IITA Transforming African Agriculture CGIAR	**Crop:** Cassava **Function:** Phenotyping Data Management	**SOP #**	IITA-CS-SOP05
		Revision #	IITA-CS-SOP05-01
		Implementation Date	6/10/2024
Page #	1 of 18	**Last Reviewed/Update Date**	5/6/2024
SOP Owner	Prasad Peteti	**Approval Date**	5/28/2024

STANDARD OPERATING PROCEDURES FOR PHENOTYPING DATA MANAGEMENT IN IITA CASSAVA BREEDING PROGRAMME

Authors & Contributors

Prasad Peteti, p.prasad@cgiar.org;

Afolabi Agbona, a.afolabi@cgiar.org;

Ukoabasi Ekanem, u.ekanem@cgiar.org;

Toyinbo Seyi, o.toyinbo@cgiar.org;

Dhikrullah Mahmud, m.dhikrullah@cgiar.org;

Adetuyi Mofeoluwa, m.adetuyi@cgiar.org;

Mberu Esther, e.mberu@cgiar.org;

Ismail Rabbi, i.rabbi@cgiar.org

Peter Kulakow, p.kulakow@cgiar.org;

1. Introduction

Data management involves data collection, curation, analysis, and storage to ensure accessibility to quality data. The goal of data management is to provide breeders with relevant information and tools that help to take better breeding decisions.

2. Purpose

This document describes the procedure for phenotypic data collection and management in cassava breeding activities.

3. Scope

This SOP contains the descriptive procedure required for day-to-day data management in cassava breeding. It covers the phenotypic data management like designing field layouts, generating barcode labels, data collection, curation and storage. This SOP version does not cover genotyping, crossing, soil and weather data managment.

4. Definition of terms

Please refer to the Annexure

5. Roles and Responsibilities

Breeder: responsible for planning, implementing data management strategies, trial design based on product profiles and deciding on the number of clones, environments and key traits to be evaluated.

Data Manager: oversees various data management activities within a breeding programme and ensures the deployment and use of the breeding database (www.cassavabase.org). Training students, technicians and supervisors on data management using Cassavabase.

Data Analyst: Biometric support for analysis and selections from phenotyping trials using appropriate statistical software. Backstops staff members and research partners develop field design and conduct statistical analyses of field and laboratory data.

Data Entry Technician:

Carry out digitization and data curation before analysis and uploading data to Cassavabase. Assist the data manager and analyst in data management activities.

6. Procedure/Protocols

Steps involved in phenotypic data management:

6.1.1. Experimental Design

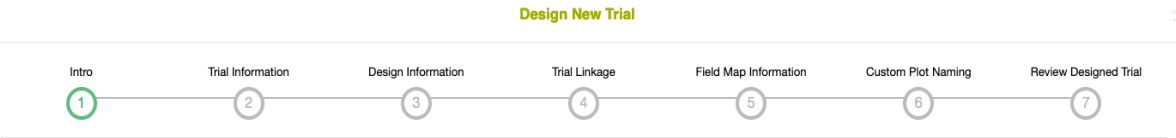
Experimental type	Design	Number of plants per plot	No of reps	No of locations
Seedling Nursery		1	1	1
CET	Augmented	5	1	1
PYT	Alpha Lattice	5 or 10	2	2

AYT	Alpha Lattice	20 or 28	3	4
UYT	RCBD	42	3	6
NCRP	RCBD	42	3	10

6.1.2. Create New Trial in CassavaBase

Log into CassavaBase to add a new trial. The steps followed are:

- Navigate to "*Manage*" and select "*Field Trials*".
- Begin the "Design New Trial" workflow by clicking on "*Design New Trial*".

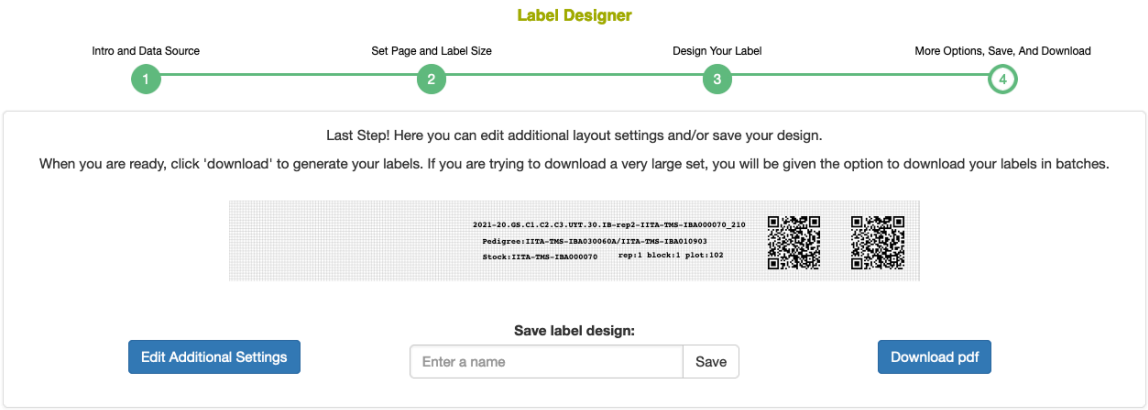


- Enter all relevant metadata in "Trial Information" (e.g. trial name, location, plot and field dimensions, trial type, design type, description)
- Enter "Design Information" (e.g. list of genotypes in the trial (list should already be defined in CassavaBase).
- Enter "Trial Linkage" information if applicable.
- Enter "Field Map Information" (here you can specify how the row and column numbers will be generated for plots in the trial)
- Custom Plot Naming (optional)
- Review Designed Trial
- Save a new trial in the database.

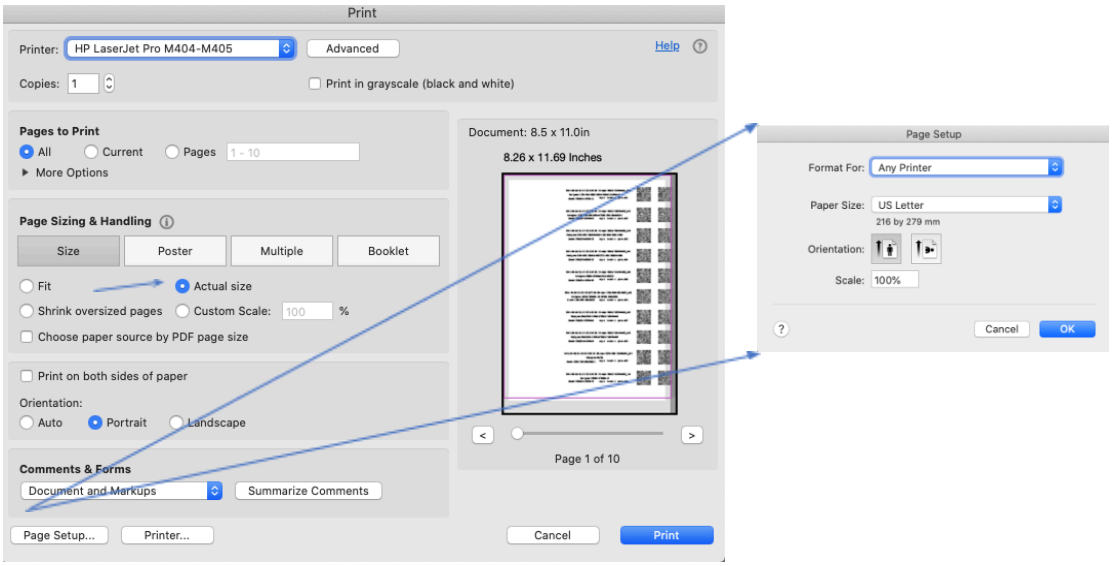
6.1.3. Print field labels

Designing a field label using CassavaBase label designer

- Navigate to “Manage” and select “Label designer”
- Select a data source and a data level then click “Next”
- Select a page format and label format. You may retrieve these settings from a saved design, otherwise, you can start a new design.
- Design your label by selecting the type of barcode which should be a 2D Barcode (QR code). Also, select the field and size, and then click “add” to generate your new label element in the draw area.
- Change the “type” to “Text (PDF)” to add other fields.
- The exported file will contain amongst other fields the plot name, pedigree, accession name, plot number, block number and replication number.
- Edit additional settings and Download the pdf



- Always open the downloaded pdf in Adobe Acrobat reader and check if the labels look fine
- Press the Ctrl and P buttons (Windows) or Command (⌘) and P button (Mac) on your keyboard at the same time to print



- You must always select the actual size in the “Page Sizing & Handling” options, and under “Page Setup”, select the page you have chosen in step 2 while designing the label

Note: Always try to print the first page only at first to see if labels are printed correctly. After confirmation, you can go ahead and print other pages, otherwise adjust your design and try again.

6.1.4. Tagging of labels

Shows how the labels can be properly tagged in the fields for easy and accurate data collection to improve the quality

Tagged barcode labels in the field

6.1.5. Field data collection using the Field Book App

Two input files are required for data collection

1. Field Layout
2. Trait file

The files can be generated on CassavaBase

6.1.5.1 Creating a field layout file based on the design of the field trial

- Navigate to “Manage”, click “Field Trials” and select the desired trial to open the “trial detail page” to achieve that, double-click the desired trial.
- Scroll down the page and expand the Upload Data Files section.
- On the Android Field Book Layout row click on “Create Field Book” to generate the layout file.

- A new dialogue box will appear with the trial details and options to select the spreadsheet format and desired data level.
- Click "Submit" and click on the link that pops up to download the layout file.

6.1.5.2. Creating a trait file from the list of traits

- Navigate to **menu** item Search Traits to open the "**trait search**" page
- Search desired traits by ID, name or description
- Select the desired trait in the results.
- Add selected results to a trait list (This could be an existing trait list or a new trait list. It is possible to create a new trait list on the fly).
- Navigate to the menu item "**Manage**" then "**Field Book App**".
- Find the heading "**Trait Files**" and click on "**New**" to create your file.
- A new dialogue box will open. Select the trait list that you created from the drop-down list.
- Check the box titled "**Include Notes Trait**" if you would also like to record and upload general plot notes in the field.
- Type in an appropriate file name and click on "**Submit**" to download your traits file.

6.1.5.3. Downloading the field layout file and trait file from the database to your computer

Your field layout and trait files are now downloaded to your computer. You can copy them onto your tablet or copy them to your team's cloud storage like dropbox or OneDrive

To copy the field layout file and trait file to the tablet with the Field Book App

- Connect your android tablet to your computer using a USB cable
- Copy the field layout file into the Field Book > field import folder in the tablet

- Copy the traits list files into the Field Book > trait folder
- Open the Field Book App to import the new field layout and Traits files.

Note: You can also import the field layout and trait files directly from your team's dropbox or OneDrive cloud storage.

6.1.5.4. Collecting phenotypes

An important aspect when collecting phenotypic data is to use barcode labels to avoid errors. You can use different trait files for various stages of evaluation, e.g. 3 MAP, 6 MAP, Harvest, etc. The ranges specified for each trait provide validation for quality data collection

Note:

1. Never use zero as missing values
2. Always export or save your data after data collection
3. Regular back-ups should be taken on the computers,cloud server.etc.

6.1.5.5. Exporting phenotypes from Field Book App to your computer

1. Connect your android tablet to your computer using via a USB cable
2. Tap "Export" under settings to export data collected
3. You can either export in database format, or table format. Table format is in the form of spreadsheet.
4. The file name can be edited under "Filename" to suit the needed names to be given to a trial based on its type and other information about the trial.
5. Access the "field export" sub-folder in the "Fieldbook" folder
6. Data exported in table format can be accessed in spreadsheet format (.xls) on the PC.
7. All trial data collected are in the "field export" sub-folder in the "fieldbook" folder.

6.1.6: Data Curation

It is important to curate data by checking it for inconsistencies and correcting it if necessary. Currently, In-house R scripts is use to perform this task. Data are merged on computer, curated and then uploaded to CassavaBase. In the future, data curation will be made possible within Cassavabase.. However,. Data curation steps are as follows :

1. Metadata – trial name, source, no of clones, reps, plot size, spacing, DOP, location, experiment design, the person(s) responsible
2. Check the layout file heading, it should include: plot_name, accession, plot, block_number, is_a_control, rep_number, row_number, col_number
3. Check the trait name headings
4. Validating the layout content
    a. plot_name – it should be combination of year + trial name + rep + accession + plot
    b. accession – It should be validated through cassavabase using a list
    c. plot number – use 101,102... For AYT, PYT, UYT trials 1001,1002... for SN and CE
    d. source – it should be available trial name(s) from cassavabase
    e. rep – check for equal number of accession in each rep, check for clone names in each rep
5. Validating the trait values
    a. Check the file for special characters and remove – NA, “.”, dead, “?”
    b. Check for the out-of-range values – each trait should have min and max values

6.1.7. Upload collected data to CassavaBase

All field trials data are uploaded to Cassavabase immediately after curation.

6.1.8 7: Data analysis

The trial data are downloaded and analysed using R statistical tools for trial data analysis. Some key analysis tools are embedded in Cassavabase and can be used as well. The following data analyses are performed on trials after each evaluation.

a. Mean of each accession
b. Overall mean for each trait
c. Min and Max value of each trait
d. Standard deviation of each trait
e. Standard error of each trait
f. Distribution and outlier detection of each trait
g. Coefficient of variance of each trait
h. Probability of the F statistic: Pr > F

Annex A. Cassava naming conventions

1. Abbreviations

1.1. Institution code

IITA - International Institute of Tropical Agriculture

1.2. Species code

TMS - Tropical *Manihot* Species

TME - Tropical *Manihot esculenta*, a prefix for all IITA landrace accessions

1.3 Trial Stages

SN	:	Seedling Nursery
CET	:	Clonal Evaluation Trial
PYT	:	Preliminary Yield Trial
AYT	:	Advanced Yield Trial
UYT	:	Uniform Yield Trial
NCRP	:	Nationally Coordinated Research Program
OFT	:	On-Farm trial
DCT	:	Demand Creation Trial

1.4 Locations

IBA	:	Ibadan	ABU	:	Abuja
IKN	:	Ikenne	ZAR	:	Zaria
AGO	:	Ago-Owu	MOK	:	Mokwa
UBJ	:	Ubiaja	KAN	:	Kano
ONN	:	Onne	MM	:	Malam Madori

1.5 **Experimental Design**

RCBD: Randomized Complete Block Design

CRD: Completely Randomized Design

1.6 Others

MAP: Months after planting
DOP: Days after Plantin
NA: Not Applicable

3. Examples

3.1. Trial name example

21pyt48highTCIB 21 - Year of trial establishment pyt - Trial stage 48 - No. of Genotypes used highTC - Purpose of the trial IB – Location	**20.GS.C1.C2.C3.UYT.30.IB** 20 - Year of trial establishment GS - Genomic Selection C1.C2.C3 - Cycles UYT - Trial stage 30 - No. of Genotypes used IB - Location

3.2. Plot name example

2021-20.GS.C1.C2.C3.UYT.30.IB-rep1-TMS17F1377P0077_103

2021 - Year of Harvest
20.GS.C1.C2.C3.UYT.30.IB - Trial name
rep1 - Replication
TMS17F1377P0077 - Accession or Clone name
103 - Plot number

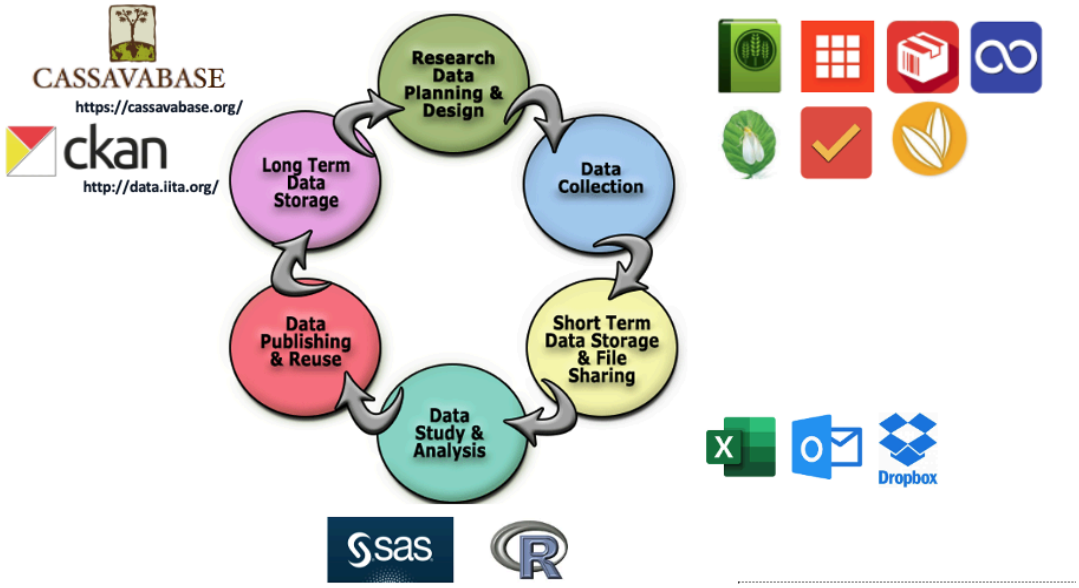
3.3. Accession name example

IITA-TMS-IBA210134 IITA - Institute name	**TMS17F1377P0077** TMS – Species name

TMS - Species name IBA - Location name 21 - Year 0134 - Four-digit Identifier	17 - Year F1377 - Family four-digit identifier P0077 - Progeny four-digit identifier

7. References

8.Annex: Forms/Templates to be used for monitoring and data collection



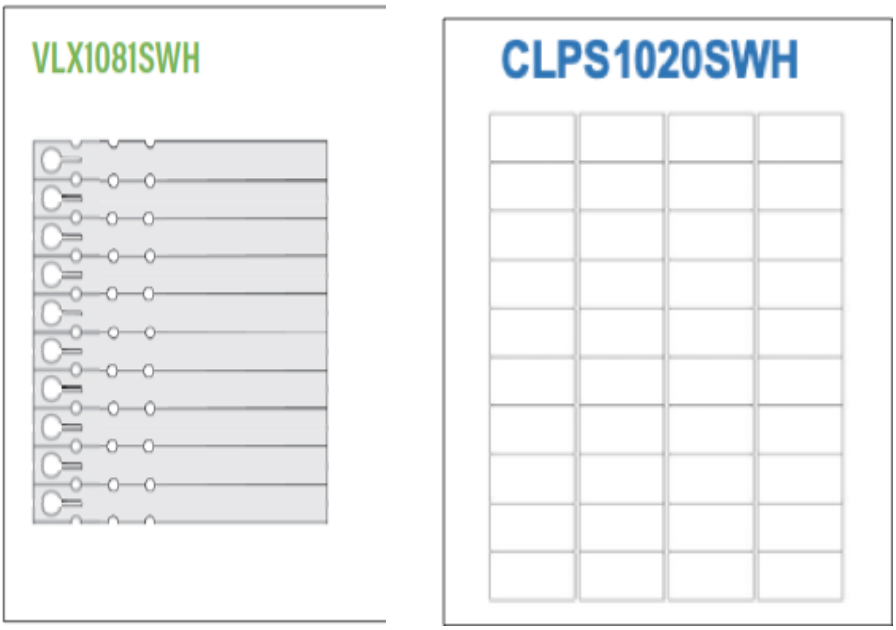
Workflow of Data management Activities in Cassava Breeding unit

Tools and Supplies

- **Tablets or Mobiles**

- **Laser jet Printer**

- **Barcode labels**



VLX1081SWH (1” x 8 ½” Laser slip on tag): Regular Labels for field use

CLPS1020SWH (2 x 1 laser label, 40 labels per sheet): Sticky labels for lab and field use

PhenoApp tools

- Fieldbook - This is used for collecting field evaluation data
- Coordinate - This is used for Genotyping sample tracking
- Intercross - This is used for collecting Crossing information
- Whitefly count -This is used for counting whitefly images on cassava leaf
- Inventory - This is used for collecting sample weights

Fieldbook

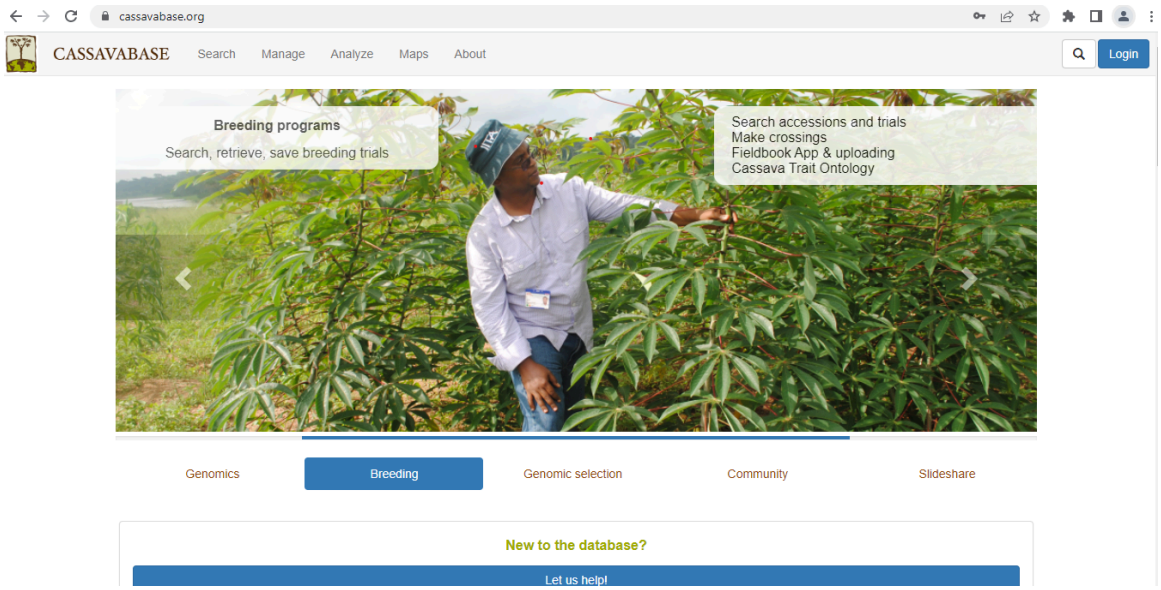
Coordinate

Intercross

Whitefly count

Inventory

Cassavabase is an open-access digital ecosystem for data management which includes trial design, data collection, analysis, and storage.



www.cassavabase.org