen Adriana Bolaños, Vianey Barrera, Luz Andrea Gomez, Maria Alejandra Bedoya, Camilo Sanchez, Carlos Ordoñez

- **Researches/Cassava Genetics Lab**

Cassava Genetics Lab Facilities and Capacities

oKtopure™

High Throughput DNA extraction

Flow cytometer

Genome size and Ploidy

MinION

Sequencing platform

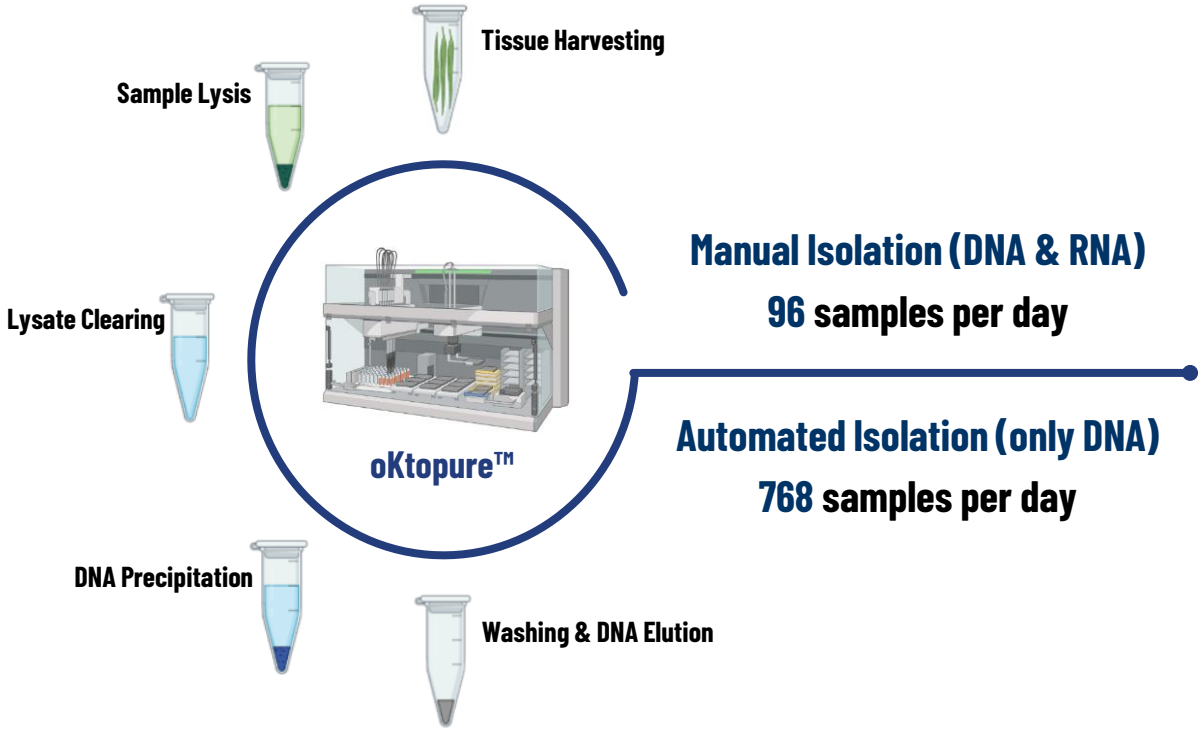
QuantStudio 5

Gene Expression

Fluidigm

Genotyping

DNA & RNA Isolation Platform



Support for breading program

DNA isolation for duplicate test

Automated isolation platform as a potential support for other crops

Cassava Genetics Lab: DNAs extraction for 10X Whole Genome Sequencing Project (Cassava Breeding, Xiaofei Zhang)

Total DNAs extracted and shipping: 6817

GROUP	Trait	#DNAS sent to BGI
Group 1	DVPRG Progenitors, Genomic prediction, GWAS	577
Group 2	DVGST, Genomic Prediction, training population	440
Group 3	BCCOB, Beta-carotene, markers for BC, DM, HCN, WAB	372
Group 4	CQQU2, Cooking Quality, markers for DM, HCN, WAB	258
Group 5	LAEPR, Waxy starch, markers for germination, DM and agronomic traits	180
Group 6	BSGN1, markers for CBSD Resistance (2 batches)	639
Group 7	S1GN2, Selfing population, S1 inbreeding depression, F1 Heterosis	918
Group 8	DVGSB, Genomic prediction, Breeding population	659
Group 11	New progenitors, Genomic prediction, GWAS, Cal. Genetic gain	106
Group 12	Genomic Prediction (2021GSF1C, GSF1T) -- training pop, breeding pop (GST, GSB)	873
Group 13	Genomic Prediction DM, CQ (2021GSF1C, GSF1T) -- training pop, breeding pop (GST, GSB)	1495
Group 28	Global Cassava Breeding (Thailand)	300
	TOTAL	6817

Alliance Bioversity & CIAT

CGIAR

Cassava Genetics Lab: Samples for Genotyping to Intertek (Cassava Breeding, Xiaofei Zhang)

Low-density genotyping with Intertek set

Set 1, DVPRG = Group 1 + Group 11
2020, 2021 progenitors , all progenitors -- validate ALL
Set 2, DVGST = Group 2 + Group 12
Genomic Prediction () -- training pop (GST)
Set 3, DVGSB = Group 8 + Group 13
Genomic prediction () -- GS breeding population – CMD2 +
Set 4, WFCOB
-- CMD + WF selection -- CMD2 +
Set 6, BSGN1 = Group 6
Families of CBSDxCMD – send to Stephan – CMD2

Set 7, DVPRG = Group 1 + Group 11
2020, 2021 progenitors , all progenitors -- validate ALL
Set 8, DMPLY
Selected half-sib families from 2020DMPLY

PacBio genome sequencing project (Xiaofei Zhang)

2022

3 progenitors of the selfing population
(done)

10 genebank accessions with different genome size
(Allele mining project) **In Process for HiC, (10 genotypes PacBio seq, done)**

4 breeding progenitors sequenced by EiB, **(Done)**
4 breeding progenitors supported by NextGen, Ed's team' **(Done).**

10x whole genome sequencing project (Xiaofei Zhang) group

2023
Group 21-29

Group 21, DVPRG
New progenitors (2022DVPAR)
-- Genomic Prediction
-- GWAS
-- Cal. genetic gains
-- Select parents
2022-Oct-02

Group 27 DMF1C
Vietnam breeding pop (2021DMF1C)
2022-Oct

Group 22 & 23, DVGST & DVGSB
Genomic Prediction (2022GSF1C, DVF1C)
-- training pop, breeding pop. (GST, GSB)
2022-Oct
Group 22 will be ready in 2023 Feb, ~800
Group 23 will be ready in 2023 Aug, ~1000

2022-Oct
Group 24 will be ready in 2023 Feb, ~400
Group 25 will be ready in 2023 Aug, ~ 600

Group 24 & 25, LAGST & LAGSB
Genomic Prediction (2022GSF1C, LAF1C)
-- training pop, breeding pop. (GST, GSB)

Group 26, LAEPR
Waxy starch (2022LAEPR)
-- for future GS
-- Camilo or Jorge Ivan sends stakes

Group 28, global cassava breeding
DVGCB from KU, IITA, Embrapa, NaCRRI, Vietnam
2022-Oct
Had KU, Embrapa..

Group 29, PPD
PPD pop from wild crosses

Cassava Genetics Tissue Culture Laboratory

Cassava Breeding Program Main Purposes

- Yield
- Carotenes-Biofortification
- Cassava Diversity
- Starch
- White fly Resistance
- Harvest Index
- CMD Resistance
- CBSD Resistance
- Drought Tolerance
- Types of starches: Small granule
- Post-harvest deterioration
- Mite Resistance
- Protein content

Introduction of cassava materials to in vitro

Delivery stakes of the materials to be preserved in vitro in the Cassava Genetics Tissue Culture Laboratory

Entry of material to the laboratory from greenhouse and field

There are approximately **3100** genotypes (10-15 copies each) corresponding to **22** traits each with a purpose.

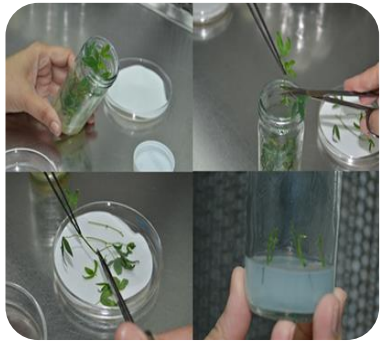
Alliance Bioversity & CIAT

CGIAR

Cassava Genetics Tissue Culture Laboratory

Thermotherapy to clean

Multiplication of genotypes to conserve



Multiplication for shipments to Greenhouse and/or partners

Genotypes are planted, then hardened in a green-house and subsequently established in the field

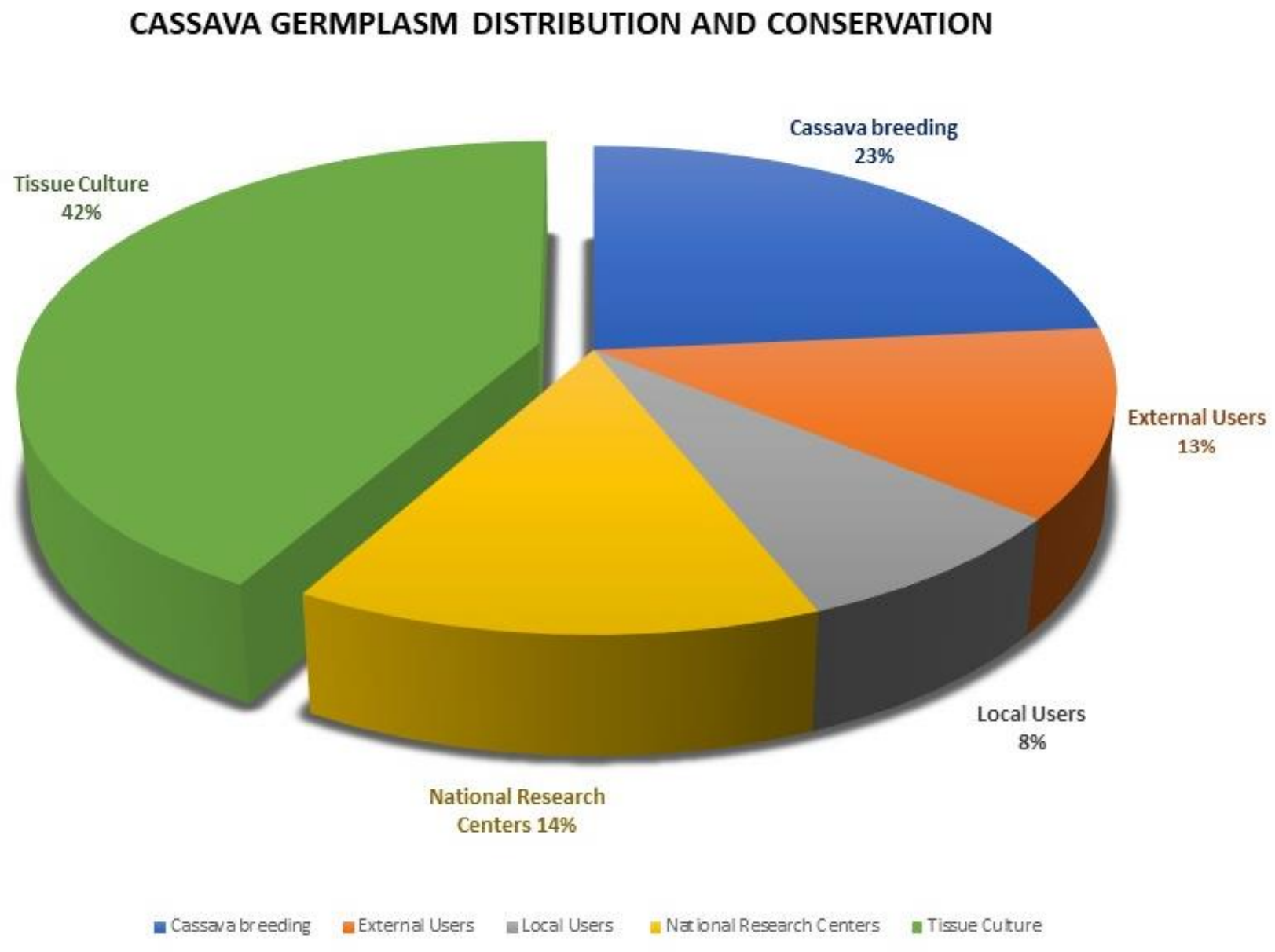
Shipments to partners

Inventory and all processes that are carried out in the laboratory we make supporting us with platform: CIS
https://isa.ciat.cgiar.org/CIS/includes/contents/vistaProcesos.xhtml

Alliance Bioversity & CIAT

CGIAR

Ongoing projects in the Cassava Genetics Tissue Culture Lab

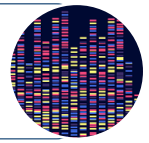


Varietal identification and duplicate test Cassava Genetics Laboratory

Luz Gómez-Martínez, Vianey Barrera-Enríquez, Tatiana Ovalle, Adriana Bohorquez-Chaux, Luis Becerra Lopez-Lavalle

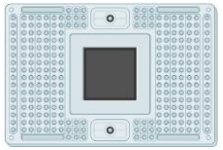
Cassava duplicates test

What is this for?

- Detect genetic duplicates 
-  Genetic diversity studies
- Verify crosses in breeding programs 
-  Select true-cross progeny
- Guide the selection of parental lines 
-  Help to assess diversity, and conservation

How does it work?

Genotyping



Data Collection
96 SNPs

Python 3.8.8

Distance Matrix
Hamming distance

DBSCAN clustering

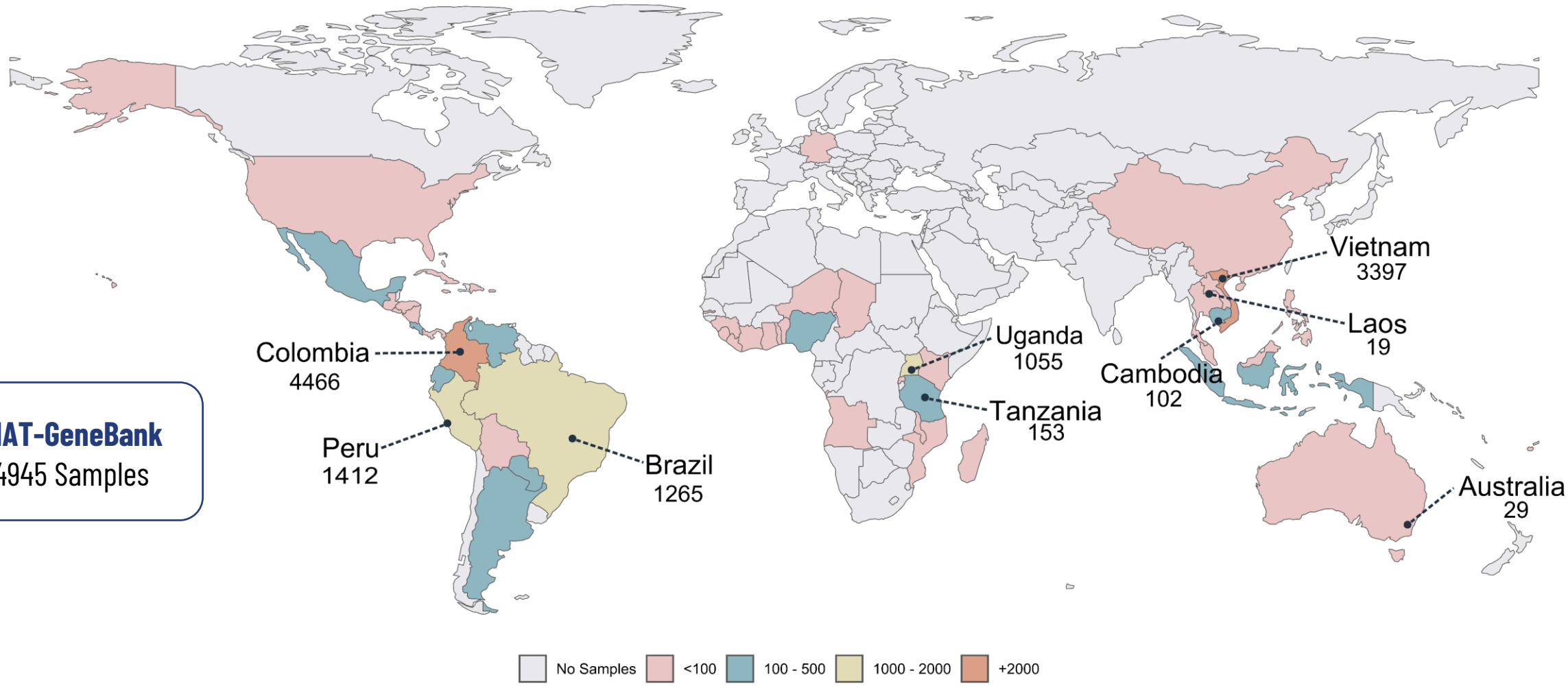
Same Fingerprint = One cluster
Fingerprint without match = Unique

Cassava Reference Data set > 14500 samples

Alliance Bioversity & CIAT

CGIAR

Cassava Reference Data set > 14500 samples



CIAT-GeneBank
4945 Samples

Distribution of samples from the Cassava database in the duplicate test according to their geographical origin.

DNA fingerprinting reveals varietal composition of Vietnamese cassava germplasm (Manihot esculenta Crantz) from farmers

John Ocampo · Tatiana Ovalle · Ricardo Labarta · Dung Phuong Le · Stefan de Haan · Nguyen Anh Vu · Le Quy Kha · Luis A. Becerra Lopez-Lavalle

Varietal composition

92 Vietnamese villages ⟶ **1579** samples
From 6 agro-ecological areas

↓ Varietal identification test

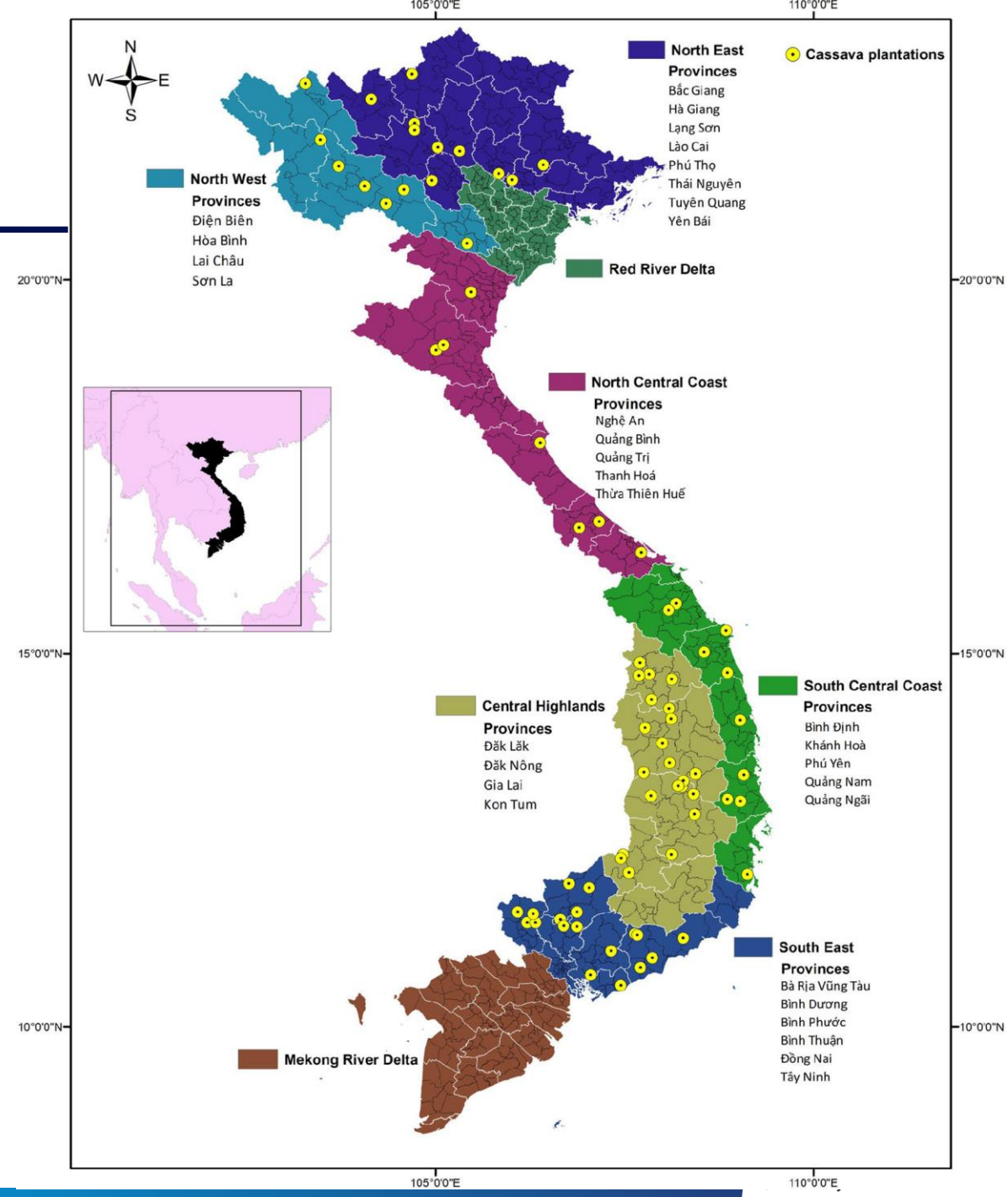
Germplasm reference data set
Latin-American (CIAT, Colombia)
Asia (HLARC, RCRDC, AGI Vietnam)

↓ Cluster by DNA fingerprint

85 Cassava genetic groups
98% have a least one duplicate

↓

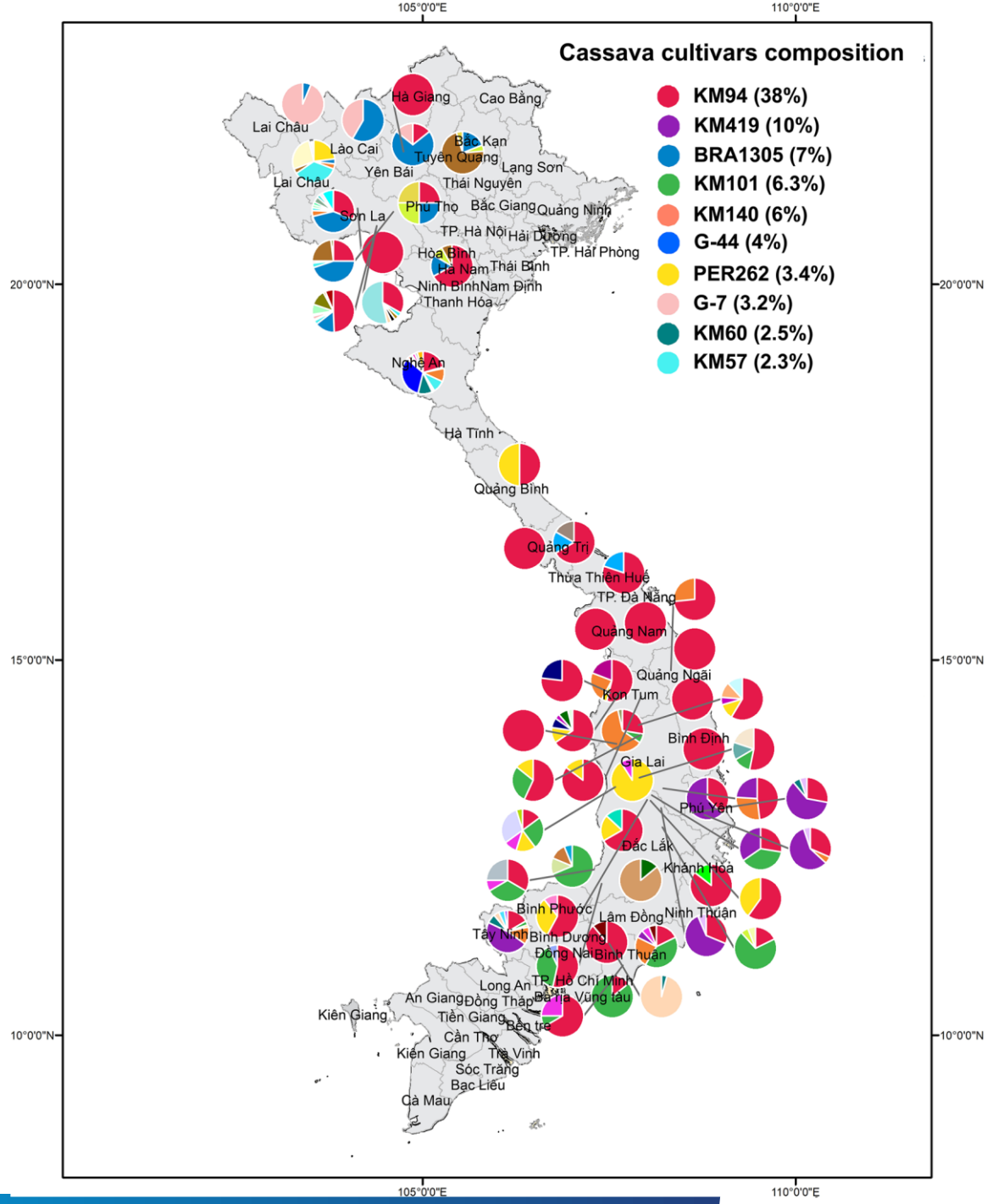
10 genotypes ⟶ **82%** of the distribution
75 genotypes ⟶ **18%** of the distribution



Varietal composition

Highlights

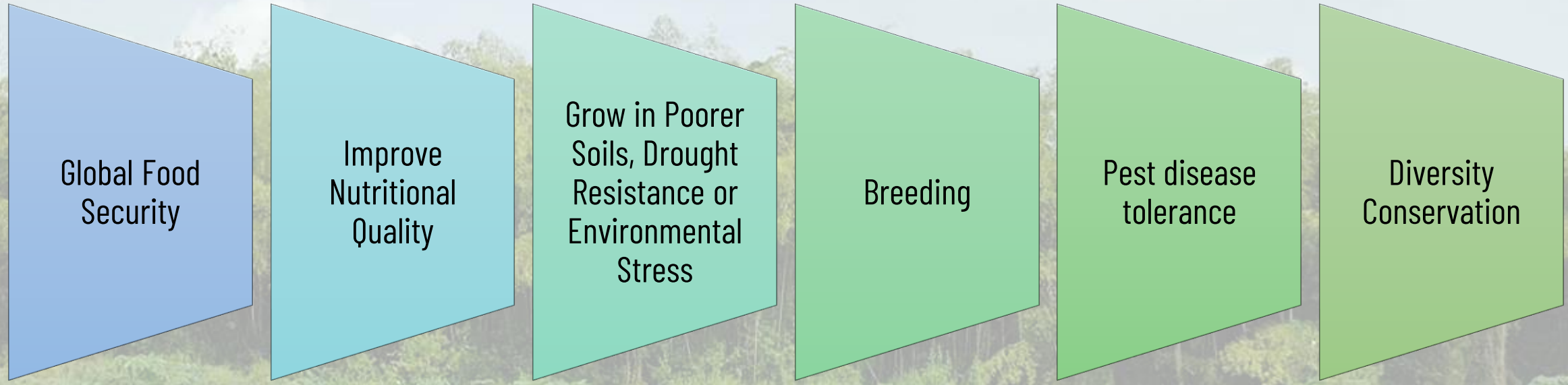
1. KM94 ⟶ KU50 is grown by 38% of farmers across all agro-ecological zones
2. BRA1305 and PER262 represent 10.4% of de frequency distribution
3. **G-44** in North Central Coast ⟶ "High yielding" and **G-7** in North ⟶ "Hybrid"
4. 97 different names for 85 genetics groups
5. 31 unique genotypes ⟶ landraces **Enrich the primary gene poll**



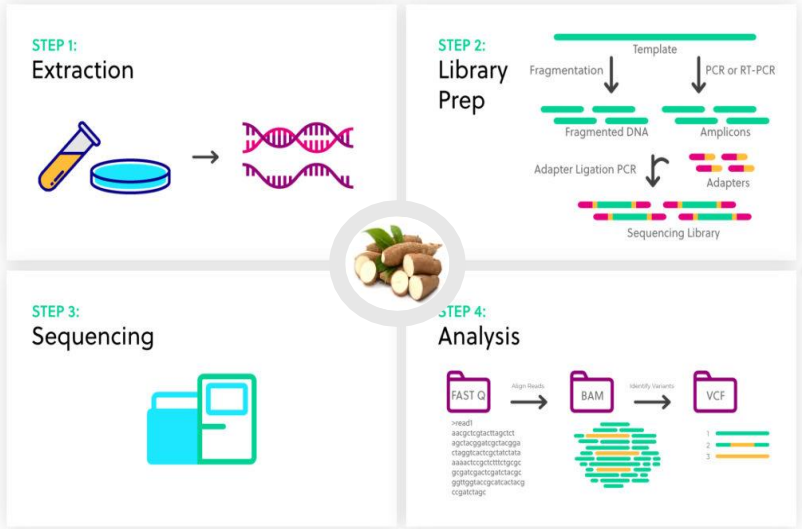
Cassava Bioinformatics Platform

Genomics to untangle relevant traits

Vianey Barrera-Enriquez
Camilo E. Sanchez
Adriana Bohorquez-Chaux
Xiaofei Zhang


Global Food Security
Improve Nutritional Quality
Grow in Poorer Soils, Drought Resistance or Environmental Stress
Breeding
Pest disease tolerance
Diversity Conservation

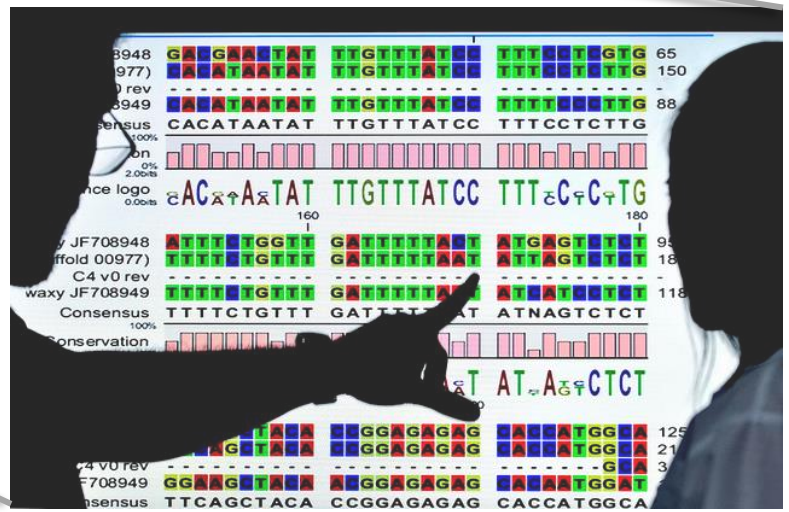
From Field/Greenhouses to Genomic Data



Identification of genes

Identification

Discovery of molecular markers for early diagnosis

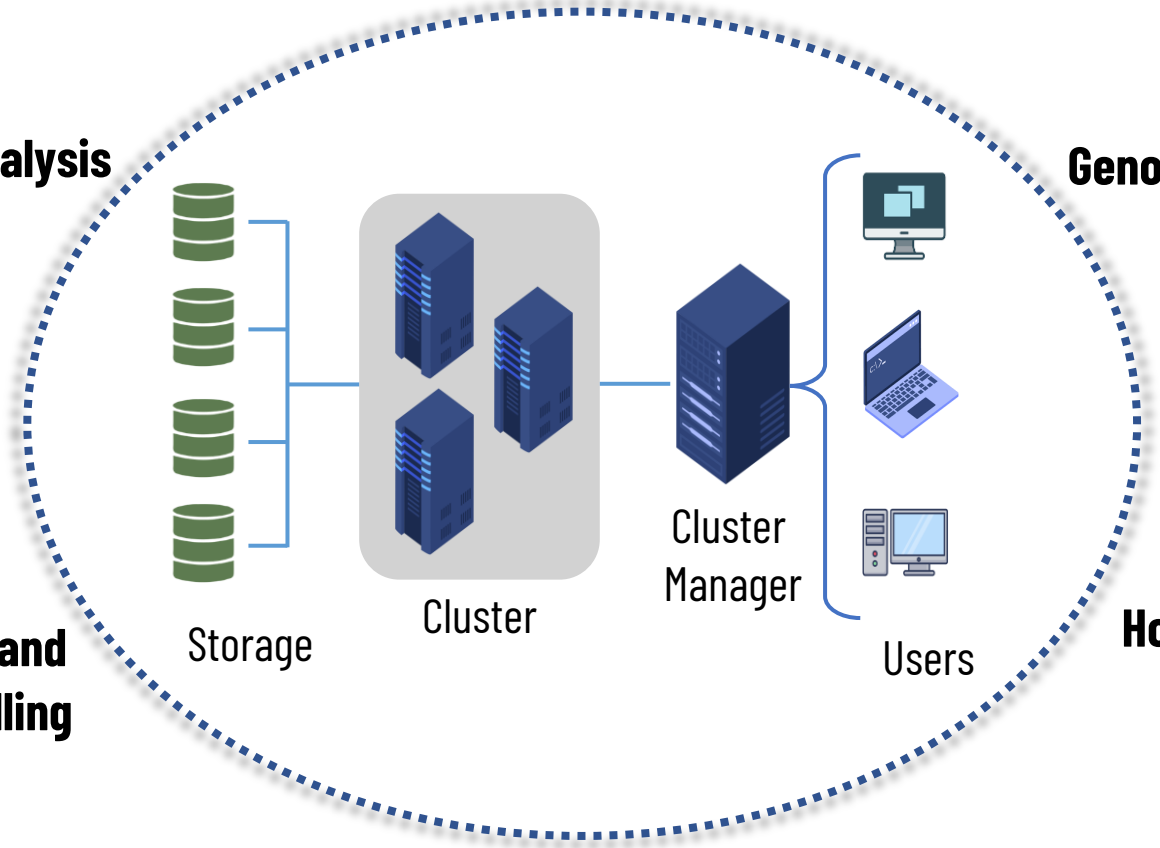


Prediction

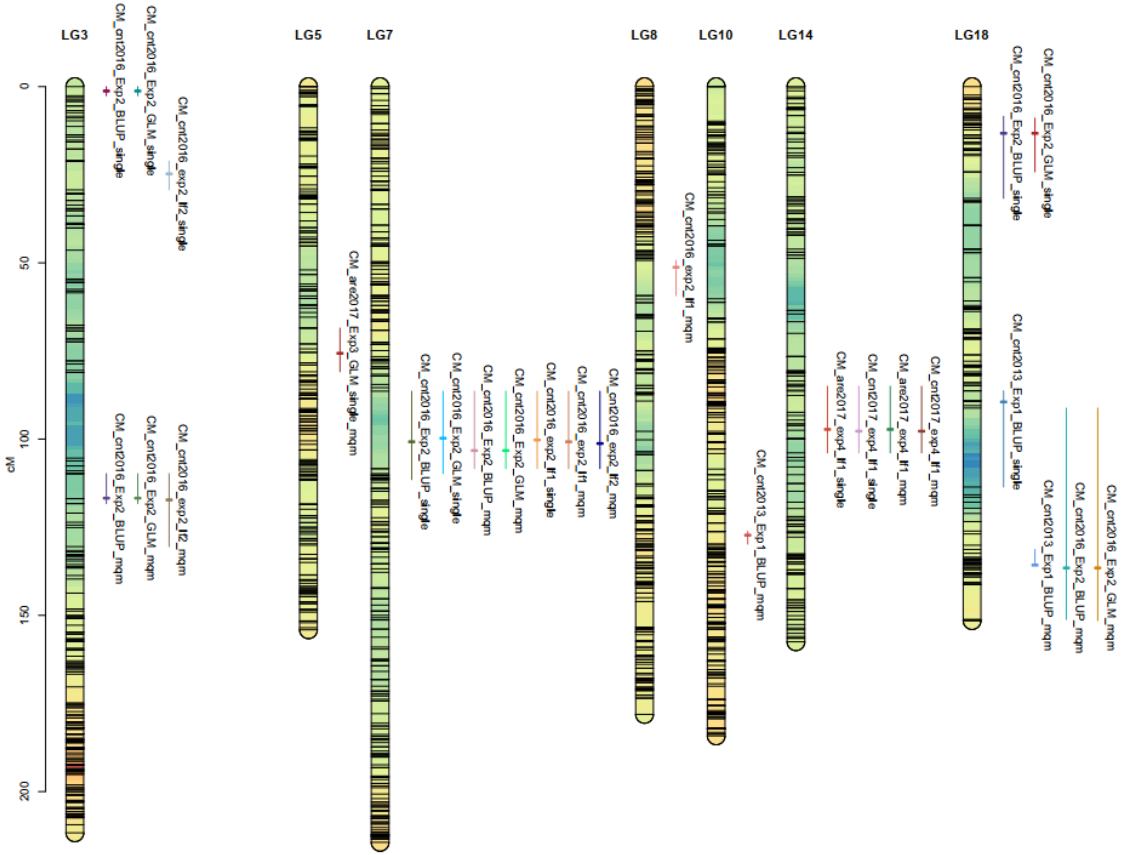
Understand the genetic architecture behind a trait

From Data to untangle relevant traits

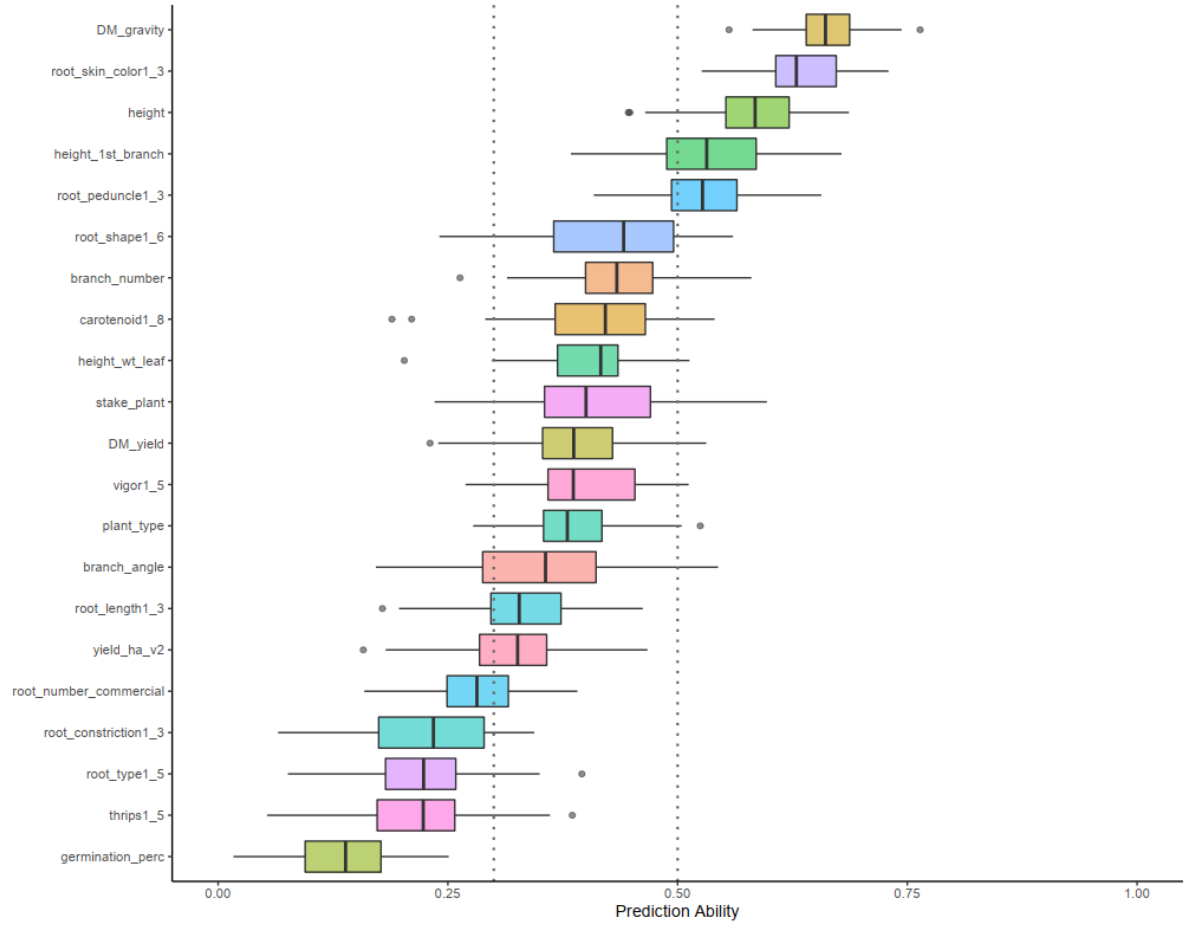
Cassava NGS collection: ~10000 (Whole Genome, RADseq, DartSeq, RNAseq)
Cassava Genotyping Collection: ~14000



Some results from Genetics Lab and Breeding Team



Whitefly CM8996 - QTLs



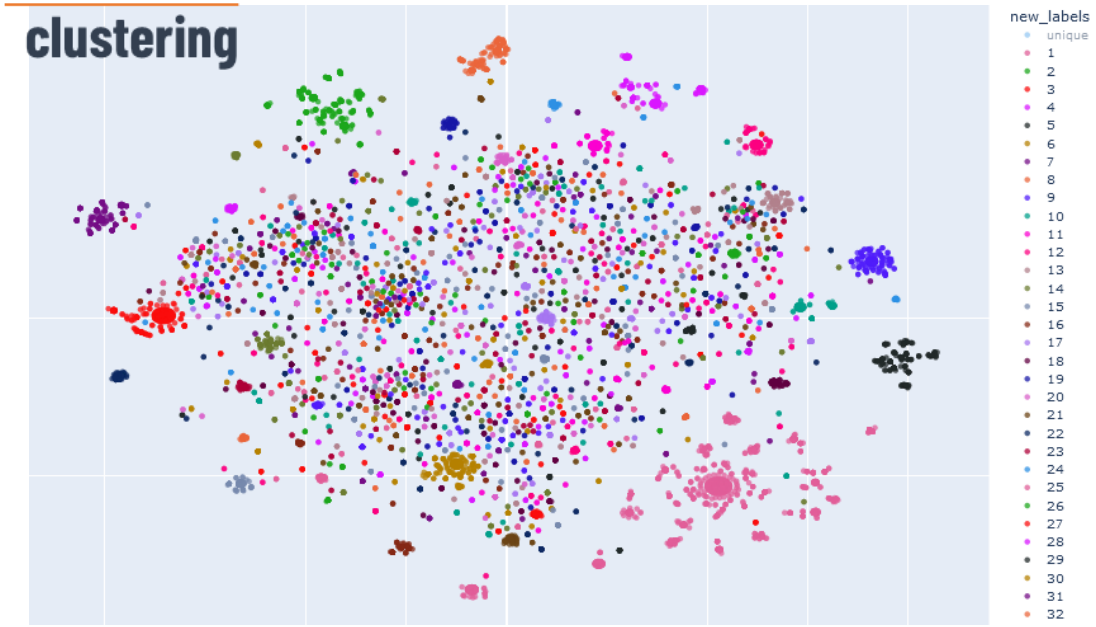
Genomic Prediction

Alliance Bioversity & CIAT

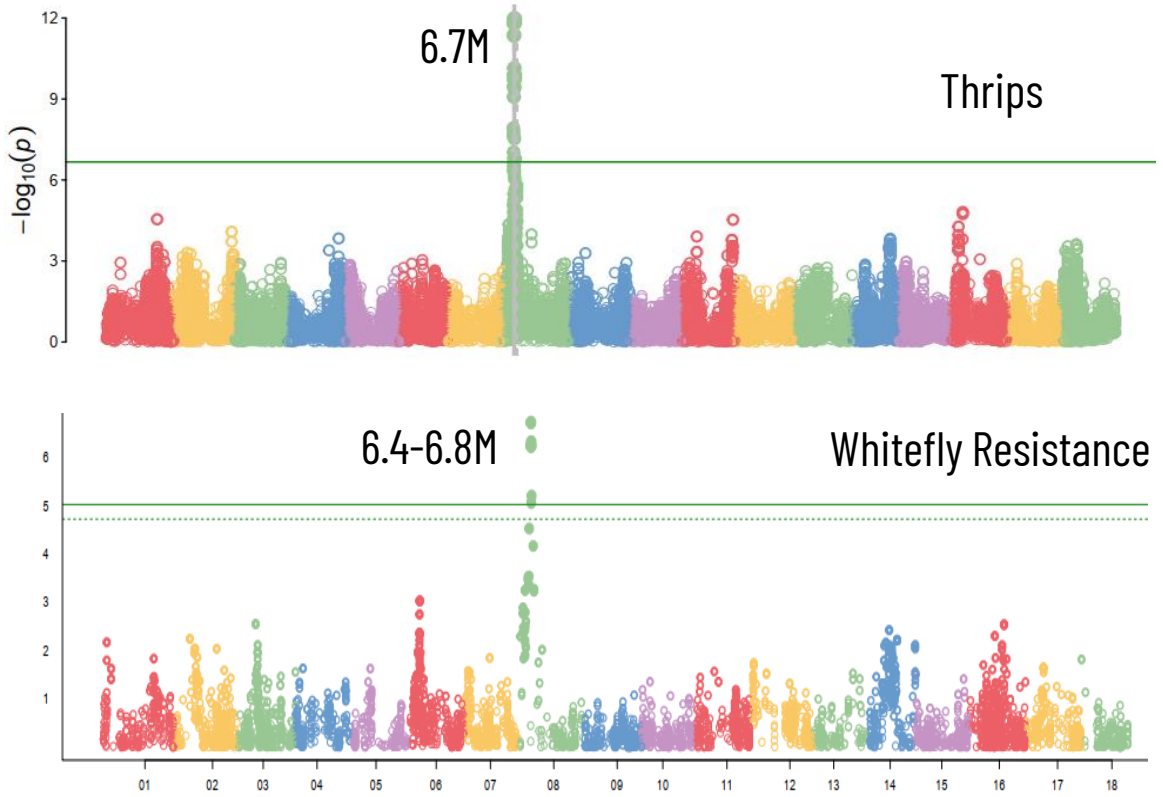
CGIAR

Some results from Genetics Lab and Breeding Team

DBSCAN clustering



Duplicated Analysis



GWAS

Alliance
Bioversity & CIAT

CGIAR

Thanks!