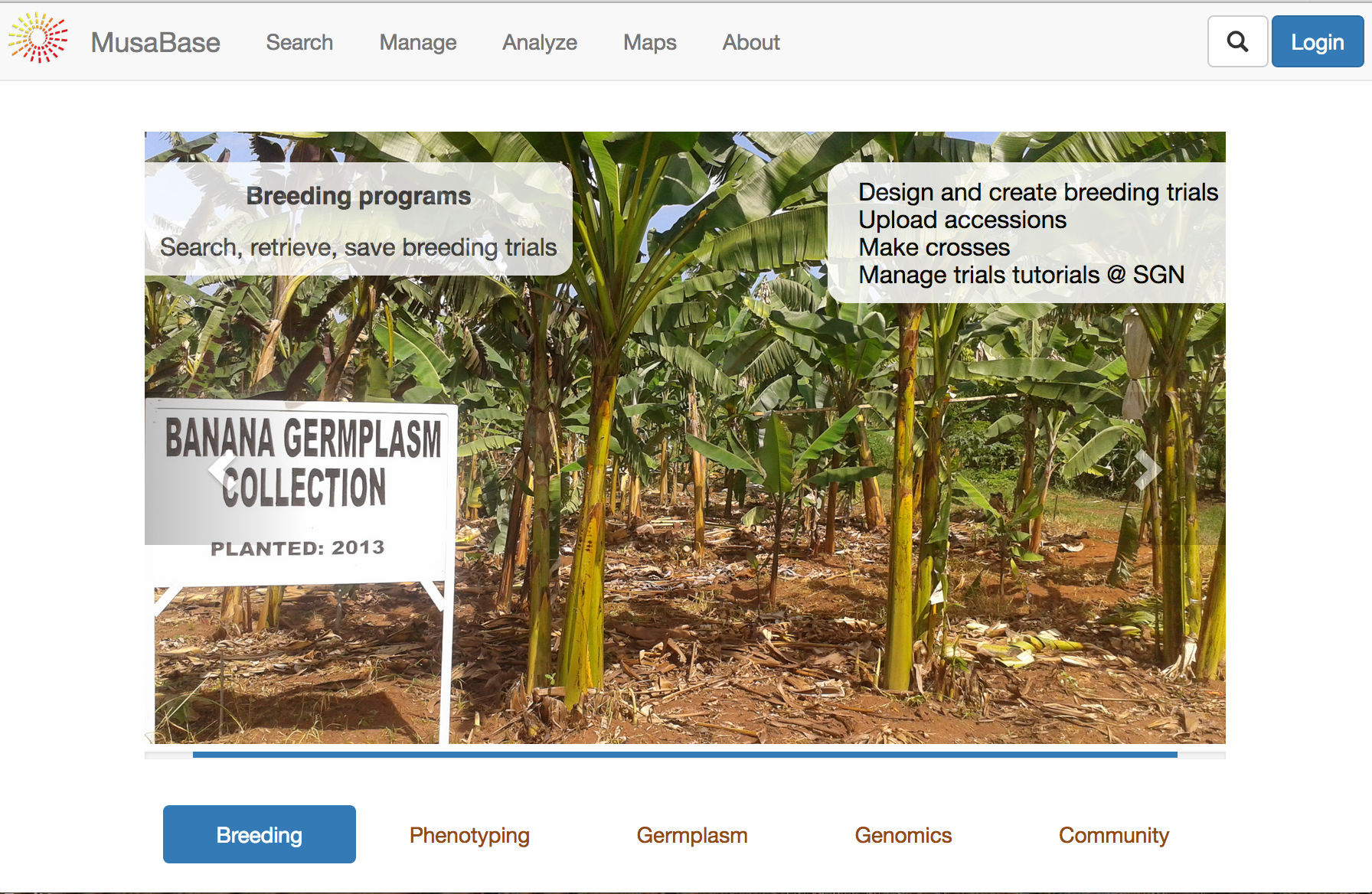


IMPROVEMENT OF BANANA FOR SMALLHOLDER FARMERS IN THE GREAT LAKES REGION OF AFRICA

WP5 Progress report : ‘Harnessing data’

Reporting period: 01 OCTOBER 2016 – 31 MARCH 2017





2017

Table of Contents

[1. Executive summary 3](#_Toc480758477)

[2. Primary and intermediates outcomes 3](#_Toc480758478)

[Table 1. Framework and result tracker WP5. 4](#_Toc480758479)

[3. Results to date 6](#_Toc480758480)

[Table 2. Progress for WP5 year 3 milestones. 6](#_Toc480758481)

[4. Challenges encountered 10](#_Toc480758482)

[5. Lesson learned 10](#_Toc480758483)

[5. Workplan 10](#_Toc480758484)

# Executive summary

BTI continued to work on the IITA funded “Improvement of Banana for Smallholder Farmers in the Great Lakes Region of Africa” with the objective of developing a banana breeding database that provides project partners and Musa researchers and breeders a virtual hub for information exchange, R&D collaboration and enhanced adoption of new hybrids.

The East African Highland banana breeding project annual meeting took place in April 2016 and identified several components where development and improvement where the breeding database work package (WP5) could contribute effectively:

1. Musabase code development to Include in experimental data (field layouts) in the database a sub-plot level to include developmental Cycles, specific to banana breeding.
2. Initiate and co-lead with WP4 discussions across all work packages to establish an appropriate banana trait ontology that can cover breeding (WP1,2,3) and other research aspect (WP4).
3. Provide trainings to the different breeding programs with an emphasis on data collection (using tablet and field book application) and trial management through musabase. Take advantage of the training sessions to discuss future developments.
4. Trial data upload and germplasm curation with breeders.

# Primary and intermediates outcomes

1. Musabase code developments: In 2016, the database was set up to include in experimental data (field layouts) and a sub-plot level to include developmental Cycles, specific to banana breeding using the fieldbook application and a a new crossing tool was released. Database trainings were initiated to start data uploads.

This year focus is on providing better germplasm curation, finalizing the ontology version released last year, provide a new crossing tool and include trial data analytics.

1. A trait ontology first version was released 2016, covering breeding (WP1,2,3) and supplemental variables (WP4). A new version is now going to be released including new traits (120) and an improved interface for cycles called “postcomposing”.
2. Workshops were held in **October** at IITA Ibadan to the Banana breeding team (3 attendees) and in **April** 2017 a second workshop was held at IITA Kampala to present musabase to WP2 banana pathologists participants (8 attendees). An emphasis was made on introducing the use of digital data collection using the Fieldbook app and its interface with musabase. Pathology traits in the current ontology were discussed as well.



* **Table 1** gives the Framework and Results Tracker for WP5, as presented in the project document.

# **Table 1. Framework and result tracker WP5.**

| **ID** | **Outcome/ Output** | **Indicator(s)** | | **Data Sources  (If Applicable)** | **Targets/ Milestones** | **Actual** | **Variance** | **Year 1** | **Year 2** | **Year 3** | | **Year 4** | **Year 5** | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 8 | Primary Outcome 8   Create a banana Breeding database | Database online | clone cassavabase system with empty database | | Database available online |  |  | - | - | 2017: New version release | - | | | Database available online | |
| 8.1 | Intermediate outcome 25   New database/website created based on Cassavabase software | Database online and accessible by breeders | clone cassavabase system with empty database | | Database available online |  |  | - | - | - | - | | | New base/website based on Cassavabase software | |
| 8.1.1 | Output 60   Banana database/website developed  (Cornell univ.-IITA-NARO-EMBRAPA) | The website can be accessed by breeders | website access logs | | Database available online | NARO and IITA as a partner compiled and submitted a list of Musa accessions available in the NARO fields for the construction of the Musabase by the Cornel University partner: https://musabase.org/ | 0% | na | na | na | na | | | Banana Database/website assessed by banana breeders | |
| 8.2 | Intermediate outcome 26   Adapt user interface/tools for the banana database | Adapted interfaces | feedback from banana breeders | | - |  | na | na | Interfaces adapted based on banana breeder input | na | na | | | Adapted user interface/tools for the banana database | |
| 8.2.1 | Output 61   Customized interfaces/tools developed for banana breeders  (Cornell univ.-IITA-NARO-EMBRAPA) | new interfaces available | feedback from banana breeders | | na | na | na | na | Interfaces adapted based on banana breeder input | na | na | | | Customized interfaces/tools | |
| 8.3 | Intermediate outcome 27   Database populated with historic data from NARO and IITA breeding programs | new database contents | Data provided by breeding programs | | na | na | na | na | Database up to date with historic informatio from IITA and NARO | na | na | | | New database populated with historical breeding data | |
| 8.3.1 | Output 62   Database populated with historic data from NARO and IITA breeding programs  (Cornell univ.-IITA-NARO) | new database contents | database contents | | na | na | na | na | Database up to date with historic informatio from IITA and NARO | na | na | | | New database populated with historical breeding data | |
| 8.4 | Intermediate outcome 28   Banana database integrated in workflows | Breeders use database to advance their breeding program | website access logs, database contents | | na | na | na | Breeders use crossing/field design/phenotyping tools | Breeders use crossing/field design/phenotyping tools | Breeders use crossing/field design/phenotyping tools | Breeders use crossing/field design/phenotyping tools | | | Banana breeding database integrated in workflows | |
| 8.4.1 | Output 63   Use banana database to manage crosses, trials, etc.  (IITA-NARO-Cornell univ.) | database contents | website access logs, database contents | | na | na | na | na | Breeders use crossing/field design/phenotyping tools | na | na | | | Banana database used to manage crosses, trials, etc. | |
| 8.5 | Intermediate outcome 29   Train breeders / IITA staff | pre- and post course surveys | Course attendance lists, surveys | | held yearly | na | 0% | held yearly | held yearly | held yearly | held yearly | | | Breeders / IITA staff trained | |
| 8.5.1 | Output 64   Breeders trained  (Cornell univ.-IITA-NARO-EMBRAPA) | Provided training | Course attendance lists, surveys | | training course held | na | 0% | training course held | training course held | training course held | training course held | | | Breeders trained | |

| **ID** | **Outcome/ Output** | **Targets/ Milestones** | **Progress** | **Variance** |
| --- | --- | --- | --- | --- |

# Results to date

In **Table 2**, we give an overview of the progress of activities to date for the outputs with a milestone in year 3. A more detailed description by primary outcome is given below.

# **Table 2. Progress for WP5 year 3 milestones.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Outcome/ Output** | **Targets/ Milestones** | **Progress** | **Variance** |
| 8 | Primary Outcome 8   Create a banana Breeding database | na | na | na |
| 8.1 | Intermediate outcome 25   New database/website created based on Cassavabase software | na | na | na |
| 8.1.1 | Output 60   Banana database/website developed  (Cornell univ.-IITA-NARO-EMBRAPA) | Interfaces adapted based on banana breeder input | na | na |
| 8.2 | Intermediate outcome 26   Adapt user interface/tools for the banana database | Interfaces adapted based on banana breeder input | na | na |
| 8.2.1 | Output 61   Customized interfaces/tools developed for banana breeders  (Cornell univ.-IITA-NARO-EMBRAPA) | na | na | na |
| 8.3 | Intermediate outcome 27   Database populated with historic data from NARO and IITA breeding programs | Database up to date with historic information from IITA and NARO | na | na |
| 8.3.1 | Output 62   Database populated with historic data from NARO and IITA breeding programs  (Cornell univ.-IITA-NARO) | Database up to date with historic information from IITA and NARO | **-1-** Current trials, field layouts and phenotypic data from IITA Arusha and Sendusu were added during worskops.  **-2-** Field layouts were provided by WP4 and added.  **-3-** NARO trial data is still pending **--4-** More historical data from WP2 is still to come.  **-5-** MGIS germplasm data curation initiated with IITA and NARO. Feedback pending. | 50% |
| 8.4 | Intermediate outcome 28   Banana database integrated in workflows | Breeders use crossing/field design/phenotyping tools | na | **na** |
| 8.4.1 | Output 63   Use banana database to manage crosses, trials, etc.  (IITA-NARO-Cornell univ.) | Breeders use crossing/field design/phenotyping tools | **-1-** Banana trait ontology was established in collaboration with WP1,2,3 and 4. Trait ontology is now available on the test site: http://musabase-test.sgn.cornell.edu/tools/onto/ and received additional comments and suggestions from the workshops (see section 8.5).  => Additional traits added to the current Ontology for WP4. The final ontology release will be done after Annual meeting review including “Post composing” to allow a better handling of Cycles.  **-3-** Development of new nursery and crosses manager to address diallele and multiple crosses. A portable tool, “CIPcross” has been requested to complete the pipeline  **-** | 20% |
| 8.5 | Intermediate outcome 29   Train breeders / IITA staff | held yearly | na | 0% |
| 8.5.1 | Output 64   Breeders trained  (Cornell univ.-IITA-NARO-EMBRAPA) | training course held | Trainings on Fieldbook data collection tool and Musabase data management were provided at two research stations (IITA Ibadan (Nigeria) for WP1 and IITA Kampala for WP2 (Uganda). | 0% |

* 1. Improvement of germplasm curation tool and related information has been implemented on both live and test sites: a fuzzy search allows to find exact and non exact matches with accessions already in the database and synonymize names when useful. A trial comparison tool is currently being developed.
  2. Ontology was over project partner and breeding programs in a unified format and released on the musabase test to be assessed late 2016. The final version with an extra set of traits will now be released on the live site.
  3. Code development has been done to adapt the musabase workflow to the breeding needs: cycles (post composing), nursery and crossing.
* Further developments to improve the data collection workflow and include a recent request for the CIPcross tool.
* Integrate WP4 data
* Keep the ongoing efforts of adoption of tablet-based phenotyping solution and database management for WP2 and WP3 by providing weekly support (skype) and onsite trainings.

# Challenges encountered

-A crossing tool was designed in the database. An additional need for handled device “CIPcross” tool was requested to complement it while in the field.

-Heterogeneity of germplasm naming requires harmonization across work packages. The fuzzy search will help data managers to tackle the problem.

-Reach each work package needs and support: ensure support for digital data collection to WP2 and WP3 in the coming 6 months.

# 5. Lesson learned

Ensure good curation of germplasm through database tools and active discussion between stations/germplasm resources center.

Adapt to the various audiences and needs in the project without losing the interoperability of database tools.

# Workplan

* Integrate WP4 specific data collection tool input format and the new crossing tool “CIPcross” tool input format to musabase
* Embark pathologist in the digital data collection
* Keep developing the ontology and database features specific to banana breeding: ploidy level and germplasm management, develop more trial analysis tool and germplasm management.