	Crop: Yam **Function:** Data Collection	**SOP #**	IITA-YM-SOP03
		Revision #	IITA-YM-SOP03-01
		Implementation Date	22-08-22
Page #	1 of 8	**Last Reviewed Date**	22-08-22
SOP Owner	Lead Breeder (Amele Asrat)	**Approval Date**	22-08-22

Standard Operating Procedure for Yam Data Management

Authors & Contributors:

Agre Paterne P.Agre@cgiar.org

Alex Edemodu A.edemodu@cgiar.org

Adenike Olatunji aa.olatunji@cgiar.org

Amele Asrat A.Amele@cgiar.org

1. *Introduction*

The yam breeding generates thousands of genotypic and phenotypic data every year from genetic studies and breeding trials. There is a need for data centralization, curation, management, and sharing. This data-driven approach hinges on strong data management, quality control, and analytics foundations. The breeding program uses data management system called YamBase. The YamBase is used to: (1) track breeding materials, (2) store experimental evaluations, (3) record phenotypic measurements using consistent ontologies, (4) store genotypic information, and (5) manage crossing information and (6) implement algorithms for analysis (Figure 1).

2. *Purpose*

The purpose of this document is to outline the roles, responsibilities, and procedures to be followed in executing workflows and management of trials and phenotyping using the YamBase and digital tools. This SOP also intends to guide the staff in recording data using the breeding data management systems.

3. Scope

This SOP covers trial creation and management in YamBase, list creation, and phenotyping using digital tools. It also covers data validation checks to ensure the accuracy and integrity of the data that will be stored in YamBase.

4. Definition of terms

CG	**Consortium group**
EBS	**Enterprise Breeding system**

YamBase is a relational database that can manage many breeding workflows and data types. It is a type of BreedBase developed by the Boyce Thomson Institute (BTI) based at Cornell University. YamBase can perform the following actions: define a germplasm list, design a field trial, select a trait list based on the standard yam ontology and generate a field book compatible with Field Book App. YamBase also generates unique plot names for every trial created in the database.

Field Book App is an Android-based mobile application for plant phenotyping. The app allows users to import a trial layout file in CSV format and a trait list text file to set up data collection.

Android handheld device/Tablet: A tablet or smartphone that runs the Android operating system and supports almost all the critical features of a regular tablet PC.

Accession/Genotype: a germplasm entry for evaluation in a trial.

Accession list: Group of germplasm materials for an experiment.

Breeding program: An entity that is responsible for the breeding operations. All uploaded data should be associated with a breeding program. Users are also linked to a breeding program.

Data validation: This involves checking whether the observation recorded coincides with the accepted data type or list of accepted values. It also automates formula computation to the traits as necessary.

Data Curation: Removing inconsistent and redundant data from trial data and ensuring the accuracy and integrity of data is stored in the EBS.

Experimental design refers to how the different planting materials are assigned to the available plot positions in the field, depending on the type of trial (mainly on the number of accessions or treatments).

Location: Refers to a stable place such as a: Breeding station managed by a CG Center, Experimental station managed by a NARS partner, Agricultural university, Seed company facility, or Farmer’s field.

Trait list: Group of traits to be evaluated in an experiment.

The trial is a top-level activity in the EBS. Experiments are designed to test or generate germplasm for specific qualities or to observe the performance of existing germplasm. Experiment and Trial are used interchangeably in this document.

Zebra ZM/ZT series Printer, Laser printer: Devices for printing labels.

Zebra Designer Pro 2: Desktop software for printing field and lab barcode labels.

5. Roles and Responsibilities

Scientist: The scientist is responsible for implementing the development of new traits and defining the methodology.

Research Associate/Manager: In the yam breeding, the research associate is responsible for the data curation and uploading into the YamBase.

Supervisor/Technician: Responsible for data collection, cleaning and sharing with the Associate Scientist.

Field Workers: The Field workers are responsible for data collection.

Laboratory Technician: Responsible for leave sampling, genotypic data tracking and linking with the YamBase.

6. Procedure/Protocols

A trial must be created in the YamBase before carrying out the tasks of data collection. The **Scientist** and the Research Associate must ensure that all the agreed agronomic traits are defined in the yam trait ontology. The **Supervisor/Technician, Field Workers** and **Laboratory Technicians** must use the Standard Evaluation System for Yam as their reference in evaluating and scoring the phenotypic data. Manual collection of data using pen and paper is NOT allowed. Data collection must be done using digital tools like the Field Book App. The collected data must be downloaded from the Tablet/Handheld device the same day after collection to a computer or online storage, preferably to avoid data loss. The data must then be uploaded to the Yambase for permanent storage as soon as possible. A typical yam phenotyping workflow is shown below (Figure 1) using YamBase and the Field Book App.

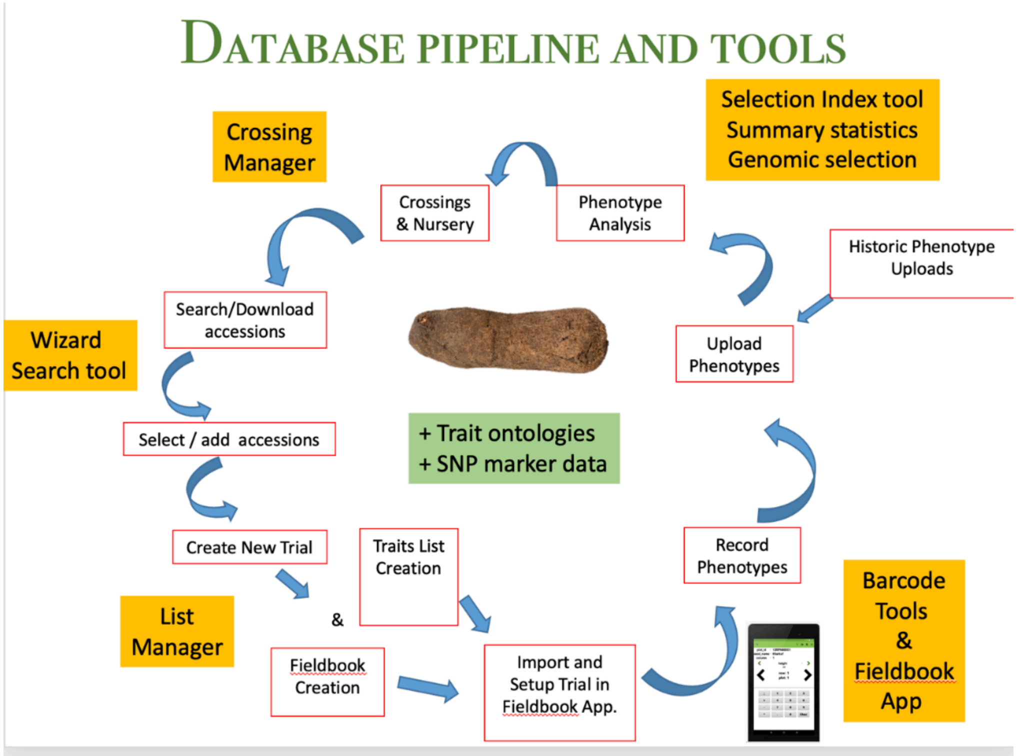

Figure 1: Yam phenotyping workflow

6.1 Creating Trials in YamBase

All trials must be created in the Yambase before carrying out the tasks of data collection. The Research Associate/Manager must ensure that all the agreed trial design protocols and agronomic traits are defined in the yam trait. A trial can be created using the "Design New Trial" tool or the "Uploading Existing Trial(s)" option in the YamBase.

6.1.1 Creating Trials Using Design New Trial Tool

Step 1: Login to YamBase

Step 2: Navigate to Manage Trials.

Step 3: Begin the "Design new trial" workflow by clicking on the design new trial.

Step 4: Enter all relevant metadata in “Trial Information” (e.g. trial name, description, location, plot and field dimensions, trial type, and design type

Step 5: Enter “Design Information,”, e.g. list of accession (genotypes) in the trial (list should already be defined in YamBase)

Step 6: Enter “Trial Linkage” information if applicable.

Step 7: Enter “Field Map Information” (keep the field map display checked and enter the number of plot rows in the field). Custom Plot Naming (leave plot prefix empty, choose plot start number 1 and plot number increment 1)

Step 8: Review Designed Trial (check that the design layout corresponds with the situation of the field).

Step 9: Add Field Management Factors to your design (Optional by clicking the button below). This is required when treatment factors are applied to this experiment. The Field Management Factor Name is most important and will appear later in the field book. Use for this factor name standard treatment names like: “treatment_with_NPK”, “treatment_without_NPK”, “early-seasondrought”, “mid-season-drought”, “late-season-drought” and “irrigation”.

Step 10: Save a new trial in the database

NOTE: The accessions to be used in the new trial must be selected as a defined list of acquisitions (a new list of assets can be determined by navigating to Search Accessions and Plots, searching the intended accessions and clicking add to the new list). The accessions themselves also already must exist in the database. Thus, they have to be uploaded directly to the database as described in **section 6.3**

6.1.2 Uploading an Existing Trial

Sometimes, it might not be possible to design your trial using YamBase due to limited trial design type options or if the design was generated using different software for some reason. Then, the trial design can be uploaded to the database. This can be done by navigating to ***Manage Trials and clicking on Upload.***

Existing Trial(s) and running through the wizard (we recommend uploading every trial separately by choosing ***Single Trial Design***; the required ***.xls file format template*** can be obtained).

6.2 Managing Phenotypes

The list of traits to collect data on depends on the phenotyping experiment's objectives. The **Scientist must ensure all expected traits** for the yam breeding program's product profiles are loaded in the YamaBase. The **Research Associate/Manager** shall list key traits and share them with the Field and Lab Technicians for phenotyping.

6.2.1 Field data collection using the Field Book App

6.2.1.1 The field layout file based on the design of the field trial must be created first for the tablet by following the steps below.

- Navigate to Manage Trials and select the desired trial to open the "trial detail page" for the desired trial by browsing to and double clicking on it.
- Scroll down the page and expand the Upload Data Files section.
- On the Android Field Book Layout row, click Create Field Book to generate the layout file.
- A new dialogue box will appear with the trial details and options to select spreadsheet format and desired data level.
- Click Submit and click on the link popping up to download the layout file.

6.2.1.2 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 go).
- Navigate to 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 traits 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 Submit to download your traits file.

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

After the field layout and trait files are now downloaded to your computer.

The next step is to copy these to your tablet. Follow the steps below to *Copy the field layout and trait file to the tablet with the Field Book App. To copy the field layout and trait file to your android tablet:*

- Connect your android tablet to your computer using a USB cable.
- Copy the field layout file into the fieldBook>field_import folder in the tablet.
- Copy the traits list files into the fieldBook>trait folder.
- Open the field Book app to import the new field layout and Traits files.

6.2.1.4 Collecting phenotypes

An important aspect when collecting phenotypic data at harvest is dealing with missing values. Ideally, rules for that must be included in the data collection device, like Field Book App, but if that is not the case, consider the following guidelines. Missing values occur when:

• All the plants died in a plot due to external reasons (not related to the quality of the genotype).

• For some reason, it was impossible to record a value in a random plot (again not related to the genotype assigned to that plot. yield of a plot that simply was forgotten to be recorded by a technician).

• There were no tubers to evaluate a trait that tubers. For example, tubers are needed to evaluate the dry matter, so if there are no tubers, dry matter is a missing value. A frequent error is to write zero for a missing value or the other way around. A zero and a missing value are different because a value is observed for the trait in the first case, while the second indicates that the value is unknown. This makes a big difference for data analysis and is not a detail! Two important but very different cases are the following:

• If all the plants die in a plot because the genotype was not strong enough, set zero for the number of harvested plants and all the yield traits observed at harvest. However, post-harvest traits like the quality traits that need some tubers for evaluation should be recorded as missing value when the tubers are available to get the evaluations.

• If for some reason, the results of a random plot were impossible to observe (e.g. all the plants died because of some external reason like flooding but independent from the genotype that was assigned to the plot, the plot was destroyed by a truck going off the road, …), a missing value must be assigned to the number of harvested plants and all the traits that would be observed at harvest and post-harvest.

6.2.1.5 Exporting phenotypes from Field Book App to your computer

Step-by-step instructions on how to achieve the above are available here. The Field Book App is available on Google Play for free. To download the app, you must be connected to the internet. Internet connection is not necessary during data collection. The App allows you to disable moving to the next field during data collection if no data has been collected. This can be enabled in the advanced settings after installation. Also, GPS location data can be recorded with Field Book App, but mure is essential to have the trial location defined in Yambase when describing the trial. Most Android phones and tablets can be used for field data collection. A typical 7” to 8” android tablet will be enough for the data collection needs of most breeding programs.

6.3 Upload collected data to YamBase

All curated field and lab data of a trial must be uploaded as soon as the harvesting of that trial. To upload trial data:

- Once logon to the Yambase, navigate to “Manage” then “Phenotyping.”
- Under the “Uploaded Phenotype Files”, Click on “Upload”, Choose the “Fieldbook File” option, and browse the .xls file that contains the phenotypic data.
- Click on “Verify” to verify the file format and data quality.
- If successfully verified, click “Store” to upload and save the data.

6.4 Data analysis

Trial data can be analysed by downloading the data to statistical software like R or SAS.

6.5 Downloading Trial Data

- Once logon to the YamBase, navigate to “Manage” then “Field Trials.”
- Click on the “+” by your breeding program name to see the trials.
- Navigate to the specific trial folder or trial name to access the trial.
- Double click on the “Trial” to open the trial detail page.
- Scroll down to and expand the “Phenotyping and Summary Statistics.”
- Click on download trial data to download.

7. References

https://yambase.org/

Rife, T. W., & Poland, J. A. (2014). Field book: an open-source application for the field data collection on android. Crop Science, 54(4), 1624-1627.

International Potato Center. (2020). Standard operating procedures for sweet potato breeding data management. COP Breeding Data Management SweetGAINS. International Potato Center: Lima, Peru.

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