BILL & MELINDA GATES foundation

Final Narrative

Use this form to provide your final update to your foundation program officer regarding the results achieved for the entire project. In addition, please provide your perspective on key lessons learned or takeaways and input on the foundation's support of your work to ensure that we can capture and share learnings as appropriate both internally and externally.

The Final Narrative must be submitted in Word, as PDFs will not be accepted.

General Information

Investment Title	Improvement of Banana for Smallholder Farmers in the Great Lakes Region of Africa		
Grantee/Vendor	International Institute of Tropical Agriculture		
Primary Contact	Danny Coyne	**Investment Start Date**	October 1, 2014
Feedback Contact[1]	Danny Coyne	**Investment End Date**	November 30, 2019
Feedback Email[1]	D.Coyne@cgiar.org	**Reporting Period Start Date**	October 1, 2014
Program Officer	Jim Lorenzen	**Reporting Period End Date**	November 30, 2019
Program Coordinator	Amy Pope	**Reporting Due Date**	January 31, 2020
Investment Total	$13,873,600.00	**Opportunity/Contract ID**	OPP1093845
Remaining Funds (If applicable)	$ 0.0		

[1] Feedback Contact/Email: the full name and email of the contact whom foundation staff queries for various surveys.

Submission Information

By submitting this report, I declare that I am authorized to certify, on behalf of the grantee or vendor identified on page 1, that I have examined the following statements and related attachments, and that to the best of my knowledge, they are true, correct and complete. I hereby also confirm that the grantee or vendor identified on page 1 has complied with all of the terms and conditions of the Grant Agreement or Contract for Services, as applicable, including but not limited to the clauses contained therein regarding Use of Funds, Anti-Terrorism, Subgrants and Subcontracts, and Regulated Activities.

Date Submitted	[February 28 2020]	**Submitted by Contact Name**	Kayode Awobajo
		Submitted by Contact Title	Project Resource Manager | Project Development and Administration Unit
		Submitted by Contact Email	K.Awobajo@cgiar.org
		Submitted by Contact Phone	+234 2 700800 IITA

Progress and Results

1. Final Progress Details

Provide information regarding the entire investment's progress towards achieving the investment outputs and outcomes. In addition, submit the Results Tracker with actual results as requested.

Background:

1. This project focuses on drastically improving the speed and efficiency of breeding bananas, in particular starchy staple food bananas in Eastern Africa: it aims to dramatically upscale existing breeding activities, build a breeding and selection pipeline, improve data management, and increase the pace and efficiency of breeding by conducting research for identifying methods for achieving higher rates of seed set and developing molecular markers for early selection of priority traits. We will combine genetic and genomic studies on segregating populations for mapping sources of resistance to the target pests and diseases. This will be complemented by improved characterization of the spread and virulence of the four target pests and diseases in five testing sites and the development and application of faster bio-assay screens. Also we develop a system for better tailoring breeding products and increasing adoption of new cultivars through end-user feedback systems and

participatory evaluation of improved germplasm. All information obtained is deposited in an open-source database to improve breeding efficiency and intensify internal collaboration.

2. Structure of the project: structured around five strategic goals and one management goal:
    1. **Banana Breeding Pipeline:** Enhanced performance of breeding systems to deliver improved East African highland bananas with increased levels of pest and disease resistance, higher yields, and better consumer acceptability.
    2. **Pest and Disease Control:** Enhanced host plant resistance to major pest and disease constraints through improved pathogen identification and accelerated early stage screening of resistance.
    3. **Leveraging Genetics:** Improved breeding efficiency through molecular-based genetic studies for increased understanding of underlying genetics and development of DNA marker-based early selection.
    4. **Empowering End-user Evaluation:** System for better tailoring breeding products and increasing adoption of new cultivars through end-user feedback systems and participatory evaluation of improved banana germplasm.
    5. **Harnessing Data:** Driving improved efficiency of breeding systems and enhanced synergy in national, regional, and global partnership through an open-source database and tool box for banana breeders and researchers.
    6. **Governance, Research Oversight, and Management**: Coordinating breeding efforts integration, capacity building, communication, and dissemination undertakings and to ensure long-term project impacts, through embedding breeding and research planning, and review in a users' perspective forcing all in the research-for-development process into an adoption-orientated focus.

Executive Summary

Breeding bananas is a long-term commitment, given its perennial nature and numerous issues relating to sterility. Over the course of five years, this project (BBB) has established a formidable network of researchers and breeding programs across 6 continents. This highly interactive network has resulted in opening the exchange and sharing of hybrids, expertise and information across institutes and breeding programs. Through MusaBase, a digital platform established by BBB to design field trials, conduct statistical analysis and archive all banana breeding data for open access, an institutional memory for banana breeding has been founded. In East Africa, the project has strengthened the *Matooke* banana breeding program in Uganda and underpinned an infant banana breeding program in Tanzania on *Mchare*. Indeed, it has stimulated the national program to breed bananas. Banana breeding efficiency has been significantly improved, as witnessed through the 3-fold increase in seed production, better understanding and knowledge of banana floristry and pollination, and significant improvements in embryo rescue and germination rates. This has resulted in an accelerated production of progeny and the evaluation process, amounting to crosses and consequent selections for evaluation way above the project forecast. Over 200 *Matooke* hybrids have been selected for field evaluation, with thousands more in the pipeline and 1,572 hybrid *Mchare* under early field evaluation, representing the first ever *Mchare* hybrids. This project has also set the stage for the development of molecular tools to speed selection in banana breeding. Numerous QTLs associated with resistance to pests and diseases have been identified, as well as other important traits, mapped from various genetic backgrounds and located on different chromosomes. Models for genomic prediction and the fruit filling QTL have been evaluated, which will be validated and deployed in *Matooke* breeding. A banana breeding and selection pipeline has been realized and instigated, with a tracking system designed and now fully functional that tracks the whole process and interfaced with MusaBase. Furthermore, a consumer evaluation system has been established that feeds back into the pipeline to ensure a robust mechanism to deliver new cultivars to farmers with preferred end-user traits, such as taste and cooking traits. This allows us to dispose of unsuitable progeny early in the development cycle.

More than double the number of *Matooke* hybrids (231) were selected from EET for advancement to PYT than initially projected. Of these 122 were planted in the first joint evaluation of *Matooke* hybrids independently developed by NARO and IITA. Seed set data from *Matooke* cultivars have also shown that seeds are mostly extracted from just a few cultivars. Currently, 541 *Matooke* hybrids from 3x-2x crosses and 13,598 *Matooke* hybrids from 4x-2x crosses are in EETs. Results also show that triploid hybrids next to tetraploids and diploids are generated from 3x-2x crosses, which requires further investigation, as this will help avoid the 4x-2x crosses cycle, shorten the breeding schedule and increase 'tookiness' if *Matooke*-derived diploids are used. Hybrid ploidy is now being routinely assessed at Sendusu, including for NARO hybrids. Much attention has focused on the development of diploids, using local crosses or crosses with introduced ITC and EMBRAPA material, with 8,126 diploid hybrids generated from inter-diploid crosses and being evaluated, and 22 improved EMBRAPA diploids and 62 diploids of *banksii* background (ITC) planted out. The diploids *M. acuminata* subsp. *malaccensis* accession 250 and cv Rose are the most ideal male parents to screen for female fertility. To date 1,572 *Mchare* hybrids are under evaluation, while 79 doubled chromosome *Mchare* plants were generated for 8 cultivars, of which 60 lines have been planted in the field. NRCB have also initiated chromosome doubling with AA, BB and AB diploids using oryzalin, with 6 doubled chromosome plants identified from a population of 711 plants. Using novel video techniques, the floral development of seed-fertile and seed-sterile cultivars were shown to differ and stigmas found to be more receptive before the flower opens, all of which enables more targeted pollination timings. Sugars also affect pollen germination; glucose is more efficient than fructose or sucrose, and apparently better than diluted banana nectar. Results from studies showed that DNA marker-aided selection, such as genomic prediction, would improve efficiency of selection of both parental material and progeny. High quality genomic DNA has also enabled reliable sequencing. A draft genome version of a representative of *Mchare* is under development, the first reference quality genome sequence for this type of banana, also enabling better selection of parents and progenies. The use of oligo painting FISH has revealed relatively high numbers of translocated chromosome regions in various *Musa* accessions, which is also helpful for selecting suitable breeding parents. Multiple-site genotyping across partners is enabling comparisons in genetic diversity between sites and techniques and which helps identify discrepancies and anomalies, towards improving quality control. The genotyping of 667 accessions using SSR markers, has confirmed the identity of breeding material and parents. It appears that *Matooke* originates from a single hybrid clone, with *M. acuminata* ssp. *zebrina* and ssp. *banksii* as the most probable parents. This information has also identified ssp. *zebrina* as the source of resistance to bacterial wilt and that ssp. *banksii* is important for provitamin A and possibly starchiness. The analysis of the primary genome assembly of *Mchare* failed to identify unambiguously the parental sub-genomes. Neither did the new genome sequence of *M. acuminata* Calcutta 4 lead to an unambiguous split of the initial genome assemblies of cv Huti White. Results using Bionano mapping for the analysis of genome rearrangements at the whole genome level gave promising results, however. The exchange of banana and plantain material between the breeding programs of India, Brazil and IITA has huge potential benefits for all, in particular for their reaction to Foc, which is especially valuable for Africa. EMBRAPA have sent improved diploids to IITA, 9 of which are considered resistant against TR4, further to genotyping at UQ. The identification of QTL markers for resistance against key pests and diseases has made significant progress, as well as for some important post-harvest traits. Preliminary assessments, including from Genome-Wide Association Study (GWAS), show that Chr3 is particularly important, including for TR4 resistance and probably weevil resistance,

as well as for fruit filling, an important trait for breeding. Hundreds of lines of various populations have been phenotyped and genotyped to generate SNP markers. A training population for genomic prediction was used to run a GWAS, especially for fruit-filling traits. Markers segregating in the 'Monyet' x 'Kokopo', 'Paliama' x 'Borneo', 'Calcutta 4' x 'Zebrina GF' and 'Kasaska' x 'Borneo' populations have amounted to thousands of SNPs and used in QTL mapping for the traits. For 'Paliama' x 'Borneo' population a linkage map of 2,778 SNPs was constructed, with the markers grouped in 11 linkage groups, which represents the densest banana linkage map, spanning all 11 banana chromosomes. 'Monyet' has also now been shown to be a tetraploid (4*x*), not diploid, resulting in a 4*x* X 2*x* cross providing triploid (3*x*) progeny. QTL markers for resistance against Fusarium TR4, STR4 and Race 1 are gradually being defined using a targeted approach with *Musa acuminata* ssp. *malaccensis*. A SNP haplotype of 10 SNPs was found to determine resistance and susceptibility to Foc STR4 for this QTL locus on Chr3. Twelve loci were significantly associated with rhizome discoloration, the most reliable measure for *Foc* resistance, with the putative QTL found on Chr11, 8 and 9, which also reappear for leaf symptom index. The 'Borneo' linkage map analysis shows 2 significant QTLs for rhizome discoloration on Chr6 and Chr7 and 1 QTL for leaf symptom index on LG9 (Chr9). Four traits were assessed for weevils, total cross-section damage being the most important. SNP markers indicate that two regions of the genome confer weevil resistance on Chr6, while two new significant markers on pseudo-chromosome 12 were mapped on Chr3 and Chr10, which fit within the 2 QTLs for total cross-section damage. A QTL on Chr5 appears to confer resistance to bacterial wilt, with 5 more markers significantly associated with resistance. Data for nematodes were unreliable and a more consistent and reliable method is necessary. The heterozygosity nature of banana should not be overlooked, however, which likely explains the low level of explained variance by the QTLs. Genotypic interpretation of phenotypic data has led towards establishing good prediction accuracy with genomic selection models. Although no single model was best for all the yield component traits, prediction accuracy was improved. Leaf archiving has continued with 6,364 genotypes so far completed. One study also showed that freeze-drying yielded slightly less DNA than fresh cigar leaf material but is still suitable for genotyping with DArTSeq®.

Knowledge on pest and disease populations and dynamics in the region has been improved and protocols for more rapid and efficient screening developed and implemented, and consistently finetuned. The accurate characterization of pathogens and pests has continued and used for evaluating hybrids and breeding materials for resistance. A total 220 isolates of Foc Race 1, 369 isolates of *Pseudocercospora fijiensis,* 12 populations of Weevil have now been collected, stored and used to assess genetic diversity among populations. Monoxenic populations of *Radopholus similis* have been established in Kenya, Tanzania and Uganda for routine use in screening purposes. From surveys, *P. musae* and *P. eumusae* appear to be absent, while *Pratylenchus goodeyi* and *P. coffeae* appear more prevalent than previously known and should be considered for resistance screening. Meanwhile, weevil incidence was found to be higher when combined with maize cultivation. Improved diagnostic and screening techniques have been tested or developed, such as assessing Foc DNA in banana tissue, using RNAseq to observe transcriptome changes in response to Foc Race 4 infection, assessing different aged plants for greenhouse screening, observing sporulation of SR4, etc. Phenolic compounds could not be recommended as a reliable indicator for Foc resistance using small plant screening, while inoculation of the nematode *P. goodeyi*, prior to *Fusarium* inoculation revealed more consistent and rapid Foc screening results. Several options for improving the efficiency of screening against nematodes have been assessed and/or under assessment; the use of cocopeat plugs shows promise for standardizing and improving the process. The NARITAs have been assessed for various pest and disease resistance both in the screenhouse and the field, with good resistance being demonstrated against Foc Race 1, weevils, Sigatoka and *R. similis*. Some of these hybrids can be released to farmers in the region as an alternative to susceptible landraces. Just NARITA 12 and 16 appear to be susceptible to Foc Race 1. Improved diploids and Pisang Awak hybrids have been planted for assessment against Foc Race 1 in the field in India. Screening against weevils is underway, including diploids shown to be resistant to *R. similis*. Screening against Foc Race I in India found 3 improved diploids developed by NRCB to be resistant, which are being sent to IITA. EAHB and *Mchare* hybrids were found resistant against Foc Race 1, while all *Mchare* and *Muraru* cultivars evaluated were susceptible. Field and farmer testing of NARITA hybrids across Tanzania and Uganda have created more robust sex-disaggregated mechanisms for assessment of new cultivars with farmers and consumers, which supports the varietal release process but has also provided important feedback to the breeding program, stimulating gender-oriented banana breeding initiatives. Focus group discussions conducted across the region demonstrated the high importance of banana. Preference ranking exercises with farmers were completed in all 5 sites, using both quantitative and qualitative assessment of traits that farmers look for (or avoid) when selecting banana cultivars. A literature review on gender and trait preferences for banana cultivation also illuminated information useful for the breeding pipeline. An additional literature review on meta-analysis of multi-environment crop cultivar trials data was used to help develop a model to aid cultivar selection. Irrespective of gender, farmers tend to prefer traditional cultivars because of consumption attributes, demonstrating that new improved cultivars need to reflect local taste preferences. Most important desired visual traits are large bunch size, fruits and hand size, moderate suckering, plant height and resemblance to *Matooke*. In general, traits preferred by women reflect those of men. Across regional field-testing sites, NARITA 23 was ranked the best NARITA by farmers, followed by NARITA 2 and NARITA 12. Some NARITAs outperformed and were preferred by farmers than the local cultivars in some locations but not all. Six NARITAs have now been selected for advancement to multi-locational farmer-led trials. Importantly, the exercise highlighted the need to ensure uniformity of age across genotypes when planting, to enable consistent and comparable assessment of cultivars. These studies also highlighted the importance of organoleptic and physico-chemical assessments, which has led to the development of an analytical tool for quality and sensory evaluation. Breeding initiatives seeking to improve food security by developing and introducing productive, pest- and disease-resistant hybrids benefit from rigorous characterization within the target region for release. An open source system, Banana breeding Tracking Tool (BTracT), which accurately captures and efficiently tracks data throughout the breeding process was developed and successfully implemented and is now fully operational. BTracT is fully interfaced with the banana breeding database, MusaBase, which provides a global repository for all breeding data as it is collected in real time. FieldBook data collection, linked to BTracT, accurately records data at all stages of the breeding program, synchronizing it automatically with MusaBase to archive all information and data and make it publicly available. The wish list import-export function between MusaBase and BTracT was developed in Year 5 and together, these electronic tools and databases have become continuously more integrated during the project through upgrades, improving the synchrony and resolving issues and glitches. BTracT is now being transferred to the plantain breeding program in Nigeria and EMBRAPA, Brazil.

Over the course of the project, annual planning meetings have been held, as well as regular Steering Committee meetings. A Scientific Advisory Group was formed, who interact with project members and attend the annual meeting. The management team has developed and maintained the project website, organized or facilitated various training events and raised the profile and visibility of the project activities and achievements through various media events. A wealth of publications, presentations and outputs have been produced and disseminated, while significant capacity building through post-graduate student involvement has been achieved. An IITA internal audit and a BPAT review led to improvements but also highlighted the successful implementation and coordination of the project. A draft product replacement strategy and improvement plan

(linked to BPAT) was developed to fit in the scope of EiB. Interaction and links with external banana projects has gradually broadened the project reach, which will extend into the new Phase II (ABBB), which was approved in September 2019.

3. Organization of the strategic research goals:
   Part A. The first 5 strategic goals are organized around 5 work packages (WP), while strategic goal 6 is implemented via a project leader, a steering committee, and a science advisory group (SAG).

   1. Banana Breeding Pipeline: led by NARO, in collaboration with IITA, KULeuven, EMBRAPA, and NRCB.
   2. Pest and Disease Control: led by Univ. of Stellenbosch (South Africa), in collaboration with Univ. of Queensland (Australia), NARO, IITA, TARI (Tanzania), and Bioversity International.
   3. Leveraging Genetics: led by IITA, NARO, SLU-Sweden, in collaboration with Univ. of Stellenbosch, Univ. of Malaya, Univ. of Queensland, EMBRAPA, IEB, UNCC, DHMRI, and WCMC.
   4. Empowering End-user Evaluation: led by Bioversity International, in collaboration with TARI-Tengeru, NARO, and IITA.
   5. Harnessing Data: led by BTI, Cornell Univ., in collaboration with IITA, NARO, NRCB, and EMBRAPA.

   There were no changes in members in this project in year 5.

   Part B. Status of Contracts and consequences:
   1. In the Results Framework and Results Tracker we provide information on the progress of each contract and variation of each output, and presented based on primary outcomes 1–9.
      a. primary outcomes 1: WP on Banana Breeding Pipeline
      b. primary outcomes 2: WP on Pest and Disease Control
      c. primary outcomes 3: WP on Leveraging Genetics
      d. primary outcomes 4–7: WP on Empowering End-user Evaluation
      e. primary outcomes 8: WP on Harnessing Data
      f. primary outcomes 9: WP on Governance and Management

      This will be discussed in detail, with additional details provided in via links to additional documents.

4. Organization of the Governance, Research Oversight, and Management: http://breedingbetterbananas.org/index.php/staff/
   1. Management: the Project Management team consists of Project Coordinator and technical lead: Rony Swennen; Project Manager: Danny Coyne; Project Administrator: Scola Ponera; who work closely as a team. The original Project Manager was replaced in Year 2 when the Project Administrator was also recruited. Significant activities during Year 5 have included organizing the Annual Project Planning Meeting, maintaining the project website, coordinating SC/Management meetings and SAG interaction with Project members and preparing a proposal for Phase II.
   2. Project Planning Meetings were held successfully annually, including for Year 5 in a key banana producing area in Mbarara, Uganda 27-29 May 2019. This meeting was opened by Dr Ambrose Agona, DG NARO who welcomed the opportunity to attend and hear of the positive progress the project is making. Dr Geoffrey Mkamilo, DG TARI was unable to attend but is pleased to see greater regional activity on banana, while TARI itself has now set sights to develop a banana breeding program, based on BBB. This meeting acted to provide reflection on the project progress as well as areas for improvement, by project partners and invited additional partners linked and related to the banana value chain, especially other BMGF supported projects, such as RAPID, BBTV project, RTBFood. Six project highlights/achievements were also emphasized and used as a basis for meeting presentations (https://breedingbetterbananas.org/index.php/scientific-publications/). The annual meeting also served to develop a plan and way forward for a second phase of BBB.
   3. A proposal for a follow-on phase of the current project was invited by BMGF, which was submitted timely (OPP1213871), and subsequently approved in September 2019 enabling an uninterrupted flow from Phase I to Phase II.
   4. Steering Committee (SC): consisting of nine members including the secretary and a BMGF non-voting member met at the Annual Meeting in both separate and combined meetings with the Science Advisory Group (SAG), chaired by Klaus Koehler in the absence of Steve Rounsley.
   5. SAG: Jane Gibbs stood down as SAG member during the final year of the project, while Hale Ann Tufan was unable to attend the Annual Meetings due to circumstances outside of her control.
   6. SAG members attended the Annual Meeting in person with apologies from Steve Rounsley and Hale Ann Tufan. The SAG has been in contact with WP leaders and staff on an *ad hoc* basis to review project outputs against the existing Results Framework and Results Tracker. Greater interaction has been stimulated between project members and the SAG as the project has progressed.
   7. The IITA banana breeding program has received positive evaluations from both the BPAT (Breeding Performance Assessment Team) and by multiple module leaders of the Excellence in Breeding (EiB) program. Seventeen recommendations for improvement were made by the BPAT team, which have been thoroughly addressed through an improvement plan submitted last year to EiB. We await their evaluation of the plan. In addition, a scheme for reorganizing the breeding pipeline to focus on improving genetic gain has been submitted to EiB and has received their positive support.
   8. An Internal Audit of the project in 2017 by IITA highlighted a positive project implementation level; good interactions and common understanding of the mission among all stakeholders; good monitoring of financial transactions within IITA through periodic postings to IITA budget holders; good project investment in lab facilities; great synergy, team work and coordination among the project staff and implementing partners; development of the project website and banana database, MusaBase (https://musabase.org/). A small number of risk issues that were identified were addressed towards improved implementation.

5. Capacity building and communications:
   1. Overall, a broad range of outputs have been produced (Annex 1), or remain in preparation, including book or journal articles, technical briefs, conference oral and poster presentations at international meetings as well as at the BBB Annual Meetings (https://breedingbetterbananas.org/index.php/scientific-publications/) and communicated through various channels. Some outputs have been part-funded by other projects, activities or result from BBB activities. Numerous media outputs have been developed, largely as a result of exposure at the BBB Annual Meetings, but also during the course of the project and released through various media outlets as feature length media stories, TV interviews, public media journals, short documentaries on breeding activities etc.
   2. To create awareness and visibility of the project and of the importance of banana in the Great Lakes region and why improving banana breeding is key to ensure food security of the local communities, two videos were produced and published to reach a wider audience and

the general public: “Journey to a better *Mchare*: Improving Tanzania cooking banana” (https://youtu.be/BzhNKTLN73c) posted September 2018 has received 339 views and created 2,700 impressions; ”Breeding Better Bananas - Improved *Matooke* for East and Central Africa” (https://youtu.be/yXbeDBg9SuI) posted December 2019, has had 354 views and achieved 1,000 impressions. Both videos are also posted on the BBB website.

3. A range of training events have been undertaken during the project, including regular trainings in relation to MusaBase and data entry, breeding training, scientific writing in addition to exchange visits between partners and breeding programs to assess technical aspects as well as foster links and relations. Specific workshops for on-farm evaluation of NARITA, genomics, science writing, breeding techniques and seed set have provided training opportunities and built capacity in partner institutions, outside of the project and for associated post-graduate students.
4. Website: The project website (http://breedingbetterbananas.org/) was developed in Year 1 and has been maintained and regularly updated during the project, providing a resource platform both for the general public as well for project members to restricted areas. To provide stability the website domain is currently registered until 30/6/2024. Statistics show that as of 31 December 2019, the site has had 4,387 users and 13,214 accumulated views. The audience is gender-balanced: females (52.7%) and male (47.3%), of mostly 25-34 aged users, mainly from Kenya, Tanzania, Uganda, Nigeria, Italy, France, Belgium, UK, USA, Mexico, India and the Philippines. Overall, the website has a very positive sentiment score. The Twitter Account @BBetterBanana, created in May 2018, has had 106 posts, has reached 210,687 users and has achieved a total of 649,671 impressions. The website links directly with MusaBase (https://musabase.org/) and other relevant domains.
5. Capacity building has been a key cornerstone of the project towards generating banana breeders and technical staff for the national programs. The project aimed to maximize student involvement both directly supported as well as using the project as a platform to support those with alternative funding, especially female students. In total 25 postgraduate and 1 graduate student have been associated with the project (Table 1-2). Three PhD students graduated during Year 5 http://breedingbetterbananas.org/index.php/scientific-publications/theses/, one of whom, Ivan Arinaitwe, has been appointed to lead the NARO breeding program, while Michael Batte is now overseeing breeding activities at Sendusu.

   Table 1. Number of students supported by and associated with the project

	Female	Male	Total	Supported by the project	Support from outside the project
PhD	3	13	16	11	5
MSc/BSc	3	7	10	4	6

   Table 2. Students supported by and associated with the project according to work package

	Work package					
	WP	**WP2**	**WP**	**WP**	**WP**	**N/A**
PhD	2	4	3	2	0	5
MSc/BSc	6	1	2	1	0	0

6. Further to the SC and SAG meetings at the Annual Meeting, management skype meetings were held on a two-monthly basis for regular contact between WP leaders and SC, an attribute highlighted by the internal audit towards interaction and team spirit.

6. Investment outputs and outcomes: the major achievements are:

Primary outcomes 1: WP 1 on Banana Breeding Pipeline

Scope and Approach: Enhanced performance of breeding systems to deliver improved East African highland bananas (EAHB) with increased levels of pest and disease resistance, higher yields and better consumer acceptability.

The goal of this work package is to increase breeding efficiency.

Therefore, the objectives are (reworded from the submitted project document):

1. Upscale the breeding activity;
2. Increase efficiency by obtaining higher seed set and higher numbers of hybrids;
3. Development of an international breeding platform.

Objective 1: upscaling breeding

A Breeding Pipeline structure was developed, which now combines and links all breeding activities. It demonstrates the different ploidy crosses, early/preliminary/advanced evaluation, pest and disease evaluation, sensory evaluation and varietal release. This scheme is necessary to help guide breeding activities, which have rapidly expanded and upscaled. *Matooke* and *Mchare* product profiles guide the development, evaluation and selection processes for the NARITAs and new hybrids. To fit in the scope of Excellence in Breeding platform (EiB), a draft product replacement strategy and improvement plan (linked to BPAT recommendations) for *Matooke* and *Mchare* was developed and submitted to EiB.

Major progress has been realized with the production of >230,000 seeds and >160,000 embryos cultured, involving different ploidy combinations with *Matooke*; the first *Mchare* hybrids were also produced and field planted. In-depth field morphological characterization of *Matooke* and NARITAs was performed. BBB has demonstrated strong progress and project ability to reach its targets for *Matooke* and *Mchare* breeding (Table 3). Sizable fields of *Matooke* landraces and derived tetraploids are now in place: 1) 15 *Matooke* cultivars in 6.0 ha pollination blocks at IITA-Sendusu and NARO-Kawanda, including Bitambi, Entukura, Enzirabahima, Nakyetengu, Kabucuragye, Kazirakwe, Nakabururu, Nakasabira, Nakawere, Nakayonga, Namande, Namwezi, Nfuuka, Rwambarara and Tereza; 2) 10 *Matooke* tetraploids in 4.0 ha pollination blocks at NARO-Kawanda and IITA Sendusu, including 917K-2, 660K-1, 222K-1, 401K-1, 376K-1, 1201K-1, 1411K-1, 199K-3, 1438K-1 & 365K-1. Seed set data from the 15 *Matooke* cultivars show that seeds are mostly extracted from Nakyetengu, Nakawere, Nakasabira, Enzirabahima, Nakabururu.

We have reached our ambitious targets in delivering the expected quantity of material (95 promising *Matooke* hybrids by Year 5) with 231 *Matooke* hybrids selected from EET for advancement to PYT. Of these 122 (92+30) were planted for evaluation in multilocation PYTs (Table

3), representing the first joint evaluation of *Matooke* hybrids independently developed by NARO (37 hybrids) and IITA (55), which addresses one of the BPAT and SAG recommendations. This combined evaluation will increase the overall efficiency of the program, while allowing each team to specialize in one or more aspects of hybrid production. Arusha is chosen for diploid improvement, Sendusu will focus on tetraploid (3x-2x crosses) development, and Kawanda on the breeding of secondary triploids (4x-2x crosses). The other 130 *Matooke* hybrids are being multiplied, ready for the second wave of testing, while additional selections in EET are ongoing.

Table 3. Overview of breeding progress

Parameter	Target	Overall numbers
Improved diploids integrated into the *Matooke* and *Mchare* breeding pipeline	70	22 parthenocarpic EMBRAPA diploids resulting from 2x-2x crosses, with a bunch weight of over 5 kg, high pollen quantity, and resistance to black Sigatoka have been selected for further evaluation and 6 selected with nematode (*R. similis*) resistance. 8,126 diploids from 2x-2x were generated. 62 diploids of *banksii* background sourced from ITC under field characterization.
Matooke hybrid under evaluation in the EET (early evaluation trial)	12,000	13,598 hybrids from 4x-2x crosses 541 hybrids from 3x-2x crosses
Matooke hybrids (beyond NARITA 1-26) tested in PYT (preliminary yield trial)	95	231 *Matooke* hybrids selected from EET for advancement to PYT. 92 of these hybrids were planted in Mbarara, Hoima and Sendusu and 30 planted in Kawanda.
Development of *Mchare* hybrids	2400 seeds for embryo rescue	1,572 *Mchare* hybrids are under evaluation in EETs in Tanzania and Uganda.

Selected *Matooke* hybrids (NARITAs and their progenitors) are now routinely delivered to the Food Biosciences laboratory (Dr. K. Nowakunda, new collaborator), NARO-Kawanda for sensory and physico-chemical evaluation. A new texture analysis equipment has been installed and a texture analysis protocol (with the new equipment) developed. Also, a product oriented sensory panel was trained for the assessment of organoleptic properties, and starch and amylose content were measured for the first time (Table 4). Total mean starch content for the NARITAs was generally over 80%, with a just a few, such as NARITA 12, below 80%. Amylose content was between 11.0 and 14.6%. This now enables a much faster evaluation of *Matooke* hybrids with acceptable culinary qualities (Table 5). This collaboration with "Breeding RTB products for end user preferences (RTBfood)/ OPP1178942" reinforces cross-project interaction with other BMGF funded projects.

Table 4. Starch content of some NARITAs (N=5)

Genotypes	Total Starch (% DB)	Amylose (%)
Mbwazirume	88.1	12.23
NARITA 15	91.4	12.11
NARITA 16	89.1	13.21
NARITA 17	91.5	12.8
NARITA 18	77.3	11.3
NARITA 19	94.1	12.00
NARITA 20	85.8	12.9
Kabucuragye	84.2	12.15
KAZIRAKWE	85.1	11.01
NARITA 23	78.9	11.2
NARITA 4	82.1	12.08
NARITA 7	91.1	11.03
NARITA 8	89.2	12.8
NARITA 12	88.1	ND
NARITA 12	75.9	14.5
NARITA 24	ND	ND
NARITA 13	84.8	12.8
NARITA 14	80.07	14.6
NARITA 21	85.9	13.7
NARITA 5	86.2	12.9
NARITA 1	79.9	14.05
NARITA 25	88.1	14.3
NARITA 6	ND	ND
NARITA 7	80.1	ND
NARITA 28	81.0	13.9
NARITA 9	79.7	ND
NARITA 10	83.1	12.9
NARITA 2	81.1	11.9
NARITA 3	79.9	ND

ND = Not yet determined

Table 5. Instrumental texture of some raw and steamed NARITAs (cooking time =90 minutes; N= 5)

Genotypes	Raw pulp (N)	Steamed (N)	Overall acceptability
Mbwazirume	24.10	0.59	4.9
NARITA 11	25.7	4.48	2.1
NARITA 12	21.95	0.98	3.0
NARITA 24	25.4	0.52	4.0
NARITA 13	21.2	0.79	1.5
NARITA 14	25.10	1.68	3.4
Kabucuragye	23.65	0.57	4.7
NARITA 15	25.75	1.72	2.3
NARITA 16	25.98	1.31	2.1
NARITA 17	25.14	0.58	4.6
NARITA 18	24.2	0.97	3.7
NARITA 19	29.6	4.25	2.7
NARITA 20	35.05	2.91	2.0
NARITA 21	33.79	5.48	1.9
NARITA 5	24.09	0.42	2.7
NARITA 1	26.12	0.99	2.0
NARITA 25	27.01	1.09	2.1
NARITA 6	26.4	1.17	2.6
NARITA 7	23.9	1.01	4.3
NARITA 28	26.1	0.57	3.5
NARITA 9	26.0	4.0	2.1
NARITA 10	26.8	3.9	1.5
NARITA 2	25.8	4.11	3.3
NARITA 3	25.0	4.5	3.1
NARITA 23	27.0	1.0	3.5
NARITA 4	26.1	1.3	3.9

Scores based on a hedonic scale ranging from extreme approval (5) to extreme disapproval (1). N = newtons (unit measure of force)

Much attention has also focused on the development of diploids, using local crosses or crosses with introduced ITC and EMBRAPA material. 22 improved EMBRAPA diploids (from 2x-2x crosses) are in both Sendusu and Arusha. Their ploidy was confirmed as diploid. Their characteristics include bunch weights over 5 kg, high pollen quantity and resistance to black Sigatoka. In Sendusu 49% of cycle 1 flowering

data were collected, 35% of cycle 1 harvest data were collected, 24% of cycle 2 flowering data were collected and 11% of cycle 2 harvest data collected.

All traits are very different (*P* <0.001) among the genotypes, indicating a high level of variation among the EMBRAPA diploids, and consequently a good basis for the breeding program. Six genotypes (BMPG-1, BMPG-10, BMPG-11, BMPG-18, BMPG-2 and BMPG-8) were resistant to nematodes (Table 6) and are recommended for use in breeding for nematode resistance. Suckers were collected by NARO for tissue culture. The resulting plantlets will be used in screenhouse and field evaluation trials to assess for multiple resistance against weevils and Fusarium wilt Race 1 using the most prevalent VCG in Uganda.

Table 6. Host response of EMBRAPA diploids to *R. similis,* compared with resistant (Yangambi Km5) and susceptible (Valery) checks

Genotype	Mean	SE	comparison with Valery	comparison with Km5	Host Response
BMPG-1	2.52	1.76	s	ns	Resistant
BMPG-10	2.25	1.51	s	ns	Resistant
BMPG-11	2.86	0.88	s	ns	Resistant
BMPG-12	6.97	0.27	ns	s	Susceptible
BMPG-13	6.93	0.19	ns	s	Susceptible
BMPG-14	7.07	0.32	ns	s	Susceptible
BMPG-17	6.51	0.23	ns	s	Susceptible
BMPG-18	2.51	1.07	s	ns	Resistant
BMPG-2	1.53	0.77	s	ns	Resistant
BMPG-21	5.82	0.24	ns	s	Susceptible
BMPG-22	5.73	0.36	ns	s	Susceptible
BMPG-4	6.13	0.19	ns	s	Susceptible
BMPG-5	5.44	0.22	s	s	inconclusive
BMPG-6	6.89	0.13	ns	s	Susceptible
BMPG-8	0.00	(omitted)	s	ns	Resistant
Valery	7.83	0.11		s	
Km5	0.66	0.40	s		

Also, 62 diploids of *banksii* background, sourced from ITC, are being characterized for pollen fertility, pest and disease response and agronomic performance at Sendusu. A further 8,126 diploid hybrids were generated from inter-diploid crosses and are being evaluated for bunch yield, pollen quantity and quality. Confirmation of their ploidy is underway. Currently, 541 *Matooke* hybrids from 3x-2x crosses and 13,598 *Matooke* hybrids from 4x-2x crosses are in EETs at IITA-Sendusu. Hybrid ploidy is now being routinely assessed at Sendusu, including for NARO hybrids, further to training of a NARO technician by IITA for preparation of samples and running the ploidy analyzer machine.

Contrary to the common belief that 3x-2x creates more tetraploids than triploids and diploids, results show that triploid hybrids next to tetraploids and diploids were generated from 3x-2x crosses. This requires further investigation, as this could help avoid the 4x-2x crosses cycle, and lead to shortening the overall breeding schedule tremendously. If a *Matooke*-derived diploid is then used in addition, more “tookeness” would most likely be obtained in the final product.

1,572 *Mchare* hybrids are currently under evaluation in Tanzania and Uganda, while 79 doubled chromosome *Mchare* plants were generated for 8 cultivars: Mlelembo=10; Hutishamba=11; Mshale=10; Nshonowa=12; Huti RB=7; Huti Shumba nyeelu=3; Kahuti=18 and Makyugu=8. Of all these, 60 lines have been planted in the field. More plants have been treated *in vitro* to select chromosome doubled plants. They are recovering and once regenerated will be tested for ploidy level. NRCB have also initiated chromosome doubling with AA, BB and AB diploids using oryzalin. So far 6 doubled chromosome plants were identified from a population of 711 plants in India (Fig. 1).

Figure 1. Oryzalin treated AB diploid and control at NRCB

Objective 2: Increase efficiency by obtaining higher seed set and higher numbers of hybrids

Floral development stages of cvs Enzirabahima and Nakitembe were characterized using video footage, a method developed in the screenhouse and successfully transferred to the field. Results revealed that bracts of cv Enzirabahima - a seed-fertile *Matooke* open faster than those of cv Nakitembe - a seed-sterile *Matooke*. Flowers of cv Nakitembe opened mostly in the evenings, and early in the morning for cv Enzirabahima. Floral development studies, especially on the patterns of seed set for Enzirabahima and Nakitembe, have shown that the stigmas are more receptive before the flower opens and that seed set occurs mostly in the lower hands. Similar results were obtained for *Mchare*, using the seed-fertile cv Mshale - and seed-sterile cv Mlelembo. The video footages were uploaded on MusaBase and all data part of a book chapter on banana flowering (accepted).

Assessment of pollen germination following sugar application showed that, depending on the sugar (combination) and concentration, pollen germination could vary enormously, i.e. between 2 - 48%. Glucose was more efficient than fructose or sucrose; combining glucose with fructose and sucrose alone or in combination had either no effect or even reduced pollen germination. A concentration of 3% glucose appears optimal,

following tests using 1-20% sugar concentrations. Using diluted banana nectar, however, could not provide the high pollen germination rates that 3% glucose gave.

Historical data of breeding *Matooke* over the past 21 years were analyzed and published. Unlike plantain breeding, it appears that seed set is not season dependent and that pollinations can be undertaken year-round, provided flowers are available. The diploids *M. acuminata* subsp. *malaccensis* accession 250 and cv Rose are the most ideal male parents to screen for female fertility.

Considering the low female fertility rates of *Matooke* (mean no. seeds/bunch between 0 and 1.5) and their derived tetraploids (between 0.2 and 39.2 seeds) and the occasional high seed set (305 seeds for Nakasabira, 2279 seeds for 917K-2), it is necessary to identify factors influencing fertility, towards improving pollination conditions and boost seed set. This continues to be researched. Similarly, the wide discrepancy between seeds harvested and embryo germination (e.g. 3 embryos from 44 seeds), demands greater understanding of dormancy, seed germination, and optimization of embryo culture protocols. This is currently being addressed with a MSc study (submitted). Analysis showed that NARITAs were obtained from 18.8% of the executed 4x-2x crosses and from 26.8% of the seeds that resulted from those crosses. A total of 17.7% of the executed 3x-2x crosses and 49.2% of the ensuing seeds contributed to the production of 4x female parents. While this highlights our need to better understand floral biology and seed germination, the data does show that DNA marker-aided selection, such as genomic prediction of breeding values (see further in this report), would improve efficiency of selection, as just 26.8% and 49.2% hybrids were selected in the last two steps of EAHB breeding. The seed-fertile EAHB cvs Nakabululu and Nakawere, which gave the highest levels of pollination success, are not currently in use but are now being reintroduced into the EAHB breeding program further to this work.

Objective 3: Build an international breeding platform

Progress in banana breeding requires collective and collegiate involvement with the best partners for strong, healthy and reliable collaborations across partners. Collaborative efforts have been further strengthened through activities not initially planned. For example, NRCB has now received all IITA plantain hybrids (PITA), most NARITAs and 7 IITA improved diploids for field testing, while EMBRAPA has established several fields in Brazil with PITAs from IITA. IITA has also received improved diploids from EMBRAPA and in 2019 signed a MTA with NRCB to receive 3 improved diploids with resistance to Foc Race 1. Both NRCB and EMBRAPA intend to use the IITA material in their breeding programs but very importantly, the feedback on the response of the PITAs and NARITAs to Foc Race 4 (NRCB) and Race 1 (EMBRAPA), is of tremendous value for the breeding efforts in Africa. Field exchange visits between the IITA and EMBRAPA breeders, and between the IITA/NARO and NCRB breeders has helped provide greater understanding between sites, genotypes and fostered relations between the teams.

Additionally, EMBRAPA further developed improved diploids for their breeding program for use in Africa. Ten new improved diploids resistant to Foc Race 1 and resistant / moderately resistant to black Sigatoka, were selected. These diploids have a low plant height, over 100 fruits per bunch, pendulous bunch and male fertility, among other positive agronomic characteristics. They were shipped *in vitro* for multiplication and local validation. A new set of 23 improved diploids (from crossing improved diploids amongst themselves) proved resistant to Foc Race 1, with the highest levels of resistance observed in CNPMF1323 [(Malaccensis x Sinwobogi)] x [(Calcutta 4 x Heva)], CNPMF0612 [(M53 x Madu) x Madu)] x SH3263, and CNPMF0534 [(Calcutta 4 x Madang)] x [(Borneo x Guyod)]. Moreover, they are short (under 2.0 m tall) and bear over 100 fruits, with over 8.0 living leaves at flowering. DNA samples from the 23 new improved diploids (and all improved diploids sent so far by EMBRAPA to IITA) were additionally genotyped by UQ to identify genotypes potentially resistant to Foc TR4. Three improved diploids, now at IITA, are considered resistant to TR4: BMPG 043, BMPG 047 and BMPG 068. Six of the newly improved diploids are also considered resistant to TR4: CNPMF 0557, CNPMF 0496, CNPMF 0536, CNPMF 0542, CNPMF 0612 and CNPMF 0731. DNA samples from the 23 new improved diploids (and all improved diploids sent by EMBRAPA to IITA) were also sent to the IEB for genotyping using simple sequence repeat (SSR) markers, molecular analysis of the ITS1-5.8S-ITS2 region of ribosomal DNA locus and the analysis of chromosomal distribution of ribosomal DNA sequences, to compare the results in genetic diversity between EMBRAPA and IEB Laboratories.

Despite the large phenotypic variation, flow cytometric estimation of 2C nuclear DNA content revealed small differences (max ~6.5 %) in genome size among EAHB clones at IEB. While no difference in the number of 45S and 5S rDNA loci was found among *Matooke* accessions, genotyping using 19 SSR markers resulted in grouping into four clusters. DNA sequence analysis indicates that *Matooke* originated from a single hybrid clone, with *M. acuminata* ssp. *zebrina* and ssp. *banksii* as the most and probable parents. However, *M. schizocarpa* also seems to have contributed to the formation of this group. These results will greatly help to focus and further finetune the breeding strategy. For example, IITA proved that ssp. *zebrina* provides resistance to bacterial wilt, *Xanthomonas campestris* (now renamed *X. vasicola*), a serious disease of banana in Eastern Africa, and also showed the importance of ssp. *banksii* for provitamin A, speculating that this ssp. contributes to starchiness.

Breeding was strengthened at NRCB: various crosses were made from 2x-2x, from 4x-2x and from 3x-2x, with now 42 Pisang awak hybrids under field evaluation and 12 genotypes under *in vitro* conditions. Also 4 Pisang awak tretraploid x diploid have now been planted in the field. Three hybrids of 3x x 2x crosses and 5 hybrids of 2x x 2x crosses were field tolerant to Foc Race 1.

IEB has provided substantial support in checking genetic analysis quality and has collated the data from IITA, EMBRAPA, UQ and the ITC. SSR genotyping was performed using a standardized platform by 19 microsatellite (SSR) markers. The SSR profile for each individual sample was checked by capillary electrophoresis on ABI3730xl DNA analyzer and data obtained were analyzed using GeneMarker v1.75 software (Softgenetics) followed by a manual check. Data obtained were merged with the “core subset” dataset and analyzed together. The results of SSR genotyping were reported to the partners as raw scores data tables and cladograms created using Darwin software. The results have helped confirm the identity of the material and to estimate the genetic diversity within banana cultivars and banana hybrid clones.

The Banana Breeding Tracking System (BTracT) is now fully operational in Arusha, started at Sendusu and is in process of activation at Kawanda, following training of the banana breeding teams. The system tracks all stages in breeding beginning with the automatic decision making of crosses in the field based on flowering plants, followed by all stages of hybrid production (from seed to embryo culture to nursery). Plants are barcoded and the data management system implemented and integrated directly with MusaBase.

Over 600 crosses, 300 bunches, 2600 seeds and 1500 embryos have now been recorded at Sendusu using BTracT. The team is ready to implement the tool in the nursery activities as soon as the required equipment is delivered, completing all breeding activities. The pollination and seed extraction teams at NARO have already undergone training in BTracT and await the required equipment. BTracT is also being

transferred for use in the plantain breeding program in Nigeria. Preparations are being made to transfer BTracT to EMBRAPA, Brazil, while NRCB has independently developed its own system.

Morphological characteristics of cultivars/genotypes, as grown under field conditions at IITA banana and plantain collections in Arusha, Tanzania, Sendusu, Uganda and Ibadan, Nigeria were shared with partners by uploading onto MusaBase 2,690 images, capturing the various different plant parts and growth stages of 185 cultivars.

The images will enable a multitude of uses, including:

1) As a reference for checking the identity of plants in the field;
2) To identify differences due to environmental influences;
3) To relate desirable traits to the morphology of the plants and fruits for breeding;
4) To provide this information to the wider community as an important resource for Big Data and image analysis.

Characterization data of 26 NARITAs and 52 accessions from the germplasm collection plot at Sendusu were collected. Cumulatively, numerical and photographic data have been collected for 89 accessions, including 11 *Matooke* lines used for pollination. All files are currently in the process of being arranged for uploading onto MusaBase to increase accessibility by other banana breeding programs.

Primary outcomes 2: WP 2 on Pest and Disease Control

Scope and Approach: Enhanced host plant resistance to major pest and disease constraints through improved pathogen identification and accelerated early stage screening of resistance.

The goal of this work package is to provide the tools to faster screen parents and hybrids towards improved efficiency in banana breeding. Pests and diseases in order of importance are: Fusarium, nematodes and weevils, and Sigatoka.

Therefore, the objectives are (now reworded from the submitted project document):

1. Establish collections of the 4 pests and diseases: in support of future screening of germplasm and development of molecular tools for diagnostics;
2. Determine the presence of the 4 pests and diseases in the testing and breeding sites: in support of field screening and genotype by environment effects and regional testing;
3. Develop diagnostic techniques for pests and diseases: for precise identification of the pest and pathogen;
4. Bioassay: for faster screening, more rapid throughput using younger plants;
5. Screen *Matooke* and *Mchare* (plus hybrids): to identify sources of resistance.

The stage of progress depends on the pests and diseases.

During Year 5 hybrids and banana breeding materials have continued to be evaluated in the field against banana pathogens and pests. To ensure accurate characterization for regional/genetic variants, pathogens and pests collected from testing and breeding sites in Uganda and Tanzania were sent for identification and characterization to SU, IITA, NARO and TARI-Tengeru. A Field Guide for Disease Diagnostics and Data Collection of banana pests and diseases was also produced earlier under BBB to provide standardized protocols for project members (and the global community) (https://breedingbetterbananas.org/index.php/all-reports/). Protocols (methods) for resistance screening of NARITA hybrids and breeding materials against target pathogens and pests including early/rapid screening methods for Fusarium wilt, *Pseudocercospora* spp., weevils and nematodes have been developed over the course of the project and finetuned during Year 5.

Objective 1. Collections of the four pests and diseases established.

Fusarium: 220 isolates of Foc Race 1 were collected in Uganda and Tanzania and added to the Foc collection at SU.

Weevil: the genetic diversity among 12 banana weevil populations from selected agro-ecologies identified 6 highly polymorphic SSR markers, which could be used to successfully genotype banana weevil populations taken from four Uganda banana growing AEZs.

Nematodes: attempts to establish monoxenic populations of nematode species were successful for *Radopholus similis*, which is now being used for screening, while attempts to establish *Pratylenchus goodeyi* and *P. coffeae* continue. Pure monoxenic populations of banana nematodes are being re-established (after being lost) by Dr Luambano, TARI-Kibaha, DAR. In Uganda a pure monoxenic population of *R. similis* is used for screening purposes at Sendusu. In Kenya, Tanzania and Uganda, monoxenic cultures are established and being used for developing more efficient nematode screening techniques, as this needs to be improved. The first monoxenic culture of *Helicotylenchus*, using carrot discs has been achieved, but took a long time (9 months) to mature and may not be practical for screening purposes. It appears that culturing of *R. similis* on cocoyam is possible and may be more efficient, offering an alternative to carrots.

Sigatoka: Mating type (MAT) 1.1 and 1.2 specific primers were used to analyze the frequency of mating type idiomorphs within 318 *P. fijiensis* pure isolates from Uganda and Tanzania. Both MATs were present with 184 (57.9%) isolates amplifying MAT 1 region and 42.1% for MAT 2. The idiomorphs occurred at ~1:1 ratio, characteristic of a naturally reproducing population. This equal distribution of MAT idiomorphs facilitates random and frequent sexual reproduction that results in high genetic diversity. Exceptions to this distribution were recorded in Mbeya, Tanzania, where only MAT 1 was found, and in Arusha and Sendusu, where the MAT 2 idiomorph was under-represented. A total of 369 pure isolates from seven sites in Uganda and Tanzania has been established and confirmed as *P. fijiensis*. These have been transferred to long term storage on glycerol in Kawanda, Dar-es-Salaam and Arusha labs.

Objective 2. Testing and breeding sites screened for the presence of the 4 pests and diseases.

The distribution of the 4 pests and diseases in the 5 testing sites in Uganda and Tanzania, and beyond, have been established and distribution maps available on the BBB website (https://breedingbetterbananas.org/index.php/2017/01/24/work-package-2-pest-and-disease-control/).

Fusarium: Fusarium wilt symptoms were identified at TaCRI (Kilimanjaro, Tanzania) and Mbarara, Uganda, and confirmed as Foc Race 1

Weevil: weevils were characterized in all three Tanzania testing sites (Arusha, Kagera and Mbeya) with three TARI staff trained by NARO-Uganda. A field study in banana-based farming systems in Arusha and Kilimanjaro Regions, revealed higher weevil incidence in combination with maize cultivation (http://breedingbetterbananas.org/index.php/scientific-publications/theses/).

Nematodes: GPS coordinated data sets of nematode distribution collected from fields and mother trial sites in Kagera, Arusha and Mbeya, Tanzania have been established (https://breedingbetterbananas.org/index.php/2017/01/24/work-package-2-pest-and-disease-control/).

Sigatoka: Specific primers for *P. musae*, *P. eumusae* and *P. fijiensis* were used to determine species distribution. No *P. musae* or *P. eumusae* were detected in Uganda and Tanzania and possibly absent. A total of 31 samples (Uganda) and 147 samples (Tanzania) were collected. *P. fijiensis* was confirmed for 16 isolates from Uganda. From Tanzania, 110 isolates were recovered. Isolates are stored at IITA-DAR.

Objective 3. Diagnostic methodologies of pests and diseases

Fusarium: At SU, qPCR of Foc DNA was used to assess its use in screening for resistance at different plant ages. In 1-month-old plants the amount of Foc DNA in the rhizomes of the susceptible cv Huti Green Bell was not significantly different from that in resistant Williams, FHIA 01 and FHIA 17 plants. The amount of Foc DNA in 2- and 3-month old plants of the highly susceptible Gros Michel and Silk bananas, however, differed significantly from that in resistant FHIA 01, FHIA 17 and Williams at 3 weeks after inoculation. Foc DNA in the susceptible genotype *Mchare* Akondro Mainty differed from that in resistant cultivars only after 5 weeks, suggesting that Foc development is slower in some cultivars than in others. qPCR of Foc DNA can therefore be used for resistance screening but requires a minimum period of 5 weeks for sufficient colonization to enable accurate results (Fig. 2).

At UQ a RNAseq experiment was performed to observe transcriptome changes in response to Foc Race 4 infection. The novelty lies in the fact that the comparison is made between the transcriptomes of the progeny of the same 'Ma851' and 'Ma852' lines. This differs from all previous RNAseq studies that have primarily assessed general resistance responses at the cultivar level. By using sibling lines from the same population that segregate for a single Race 4 resistance locus, they ensure that the transcriptome changes due to genes segregating in the genetic background are minimized. This enable the pathways that regulate Race 4 resistance to be determined in a more specific manner. Intact RNA with an RNA integrity number of at least 7 were used to generate cDNA libraries for sequencing. Sequencing was performed using the Illumina HiSeq platform at Genewiz, with a minimum 30 million reads obtained per sample. These reads will be used to identify genes that are differentially expressed during the infection process between resistant and susceptible plants.

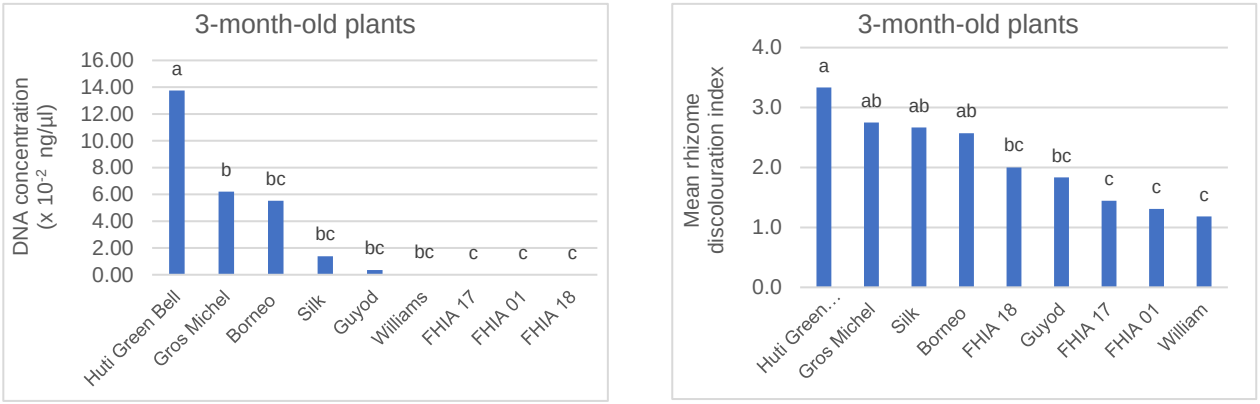

Figure 2. *Fusarium oxysporum* f. sp. *cubense* Race 1 DNA concentrations (left) and the discoloration index of rhizomes (right) of 3-month-old banana genotypes.

Nematodes: Nematode species were characterized based on morphological diagnostics and species distribution maps established.

Sigatoka: A set of 14 *P. fijiensis* SSR markers were obtained and optimized for assessing the genetic diversity. Preliminary analysis of isolates from Uganda revealed that these markers can confidently be used to elucidate genetic diversity within *P. fijiensis*.

Objective 4. Bioassay

Fusarium: The effect of plant age on the development of Fusarium wilt symptoms was investigated in the glasshouse. Disease severity did not significantly vary between resistant, tolerant and susceptible cultivars when using 1-month-old plantlets. Plantlets of 2 or 3 months clearly distinguished between resistant/tolerant and susceptible. There was no significant difference between 1-, 2- and 3-month-old plantlets in the susceptible cv Gros Michel. Very young plantlets, therefore, appear unsuitable for greenhouse screening. Two- and 3-month-old plants discriminated between susceptible and resistant bananas and are therefore recommended for screening.

Previous reports indicated that phenolic acids could act as an early indicator of banana resistance to Foc. Studies were promising but could not be confirmed, concluding that phenolic compounds cannot be recommended as a reliable indicator of Foc resistance using small plant screening. Results from a screenhouse study however, revealed that pre-inoculation of the nematode *P. goodeyi*, 14 days prior to *Fusarium* inoculation provides more consistent and faster results. The nematodes seemed to act as a predisposing factor for the fungal pathogen infection as they caused injury on the root surface as well as weakening the root tissues by causing root lesions (http://breedingbetterbananas.org/index.php/scientific-publications/theses/).

To study the infection process of Foc STR4, the presence of GFP-SR4 in 'Ma848' and 'Ma851' was visualized under a confocal microscope. At 14 dpi, 'Ma848' rhizomes showed high levels of sporulation near the xylem perforation plates. In contrast, spores were not observed nor intercellular movement of hyphae in the rhizome of the resistant 'Ma851'.

Nematodes: A number of potential options for improving the efficiency and speed of screening against nematodes have been tentatively assessed, towards extending those that show greater potential: 1) using fine sand and cotton wool infused with growth media in tubes, 2) assessment of penetration in pluronic gel, 3) coco peat plugs, 4) sterile paper pouches for TC plants or corm slices, 5) TC plantlet inoculation in liquid media. Challenges have been experienced with a number of these and modifications to overcome these continue, such as reducing contamination using pouches and cotton wool, and increasing the fluidity of liquid media. The coco peat plugs appear a potentially useful method, however, for providing uniformity and a standardized method for screening, but which does not necessarily reduce the timing, as the nematodes still need to multiply over 2-3 months. Coco peat, however, appears to interfere with nematode movement/multiplication and so

higher inoculum appears necessary, which is not an issue if good monoxenic cultures have been established. The relation between root penetration and resistance (multiplication over 2-3 months) is also being assessed to establish whether early penetration reflects potential resistance.

Weevils: the bioassay for screening weevil resistance consists of infected potted banana plants with male and female adult weevils as explained in the Technical Guide.

Sigatoka: A protocol for rapid evaluation of banana germplasm for resistance to black Sigatoka using 5 mg/ml lead to higher AUDPC (176.6) and faster disease development (19.8 days), compared to 50 mg/ml suspensions (AUDPC 90.9 and 34.4 days), and separated resistant from susceptible germplasm. Use of 5 mg/ml of mycelial fragment on detached leaves differentiated between a resistant variety (Calcutta), partially resistant (*Pisang lilin*) and a susceptible (Williams). *P. fijiensis* cell free culture filtrate was infiltrated on to a resistant (Calcutta), partially resistant (*Pisang lilin*) and susceptible (Williams) cultivars using a needle- less syringe. A clear hypersensitive reaction was observed on Calcutta 4 starting from day 4 post infiltration, while a water-soaked lesion at the point of administration was observed on *Pisang lilin* while no response was observed on Williams 10 days after infiltration. Therefore, cell-free culture filtrates was optimized for rapid differentiation of susceptible from resistant banana cultivars.

Effector proteins: Sequences of Avr4 (proteins present in *Cladosporium fulvum*) were obtained from the gene bank. A blast search for the protein homologs in *P. fijiensis* genome was undertaken to identify the gene sequences coding for these proteins and primers were developed. These were used to amplify the gene sequences for cloning and heterologous expression of the proteins in *Pichia pastoris*. The effector protein was sourced from Bon opus biosciences (USA). Testing and optimizing the use of the effector protein as a rapid screening tool is pending. These experiments have not been successful due to lack of equipment that gives controlled environmental conditions. Excessive heat might have affected stability of protein leading to inconsistent results. There is need to have a controlled environmental chamber.

A protocol for rapid production of *P. fijiensis* conidia for artificial inoculation was optimized and disseminated to Arusha and Sendusu.

Objective 5. *Matooke* and *Mchare* (plus hybrids) resistance screening

Fusarium:

Field data for Foc Race 1 symptoms on NARITA without artificial infection were collected monthly at 5 locations in Tanzania and Uganda over 2 years. Based on symptoms observed and diagnosis at SU, NARITA 12 and NARITA 16 developed Fusarium wilt at TaCRI (Kilimanjaro, Tanzania) and Mbarara, Uganda. Nineteen NARITA genotypes, including NARITA 12 and 16, were then inoculated in a pot trial (Fig. 3, Table 7) to confirm the results observed in the field.

Figure 3. Banana accessions in a pot trial for Foc Race 1 screening

Table 7. Screenhouse evaluation of NARITA hybrids for resistance to *Fusarium oxysporum* f. sp*. cubense* Race 1 at Kawanda, Uganda.

Genotype	Pedigrees for the female	Pedigrees for the male	RDI[1,2,3]	Grouping
Mbwazirume[4]	-	-	1.0 c	Resistant
NARITA 15	Enzirabahima x Calcutta 4	Tjau Lagada x Pisang Lilin	1.0 c	Resistant
NARITA 17	Entukura x Calcutta 4	Madang x Calcutta 4	1.0 c	Resistant
NARITA 21	Nakawere x Calcutta 4	SH3362 x Long Tavoy	1.0 c	Resistant
NARITA 22	Enzirabahima x Calcutta 4	Tjau Lagada x Pisang Lilin	1.0 c	Resistant
NARITA 25	Unkown	Unknown	1.0 c	Resistant
NARITA 3	Enzirabahima x Calcutta 4	SH3217 x SH3142	1.0 c	Resistant
NARITA 20	Unkown	Kabucuragye x Calcutta 4	1.0 c	Resistant
NARITA 13	Nakawere x Calcutta 4	SH3217 x SH3142	1.0 c	Resistant
NARITA 9	Enzirabahima x Calcutta 4	SH2095 x SH2766	1.0 c	Resistant
NARITA 26	Unkown	Unknown	1.0 c	Resistant
NARITA 14	Enzirabahima x Calcutta 4	SH3362 x Long Tavoy	1.0 c	Resistant
NARITA 10	Enzirabahima x Calcutta 4	SH2095 x SH2766)	1.0 c	Resistant
NARITA 19	Nakawere x Calcutta 4	SH3362 x Calcutta 4	1.0 c	Resistant
NARITA 8	Enzirabahima x Calcutta 4	SH2095 x SH2766	1.0 c	Resistant
NARITA 7	Nakawere x Calcutta 4	SH2095 x SH2766	1.1 c	Resistant
NARITA 11	Nakawere x Calcutta 4	Tjau Lagada x Pisang Lilin	1.2 c	Resistant
NARITA 4	Enzirabahima x Calcutta 4	Tjau Lagada x Pisang Lilin	1.3 c	Resistant
NARITA 16	Enzirabahima x Calcutta 4	SH3217 x SH3142	2.3 b	Intermediate
NARITA 12	Nakawere x Calcutta 4	Tjau Lagada x Pisang Lilin	2.7 b	Intermediate
Sukari Ndizi[4]	-	-	4.1 a	Susceptible

[1]The rhizome discoloration index (RDI) was scored using a scale of 1 to 6, where 1 = no internal symptoms and 6 = discoloration of the entire inner rhizome. [2]Means with the same letter within the same column do not differ significantly according to Fisher's test of least significant differences (P<0.05). [3]RDI in the screenhouse was rated 10 weeks after inoculation. [4]Mbwazirume and Sukari Ndizi served as resistant and susceptible controls respectively.

A group of 43 EAHB hybrids was evaluated in pots at Kawanda. Of the 43 EAHB hybrids, 28 were resistant, 10 intermediate and 5 susceptible to Foc Race 1. Eight *Mchare* cultivars evaluated both in the field and pots at Arusha were all susceptible. At Kawanda, however, disease severity on *Mchare* cultivars was considerably less than in Arusha, both in the field and screenhouse. The climate may have played a substantial role in disease development at the 2 locations. In addition, 8 *Muraru* cultivars and 27 *Mchare* diploids were inoculated with Foc Race 1 in the screenhouse at Arusha. All *Muraru* cultivars were susceptible, 15 *Mchare* hybrids were resistant, two intermediate and 10 susceptible (Table 8).

Table 8. Screen house evaluation of *Mchare* hybrids for resistance to *Fusarium oxysporum* f. sp*. cubense* race 1.

Number	Genotype	Parents	Mean RDI[1,2]	Grouping
	Controls			
1	Nakitengwa	Landrace (Resistant control)	1.0 l	Resistant
2	Sukari Ndizi	Landrace (Susceptible control)	6.0 a	Susceptible
	Muraru cultivars			
3	Muraru M3	Landrace	3.7 gh	Susceptible
4	Majimaji	Landrace	4.4 defg	Susceptible
5	Muraru Red	Landrace	4.7 cdefg	Susceptible
6	Mlalu	Landrace	4.8 bcdef	Susceptible
7	Muraru White	Landrace	5.0 abcde	Susceptible
8	Mraru Mchare	Landrace	5.0 abcde	Susceptible
9	TTZ 4	Landrace	5.7 ab	Susceptible
10	Njuru	Landrace	5.8 ab	Susceptible
	Mchare hybrids			
11	T.2327-1	Huti White x Cv Rose	1.0 l	Resistant
12	T.2274-6	Huti White x Calcutta 4	1.1 kl	Resistant
14	T.2273-2	Huti White x Calcutta 4	1.1 kl	Resistant
16	NM 185-1	Mchare laini x Borneo	1.2 kl	Resistant
17	T.2274-12	Huti White x Calcutta 4	1.2 kl	Resistant
18	T.1768-1	Huti White x Calcutta 4	1.3 kl	Resistant
21	T.2274-3	Huti White x Calcutta 4	1.7 jkl	Resistant
22	T.2274-7	Huti White x Calcutta 4	1.9 jkl	Resistant
23	NM 211-1	Kahuti x Calcutta 4	2.0 jkl	Resistant
24	T.2003-1	Nshonowa x Calcutta 4	2.0 jkl	Resistant
25	T.2269-1	Huti White x Calcutta 4	2.0 jkl	Resistant
26	T.2070-1	Huti White x Borneo	2.1 jk	Intermediate
27	T.2203-1	Nshonowa x Calcutta 4	2.4 ij	Intermediate
28	NM 154-1	Kahuti x Borneo	3.3 hi	Susceptible
29	T.2691-9	Huti White x Calcutta 4	3.8 fgh	Susceptible
30	NM 226-11	Mchare Laini x Calcutta 4	4.1 efgh	Susceptible
31	T.2274-8	Huti White x Calcutta 4	4.3 defgh	Susceptible
32	T.2731-2	Mchare Laini x Borneo	4.4 defg	Susceptible
33	T.2269-2	Huti White x Calcutta 4	5.3 abcd	Susceptible
34	T.2691-15	Huti White x Calcutta 4	5.6 abc	Susceptible
35	NM 226-5	Mchare Laini x Calcutta 4	5.8 a	Susceptible
36	T.2317-1	Huti White x Borneo	6.0 a	Susceptible
37	NM 226-16	Mchare Laini x Calcutta 4	6.0 a	Susceptible
	IITA hybrids			
13	T.1610-13	IITA2145/1320 x Zebrina	1.1 kl	Resistant
15	T.1608-1	IITA 2145/1320 x CV-Rose	1.1 kl	Resistant
19	T.1608-2	IITA 2145/1320 x CV-Rose	1.3 kl	Resistant
20	T.1598-2	IITA2145/1320 x CV-Rose	1.5 jkl	Resistant

[1]The rhizome discoloration index (RDI) was scored using a scale of 1 to 6, where 1 = no internal symptoms and 6 = discoloration of the entire inner rhizome. [2]Means with the same letter within the same column do not differ significantly according to Fisher's test of least significant differences (P<0.05). [3]RDI in the screen house was rated 10 weeks after inoculation.

In India, NRCB screened 17 diploid accessions against Foc Race 1. The accessions with a DSI <30 are Cultivar Rose (16.7), Tongat (23.33), Calcutta 4 (23.33), Pisang Madhu (23.33), Erachivazha (25) and Pisang Lilin (28.6). The accessions with a DSI 30 – 60 are Anaikomban (30), Matti (30.6), Type 2X (33.3), Sanna Chenkadali (33.3), Pisang Jaribuaya (33.3) and Pisang Jajee (35.7). The accessions with DSI > 60 are Poomgalli (63.3), Kunnan (66.7), Beejikela (75), Njalipoovan (80.6) and Elavazha (83.3). Three improved diploids developed by NRCB were found to be resistant to Foc Race 1 (VCG 124) in pot and hotspot field trials. The TC propagules of these improved diploids are ready and currently being prepared for export to IITA. Of 7 improved parthenocarpic diploids received from IITA, 6 are under field evaluation at NRCB for

Foc Race 1 resistance and performance. Pisang awak (ABB) is being crossed with 10 diploids. Progressive generation of 400 (cumulative) new Pisang awak hybrids were obtained, of which 32 Pisang awak hybrids were field planted to screen for Foc Race 1 resistance and yield.

AT UM a segregating open-pollinated population of *Musa acuminata* ssp. *malaccensis* has been reconstituted (after the previous was lost) and maintained in the field. Plans have been made to export this population to Uganda.

UQ has a collection of 45 recombinants selected from more than 300 F2 'Ma851' and 'Ma852' individuals. Two thirds of them showed strong associations with the markers in this region, around marker 28820. This is consistent with the QTL-Seq results. However, 15 lines did not show a clear phenotype to genotype association. The reasons for this are currently unknown.

Weevils: Compared with the local *Matooke* cv Mbwazirume with 25% outer damage, none of the NARITAs were infested with weevils in Kagera and Arusha, with average outer damage of between 0 - 2%. NARITAs continue to be screened for weevil resistance in the screenhouse at TARI-Tengeru. A set of 9 diploids, shown to be resistant to *R. similis,* was screened for weevil resistance. The diploid cvs Calcutta 4, Pisang Tunjuk, Saing Hil, Pisang Rotan and Morongo Datu showed promising resistance, while Pisang Gigi Buaya, Huwundi Vita, Gabah Gabah and SH-3142 were similarly susceptible as the *Matooke* cv Nakitembe.

Nematodes: *P. goodeyi, R. similis, Meloidogyne* spp. and *Helicotelynchus multicinctus* were recovered from field testing sites but in low densities across sites in Tanzania, with no discernible differences across cultivars. Screenhouse screening assessment is underway. NARITA hybrids were evaluated for resistance to nematodes (*R. similis*) using the single root technique with two controls, Yangambi Km5 (resistant), Valery (susceptible) and 25 NARITA hybrids. The genotypes were evaluated at 8 weeks after inoculating root segments and nematodes counted and compared with controls. This data was were additionally enriched with field survival data from a Sendusu demonstration plot, 8 years after planting. From this study, 17 NARITA hybrids are resistant to *R. similis*, 1 partially resistant and 7 susceptible. Resistance did not necessarily translate to field survival, with only limited correlation between nematode counts and field survival (r = 0.22); field survival is understandably dependent on various other factors. The experiment was repeated but disregarded as the controls did not differ significantly.

In addition, the 23 improved hybrids from EMBRAPA have also been evaluated for nematode resistance using the same technique (Table 6). Six genotypes (BMPG-1, BMPG-10, BMPG-11, BMPG-18, BMPG-2 and BMPG-8) were found to be resistant, which should be useful in breeding for nematode resistance.

Sigatoka:To identify additional potential sources of resistance to *P. fijiensis*, 95 *Musa acuminata* (41 diploids, 24 triploids, 4 tetraploids, 7 improved diploids) and *M. balbisiana* (4 diploids, 14 triploids and one tetraploid) accessions were field-evaluated in Sendusu. Three groupings were identified: those highly resistant to *P. fijiensis* were less affected and symptom development ceased at early lesion development (stage 2); partially resistant accessions that allowed the pathogen to sporulate but with a lower disease severity score and susceptible accessions where most leaves were necrotic with high disease severity. The largest number of resistant accessions belong to *M. acuminata* ssp. *malaccensis*, *M. acuminata* ssp. *zebrina* and *M. acuminata* ssp. *burmannica*. Twenty-nine accessions showed either high or partial resistance to *P. fijiensis*. In addition to Calcutta 4, Long Tavoy, Pahang, Pisang KRA, 0074 Malaccencis, M.A Truncata, Tani and Balbisiana stopped the disease at an early lesion stage (stage 2). These accessions are potential sources of resistance to *P. fijiensis*. For others, such as Pisang Lilin, Borneo, Pisang Serun and Cameroun, symptom progression ceased at Stage 3. In Monyet and Porapora symptoms developed to Stage 4 but did not progress to the late necrotic stage. Black Sigatoka symptoms on the remainder of the accessions progressed to Stage 6. Having identified germplasm that halts symptom development at different stages, hence new perspectives to identify various multiple genes that are involved in Sigatoka resistance become available.

Under field conditions 21 NARITAs were evaluated for *P. fijiensis* resistance in five AEZs in Uganda and Tanzania through three crop cycles. The disease was visually assessed by estimating the leaf area covered with disease symptoms, and the disease severity calculated as the area under the disease progress curve for each plant over time. For all sites combined there was no genotype by cycle interaction. Significant differences were observed in the response of NARITA hybrids to black Sigatoka in different environments. Most of the hybrids developed less disease than cv Mbwazirume, the susceptible control. Eight hybrids (NARITAs 2, 4, 7, 8, 14, 21, 22 and 23) consistently had only limited disease symptoms across environments. These hybrids can be released to farmers in the region as an alternative to the highly susceptible landraces. The TACRI site does not show much variation between tested germplasm and should be discontinued in the future. In contrast Mbarara was an ideal test environment for evaluating black Sigatoka and can be used as a representative location to minimize costs.

Primary outcomes 3: WP 3 on Leveraging Genetics. Genetics of resistance to *Fusarium oxysporum* f.sp. *cubense* (Foc), burrowing nematode (*Radopholus similis*) and weevil determined in banana facilitating development of molecular markers for breeding; and a SNP-based genomic model for breeding for yield and agronomic traits developed, in order to accelerate the breeding cycle.

Scope and approach: Improved breeding efficiency through molecular-based genetic studies for increased understanding of underlying genetics and development of DNA marker-based early selection.

We report about the mapped QTL for banana resistance to Foc Race 1 and weevil (and banana bacterial wilt, which was a bonus to the project after finding 'Monyet' to be tolerant to the BXW disease). The work on nematodes failed as the nematode bioassay (screening) is not accurate. The QTL for Foc STR4 was fine-mapped, and a PCR-based marker developed and validated for the QTL. QTL analysis was also carried out for fruit-filling in *Matooke*, based on the phenotypic and genotypic data of the training population for genomic prediction. Prediction accuracy was improved for the evaluated predictive models based on Mbarara data, additional to data from Sendusu, Uganda.

1. Phenotyping the mapping populations
    a. Fusarium
    b. Nematodes
    c. Weevils
    d. Sigatoka
2. Developing a genomic selection model
3. Genotyping of QTL mapping
4. Genotyping with SSR markers
5. *Mchare* sequencing

6. Leaf archiving

All mapped QTL are summarized on the DH Pahang physical map, and the implication of their locations and genetic background discussed.

1. Phenotypic and genotypic analysis

1.1. Summary of the populations

Table summarizes the traits studied, the populations used for each trait, the number of lines phenotyped per population per trait and the genotyping platform for the generation of the SNP markers used in QTL analysis. Furthermore, the training population for genomic prediction was used to run a Genome-Wide Association Study (GWAS) on fruit-filling traits.

Table 9. Summary on the populations used in the project and the phenotyped traits

Outcome	Population (response of the parents)[a]	Phenotyping	Genotyping
QTL mapping - *Foc* R1	'Paliama' (S) x 'Borneo' (R)	165 lines	DArTSeq
	'Monyet' (R) x 'Kokopo' (S)	153 lines	DArTSeq
QTL validation and fine-mapping - *Foc* STR4	UQ Malaccensis progeny + other genotypes	96 lines+	QTL-Seq
QTL mapping - weevil	'Monyet' (R) x 'Kokopo' (S)	139 lines	DArTSeq
	'Kasaska' (S) x 'Borneo' (R)	211 lines	GBS
QTL mapping - nematodes	'Calcutta 4' (R) x 'Zebrina GF' (S)	150 lines	DArTSeq
	'Kasaska' (R) x 'Borneo' (S)	217 lines	GBS
QTL analysis for BXW	'Monyet' (T) x 'Kokopo' (HS)	121 lines	DArTSeq
Predictive models - *Matooke* breeding	Training population	307 lines	GBS
GWAS - yield and fruit-filling	Training population	307 lines	GBS

[a] S: susceptible, R: resistant, T: Tolerant, HS: highly susceptible

a. Genotyping

DNA (Table) was extracted at IITA-Sendusu using the CTAB protocol, optimized for banana leaf samples. Quality of the obtained DNA was checked using a 0.8% agarose gel through electrophoresis. DNA quantity was estimated using a Nanodrop 2000 machine. DNA diluted to a final concentration of 150 ng/µl. Genotyping was outsourced, first at the Institute of Genomic Diversity (IGD), Cornell University, where genotyping by sequencing (GBS) was used for SNP discovery and genotyping of the 'Kasaska' x 'Borneo' population. This platform could not be used for other populations due to the discontinuity of the service at IGD. Later, 3 more populations (Table) were genotyped using the DArTSeq® platform through the Integrated Genotyping Service and Support (IGSS), a project under ILRI and implemented by the Biosciences eastern and central Africa-International Livestock Research Institute (BecA-ILRI) in partnership with Diversity Arrays Technology Pty Ltd. The two genotyping platforms used enzyme PstI for genome complexity reduction. The sequence reads were aligned to the DH Pahang reference banana genome. Details of the SNP genotypic data for each population are given below.

Genotyping of the 'Paliama' x 'Borneo' population

'Paliama' x 'Borneo' population was genotyped using the DArTSeq® platform through IGSS. This population was purely bi-parental, with all lines derived from 'Paliama' as the female parent, and 'Borneo' as the male parent. Together with the parents 197 hybrids were genotyped, which provided 3,621 polymorphic loci. However, 'Paliama' appears to be mostly homozygous, with only 1.3% heterozygosity. It is therefore speculated that 'Paliama' has some level of selfing, resulting in almost all the loci being homozygous. As a result, a linkage map was constructed using JoinMap®4.1 from the markers segregating in 'Borneo'. After data cleaning and limiting the missing data to 20% per locus, a linkage map of 2,778 SNPs was constructed, with the markers grouped in 11 linkage groups (LG, Figure). The groupings were the same as on the DH Pahang physical map, with some variations in the order of the markers within each chromosome. The length of the LGs ranged from 1,004 cM to 2,154 cM, with an average inter-marker distance of 6.46 cM. This 'Borneo' linkage map represents the densest banana linkage map, with the highest number of molecular markers, spanning all 11 banana chromosomes.

Genotyping of the 'Monyet' x 'Kokopo' population

This population, using 153 hybrids and the two parents, was genotyped using DArTSeq® through IGSS (Table). The female parent 'Monyet', was originally thought to be a diploid but was later shown to be a tetraploid (4x), resulting in a 4*x* X 2x cross, providing triploid (3x) progeny. Allele calling was carried out using the Genome Analysis Tool Kit (GATK) to cater for allele dosage for 'Monyet' and the progeny. 'Monyet' had 5 classes of genotypes, as expected from a 4x line, while 'Kokopo' had three classes, as expected for a 2x. The SNP data was categorized into the different genotype combinations, based on parental genotypes (Table). The markers in blue are monomorphic between the two parents and in the progeny, and the markers in purple are polymorphic but not segregating in the progeny. As a result, the markers segregating in the population amounted to 18,009 SNPs. These markers were used in QTL mapping for the traits that the 'Monyet' x 'Kokopo' population was phenotyped for, as presented in Table

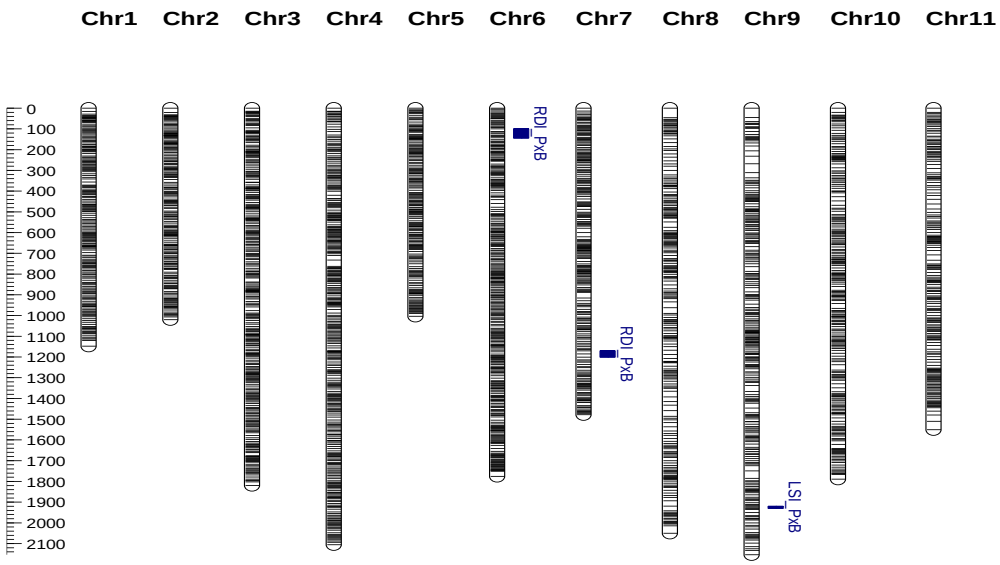

Figure 4. Linkage map of the 'Paliama' x 'Borneo' population with DArTSeq SNPs and 197 hybrids. The 11 obtained linkage groups corresponded with the 11 banana chromosomes. QTL for rhizome discolouration index (RDI) and leaf symptom index (LSI) are indicated in blue boxes on Chr6, 7 and 9.

Table 10. Segregation in the 'Monyet' x 'Kokopo' population: 0 represent the reference allele, and 1 represent the alternative allele.

Loci classes in 'Monyet'		Loci classes in 'Kokopo'		
		0	**1**	**11**
	0	499	9635	3458
	1	451	189	48
	11	2035	390	206
	111	243	109	84
	1111	5568	4619	9902

Genotyping of the 'Calcutta 4' x 'Zebrina GF' population

'Calcutta 4' x 'Zebrina GF' population was also genotyped using DArTSeq® platform through IGSS. Genotyping was performed on 144 hybrids and the two parents. After data cleaning, limiting missing data to 20%, 21,105 SNPs were segregating in the population, with a heterozygosity level of 20% in 'Zebrina GF' and 16% in 'Calcutta 4'. However, this population was not used in mapping as the nematode phenotypic data were not reliable (see further).

Genotyping of the 'Kasaska' x 'Borneo' population

'Kasaska' x 'Borneo' population was genotyped using the GBS platform at IGD. Together with genotypes 'Borneo', 'Kasaska' and an F_1 line 226 hybrids were genotyped. Within this population 18,095 markers were segregating, which, based on SSR genotyping, was a result of crosses between lines derived from 'Borneo' x 'Kasaska'.

b. Phenotyping and QTL analysis for Foc R1

i. Phenotyping for *Foc* R1

Two populations were found segregating for Foc Race 1. These were 'Monyet' x 'Kokopo', developed by and maintained at IITA-Sendusu, and 'Paliama' x 'Borneo', developed by and maintained at IITA-Arusha, which were phenotyped in Kawanda (IITA-NARO), and Arusha (IITA), respectively. An additional population was imported from UQ, a self-pollinated cross of *Malaccensis*, referred to as *Malaccensis* x *Malaccensis* population, and maintained at IITA-Arusha. Phenotyping was carried out at both Kawanda and Arusha, using the Foc-colonized millet protocol (Viljoen et al. 2016[1]) in pots using VCG 0124 in Kawanda and VCG 0125 in Arusha. The two VCGs belong to the same Foc lineage VI and are among the most abundant in the region (Karangwa et al. 2018)[2].

For 'Monyet' x 'Kokopo', 153 hybrids were successfully phenotyped together with the parents. Mbwazirume was used as the resistant check, while Kayinja (Pisang awak) and Sukali Ndizi were used as the susceptible checks (Figure). Of the three traits for Foc susceptibility (corm discoloration, yellowing of leaves, and stem splitting), rhizome discoloration is considered the most reliable trait. Phenotypic variation for corm discoloration was continuous, ranging from a score of 1 (no discoloration) to 6 (more than 50% of the corm discolored). Broad sense heritability,

[1] Viljoen, A., et al. (2016). Banana pests and diseases field guide for disease diagnostics and data collection improvement. IITA.
[2] Karangwa, P., et al. (2018). "Genetic Diversity of *Fusarium oxysporum* f. sp. *cubense* in East and Central Africa." Plant Disease **102**(3): 552-560.

as expressed in Equation 1, taking into account the number of replications, was estimated at 0.67 and 0.57 for rhizome discoloration and leaf symptom index respectively.

$$H^2=V_g/(V_g+V_e/r) \quad (Eq.\ 1)$$

Where:
H^2 =Broad sense heritability
V_g = genetic variance
V_e = error variance
r = number of replications

Figure 5. Rhizome/corm discoloration in 'Kokopo', Foc R1 susceptible male parent, 'Monyet', female resistant parent, and Kayinja, the susceptible check.

For 'Paliama' x 'Borneo' population, 162 hybrids were phenotyped together with the parents, the resistant check (Km5), and the susceptible checks (Sukali Ndizi, Gros Michel). Using Equation 1, broad-sense heritability was estimated as 0.44 for rhizome discoloration and 0.44 for leaf symptom index. In addition, 133 hybrids from the population *Malaccensis* x *Malaccensis* were phenotyped for Foc R1 in Arusha. This population showed a broad-sense heritability of 0.32 for rhizome discoloration and 0.33 for leaf symptom index. QTL analysis is not reported as the genotypic data are not yet available.

ii. QTL analysis for *Foc* R1

The genotypic data for the population 'Monyet' x 'Kokopo' were combined with the phenotypic data for Foc R1 in a QTL analysis for rhizome discoloration and leaf symptom index. Because of the 4x x 3x nature of the population, a quantitative QTL analysis approach was used, where the phenotypes were regressed against the allele frequencies instead of the genotypes. The analysis was carried out using customized scripts in GenStat Ed. 19., using the DH Pahang physical map to order the markers on the chromosomes. Twelve loci were significantly associated with rhizome discoloration, with the highest -Log(P) value of 4 (

A

Figure). Six of these markers were located on Chr11 with 21 more markers in the same region having a -Log(P) value greater than 2.5, forming the QTL region circled in green on

Figure A. Two additional putative QTL are located on Chr8 with 3 significant loci, and on Chr9. The remaining significant markers are located on chromosome 1 (1 locus) and chromosome 10 (1 locus). This trait is the most reliable scorable trait and constitutes the ultimate measure for banana resistance to Foc. The putative QTL for rhizome discoloration on Chr11, 8 and 9 reappear for leaf symptom index (

Figure B) but with 1 or 2 loci significantly associated with the trait at each chromosome. The remaining significant markers were randomly located on Chr1 (2 loci), 6 (2 loci), 7 (2 loci), 10 (1 locus).

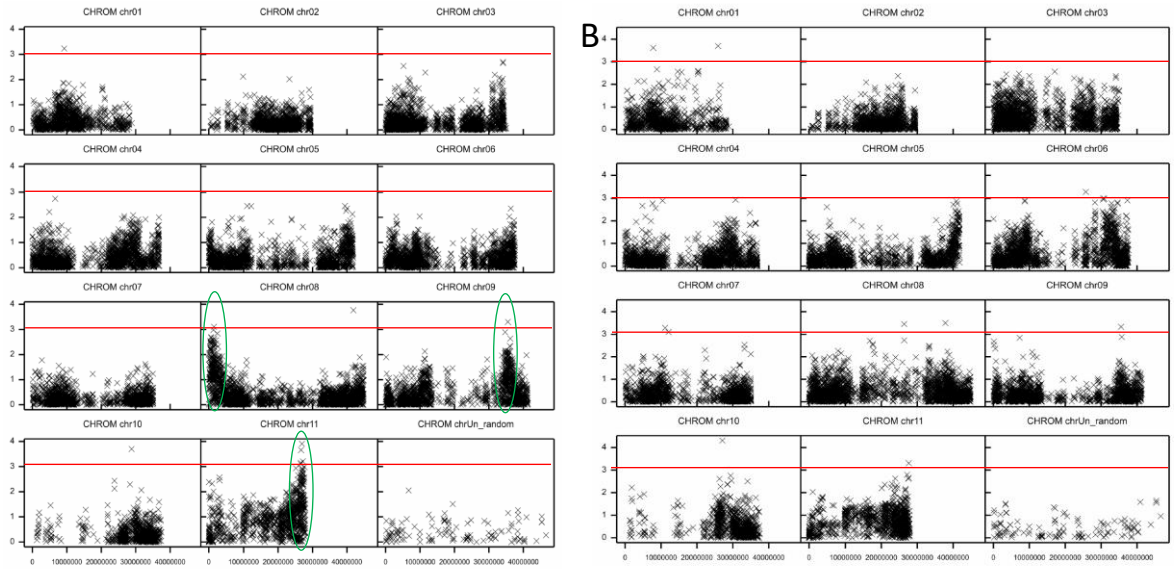

Figure 6. QTL analysis for rhizome discoloration (A) and leaf symptom index (B) in the 'Monyet' x 'Kokopo' population. The red line indicates the threshold level of -Log(p) = 3. All markers above the threshold are significant. Green shows the location of a QTL.

The ‘Borneo’ linkage map was combined with the genotypic and phenotypic data from the ‘Paliama’ x ‘Borneo’ hybrids using MapQTL® 6.0 software. The Multiple QTL Model (MQM) analysis shows two significant QTLs for rhizome discoloration on Chr6 and Chr7 (Figure , Figure A and 7B); and 1 QTL for leaf symptom index on LG9 (Chr9, Figure C). The two QTLs for rhizome discoloration had a contrasting effect. The QTL on Chr9 had a positive additive effect of 0.62, increasing susceptibility with an explained variance of 12.4%, while the QTL on Chr7 had a negative additive effect of -0.33, hence contributing to the resistance of the QTL, with an explained variance of 12.5%.

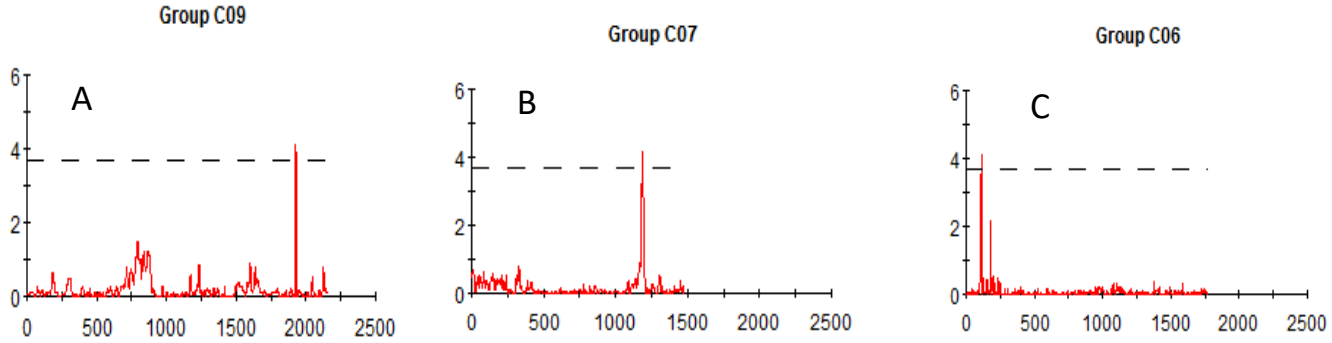

Figure 7. Multiple QTL Model (MQM) graphs for corm discoloration (A and B) and leaf symptom index (C) in the ‘Paliama’ x ‘Borneo’ population on 11 linkage groups. The horizontal dashed line indicates the 95% significant threshold of LOD value for a QTL. Only the LGs with significant QTL are shown.

1.3.3. Fine mapping of the Foc STR4 QTL

The objective in fine mapping is to accurately define the genetic intervals on Chr3 for the QTL locus using progeny testing and Foc STR4. A SNP haplotype of 10 SNPs was found to determine resistance and susceptibility for this QTL (Figure). The region is 247,887 bp in length and contains 28 putatively defined gene models, 15 of which are annotated as leucine rich repeat receptor-like kinases. The screening of the large F_2 ‘Ma851’ population (>300 screened) identified many individuals carrying meaningful cross-over events in this region. A collection of 45 recombinants was selected from more than 300 F_2 ‘Ma851’ and ‘Ma852’ individuals. Two thirds of them showed strong associations with the markers in this region, around marker 28,820. However, the rest of the lines (15) did not show a clear phenotype to genotype association. This could be due to the changing conditions, which did not favor Foc infection in the glasshouse, as some of the plants did not show the level of severity in infection as expected. These lines will be re-screened.

Furthermore, characterization of the F_2s from the inter-cross between the ‘Ma851’ and the susceptible ‘Ma848’ has begun, with 60 F_2 lines ready for testing and an additional 100 lines in the pipeline. Since the parent ‘Ma848’ is extremely susceptible to all Race 4 types, it would be interesting to see if this is due to the presence of susceptibility factors. A QTL-seq analysis will be performed on this population for new QTLs.

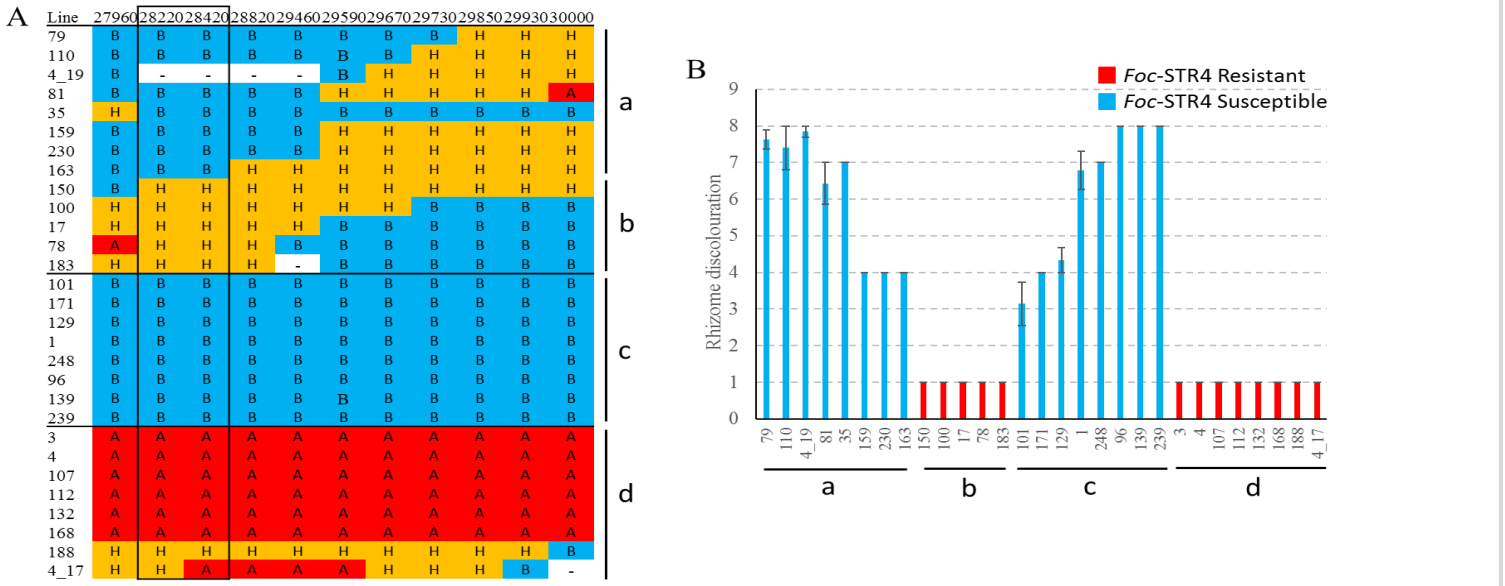

Figure 8. A. Fine mapping of the Chr3 QTL locus using progeny testing and Foc STR4. Allele marker haplotypes A, B, and H correspond to the resistant, susceptible and heterozygous marker haplotypes, respectively. a, b, correspond to the respective susceptible and resistant progeny groups carrying cross-over events in this region; c, d, correspond to the respective susceptible and resistant non-recombinant groups, respectively. B. Phenotypic assessment of 1 to 13 clones per progeny from panel A, challenged with Foc STR4.

c. Phenotyping and QTL analysis for weevil resistance

i. Phenotyping for weevil resistance

Two populations were phenotyped for weevil resistance, 'Kasaska' x 'Borneo' and 'Monyet' x 'Kokopo' in pots, using Mbwazirume as the susceptible check and Km5 as the resistant check. From the 'Kasaska' x 'Borneo' population 211 hybrids were phenotyped and 139 from 'Monyet' x 'Kokopo'. Traits considered were peripheral damage and total cross-section damage. Total cross-section damage was derived from four traits measured from the plant collar downwards: outer cross-section damage at 3 and 6 cm, inner cross-section at 3 and 6 cm, using Equation 2. Of the two traits, total cross-section damage is the most important as damage on the inside of the corm is more serious.

$$TXD = (XOD\_3cm + XID\_3cm + XOD\_6cm +)/4 \quad \textit{(Eq. 2)}$$

Where:

TXD = total cross section damage
XOD_3cm = percentage outer cross-section damage at 3 cm
XID_3cm = percentage inner cross-section damage at 3 cm
XOD_6cm = percentage outer cross-section damage at 6 cm
XID_6cm = percentage inner cross-section damage at 6 cm

Data analysis showed a broad-sense heritability of 0.33 for peripheral damage and 0.43 for total cross-section damage in 'Kasaska' x 'Borneo' population; and 0.24 for peripheral damage and 0.54 for total cross-section damage in the 'Monyet' x 'Kokopo' population.

ii. QTL analysis for weevil resistance

SNP markers from the 'Monyet' x 'Kokopo' population and the phenotypic data were combined for quantitative QTL analysis for peripheral damage and total cross-section damage, using the physical map of DH Pahang to order the markers within chromosomes. Two regions of the genome seem to confer weevil resistance at the beginning and at the end of Chr6, for total cross section damage. However, the region at the beginning of Chr6 has two QTLs. One region has a peak at 88,0914 bp with a -Log(P) of 4.09, and the reference allele has a positive effect (= reduces resistance) of 0.81 (logit-transformed). 'Monyet' carries two copies of the reference allele and two copies of the alternative allele, while 'Kokopo' is homozygous for the alternative allele. The second region has a peak at 2,795,989 bp with a -Log(P) of 3.87, and the reference allele has a negative effect of -0.81 (logit-transformed, increases resistance), with 'Monyet' carrying two copies of the reference allele and two copies of the alternative allele. The QTL at the end of the chromosome has a peak at 34681454 bp, -Log(P) of 3.27 and it contributes to resistance with an effect of -0.54 (logit-transformed, Figure A). The QTL at end of Chr6, which was found for total cross-section damage, was also significant for peripheral damage but wider, with a peak at 33,938,938 bp (Figure B). Two additional putative QTL were found on Chr7 and Chr11. Additional significant markers were located on Chr4 (1 locus) and Chr5 (2 loci, Figure B).

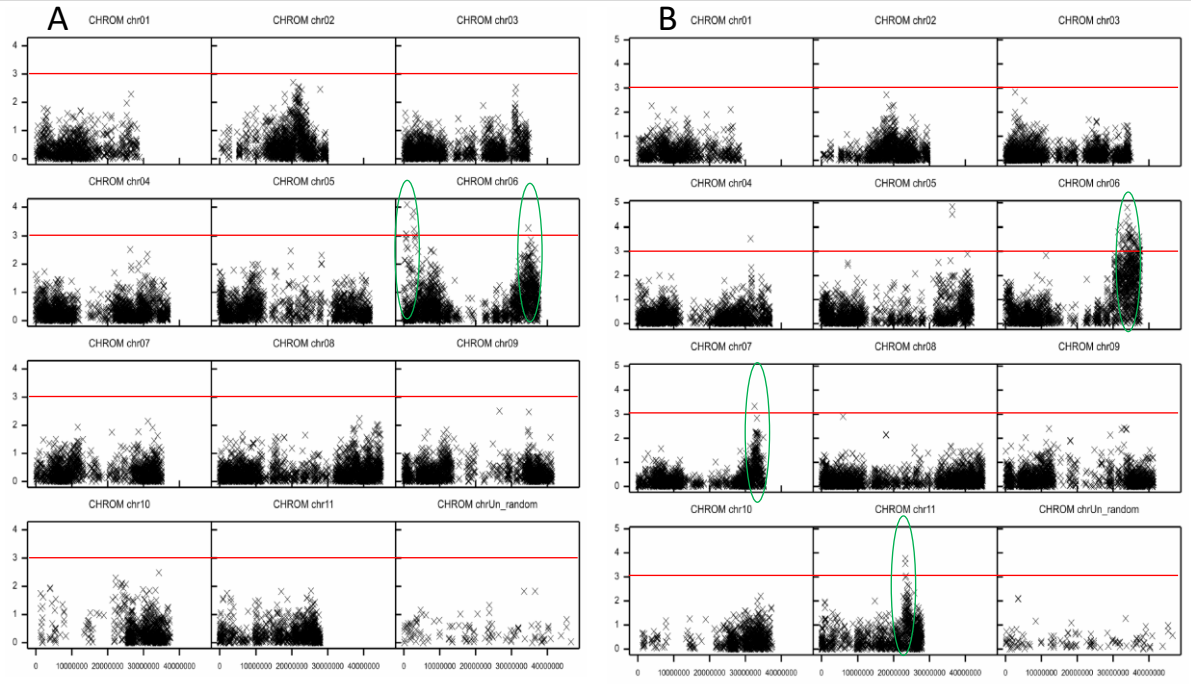

Figure 9. QTL analysis results for Total cross-section Damage (A) and Peripheral Damage (B) in the 'Monyet' x 'Kokopo' population. The traits (%) are logit-transformed. The red line indicates the threshold level of -Log(P) = 3. All markers above the threshold are significant. Green shows the location of a QTL.

'Kasaska' x 'Borneo' population had been genotyped using the GBS platform. QTL analysis was carried out using marker-trait association method as implemented in GenStat® 19th edition for peripheral damage and total cross-section damage, with 197 hybrids and 18,009 SNP markers. At a threshold of -Log_{10}(P) of 3, significant marker-trait association was identified on Chr3 for total cross-section damage, with a -Log(P) value of 7.70, at 23,329,852 bp (Figure 10A). The same region was also significant for peripheral damage (Figure B), with an additional QTL on Chr10, and a few barely significant markers on Chr1, and the unanchored contigs (pseudo-chromosome 12). Because the GBS sequence data for this population were aligned to the first version of the DH Pahang genome, in-silico mapping was carried out to convert the positions of the SNP on the 2nd version. The new positions of the QTLs were located on the new version, and the 2 significant markers on pseudo-chromosome 12 were mapped on Chr3 and Chr10, which fit within the 2 QTLs for total cross-section damage. The new positions of the QTL were used to map all the QTLs on version 2 of the DH Pahang genome.

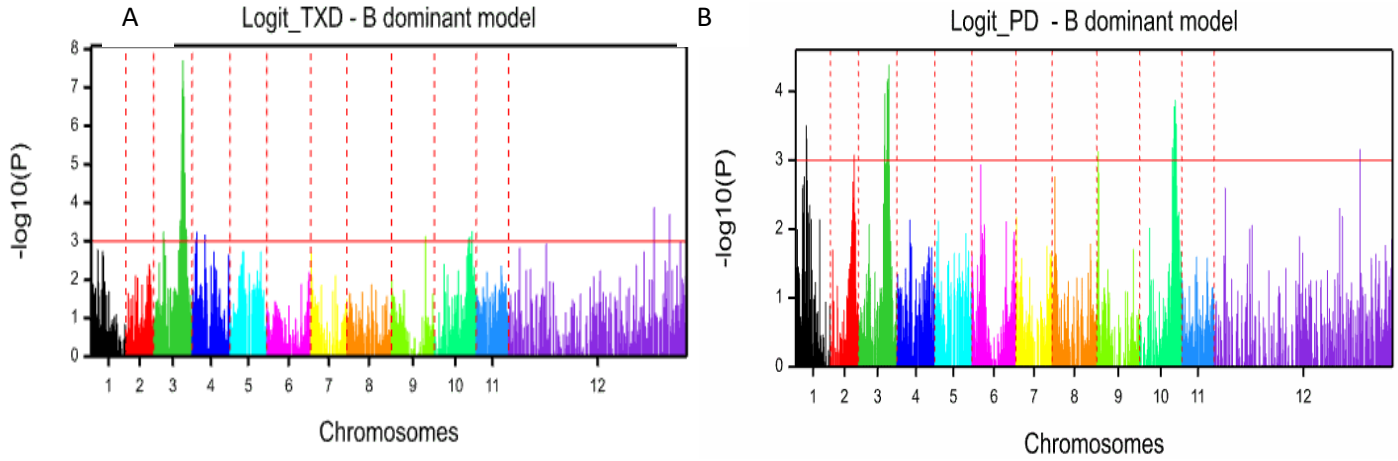

Figure 10. QTL analysis results for Total Cross-Section Damage (A) and Peripheral Damage (B) in the 'Kasaska' x 'Borneo' population. The traits (%) are logit-transformed. The red line indicates the threshold level of -Log(P) = 3. Markers above threshold are significant.

d. Phenotyping for nematode resistance (*Radopholus similis*)

Two populations were phenotyped for resistance to *R. similis*: 'Kasaska' x 'Borneo' and 'Calcutta 4' x 'Zebrina GF'. A total of 217 hybrids from 'Kasaska' x 'Borneo' and 150 from 'Calcutta 4' x 'Zebrina GF' were phenotyped in the screenhouse at Sendusu using the single-root screening method, with Km5 as the resistant and Valery as the susceptible checks. Total nematode count and necrosis of each root were scored, after inoculation of 5-8 roots per plant (Table 11). The two traits had a correlation of 0.47 in 'Kasaska' x 'Borneo' and 0.48 for 'Calcutta 4' x Zebrina GF' (two-sided test, *P<0.001*). Although significant, these correlation values are not impressive, as they mean that high number of nematodes does not necessarily translate into more root damage, and *vice-versa*. However, among the two, total nematode count is the least subjective, as necrosis was based on visual scoring as a percentage of the damaged area over the healthy part of the root. The hybrids exhibited continuous segregation for the trait, ranging from 0 to 6,022 nematodes. Total nematode count varied between experiments for the same genotype, and

from one root to another of the same plant, especially for the susceptible check, with mean total nematode counts for this genotype ranging from 130 to 54,444. Combining the lines from different experiments was unreliable and a more consistent method is necessary to reliably assess for nematode resistance. For instance, among the 60 hybrids, repeated in at least two experiments, 47% had inconsistent results. QTL analysis was therefore deemed unreliable. There is urgent need to develop a high-throughput assay for screening mapping populations for nematode resistance, and to identify reliable, objective traits, which are highly correlated with nematode damage.

Table 11. Identification of the host response to *R. similis* based on a comparison with the host response of susceptible (Valery) and resistant (Km5) checks

Comparison with KM5	Comparison with Valery	Response
Not significantly different	Significantly different	Resistant
Significantly different	Not significantly different	Susceptible
Significantly different	Significantly different	Partial resistant
Not significantly different	Not significantly different	Inconclusive

e. Phenotyping and QTL analysis for resistance to banana bacterial wilt

i. Phenotyping for resistance to banana bacterial wilt

This study was carried in the 'Monyet' x 'Kokopo' population, when 'Monyet' was determined to be tolerant and 'Kokopo' as highly susceptible to the disease. In an outdoor confined pot trial 121 hybrids were evaluated using plants generated from corms. The midribs of the youngest leaf of three-month-old banana plants were inoculated with $1x10^8$ CFU/ml of Xvm isolate SY103C (stored at -80ºC) and symptom development assessed weekly for four months. Parameters scored included time to symptom expression, leaf wilting and plant death. The disease severity scale of 0 to 3 developed by Winstead and Kelman (1952)[3] and modified by Nakato et al. (2018)[4] was used to compute disease index using Equation 3.

*Disease index (DI) = (((1*A) + (2*B) + (3*C)) / No. of plants) *100* *(Eq. 3)*

Where:
A = number of plants with inoculated leaf showing symptoms
B = number of plants with uninoculated leaves showing symptoms
C = number of wilted (dead) plants

Classification of genotypes into resistance and susceptible categories was based on a scale developed by Tripathi et al. (2008)[5] and modified : Resistant (R) – no plants wilted; Tolerant (T) – < 30% plants wilted; Moderately Susceptible (MS) – >30%- <50% plants wilted and Highly Susceptible (HS) - >50% plants wilted. Area under the disease progressive curve (AUDPC) was calculated to quantitatively summarize disease intensity over time and to analyze differences among genotypes using Equation 4.

$$\text{AUDPC} = \sum_{i=1}^{n=1} \left(\frac{\text{Yi}+\text{Yi}+1}{2}\right)(ti + 1 - ti) \quad \textit{(Eq. 4)}$$

Where:
t = time in weeks of each reading
y = percentage of affected plants at each reading
n = number of readings.
i = reading

Based on AUDPC, 'Monyet' was confirmed as tolerant and 'Kokopo' as highly susceptible (

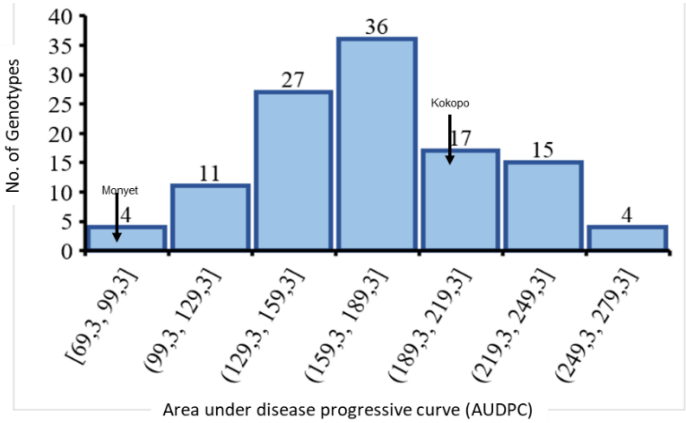

Figure). The hybrids ranged from: as tolerant as 'Monyet', to more susceptible than 'Kokopo' (

[3] Winstead NN, Kelman A, 1952. Inoculation techniques of evaluating resistance to Pseudomonas solanacearum. Phytopathology **42**: 111–4
[4] Nakato, G. V., et al. (2019). "Sources of resistance in *Mus*a to *Xanthomonas campestris* pv. musacearum, the causal agent of banana xanthomonas wilt." Plant Pathology **68**(1): 49-59
[5] Tripathi L, Odipio J, Tripathi JN et al., 2008. A rapid technique for screening banana accessions for resistance to *Xanthomonas* wilt. European Journal of Plant Pathology **121:**9–19.

Figure). Broad-sense heritability (Equation 1) for AUDPC at 112 days after inoculation was estimated at 0.69 for disease index, and at 0.63 for area under disease progressive curve.

Figure 11. Phenotypic variation for BXW resistance of 'Monyet' x 'Kokopo' hybrids expressed as area under disease progression curve. The two parents are indicated.

ii. QTL analysis for resistance to banana bacterial wilt

Quantitative QTL analysis was performed using the genotypic data of the population 'Monyet' x 'Kokopo', the DH Pahang physical map and the phenotypic data described in

Figure . A QTL for both area under disease progressive curve and disease index (DI) is found on Chr5 (Figure A and B), with a peak located at 32,425,876 bp, with a -Log(P) value of 4.75. At this locus, 'Monyet' is homozygous for the reference allele, which contributes negatively to AUDPC (-19.60), meaning that it confers resistance. Five more markers are significantly associated with AUDPC, 2 on Chr2, 1 on Chr6, 1 on Chr9 and another one on Chr11. But these markers have lower P-values than those on Chr5.

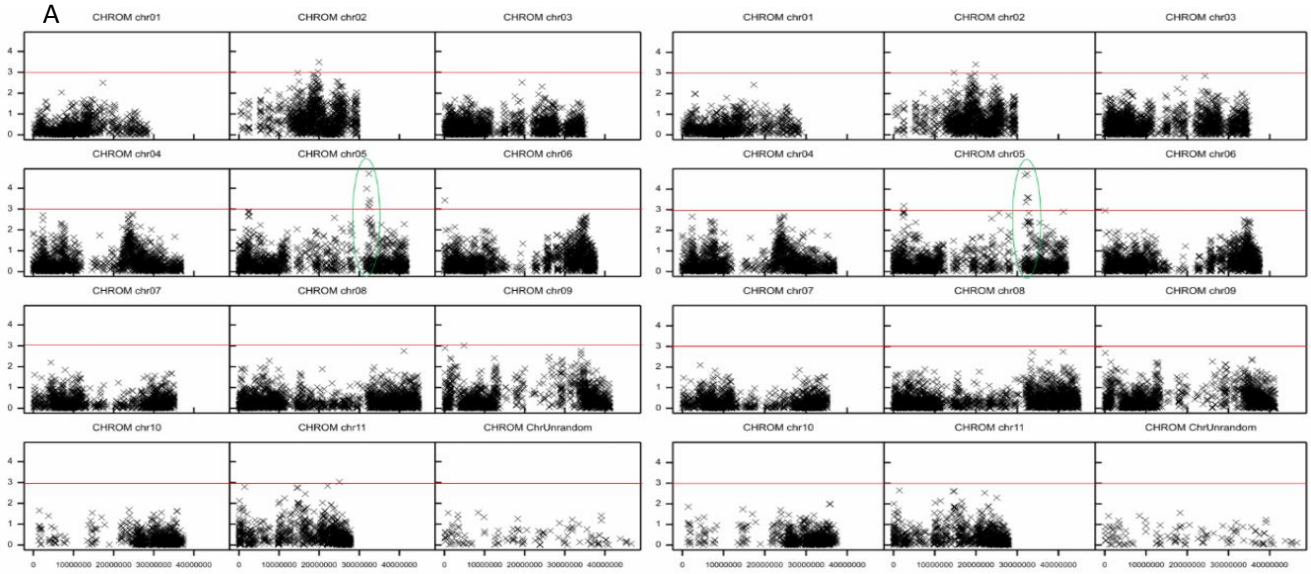

Figure 12. QTL analysis results for BXW resistance traits: AUDPC (A) and DI (B) in the 'Monyet' x 'Kokopo' population. The red line indicates the threshold level of -Log(p) = 3. All markers above the threshold are significant. Green shows the location of a QTL.

f. QTL analysis for fruit filling

Fruit filling is an important trait in banana breeding because it forms the basis for >80% of EET material rejection. Phenotypic data of the training population (307 genotypes) from the two fields in Sendusu (low-input management and high-input management) and the field in Mbarara were combined with SNP data from GBS (27,178 markers) in a genome-wide association analysis for yield components. The traits considered were number of hands and number of fruits, fruit length, fruit circumference and diameter of both fruit and pulp. Heritability of the traits, taking into account 3 cycles (term c) and 3 sites (term f) (Equation 5) ranged between 0.83 and 0.98.

$$H\text{^}2 = \frac{Vg}{Vg + \frac{Vc}{3} + \frac{Vgfc}{9} + \frac{Ve}{27}} \qquad \textit{(Eq. 5)}$$

Where:
H^2 (H^2) = heritability
Vg = Genetic variance
Vc = Variance associated with cycle
Vgfc = Variance associated with interaction between genotype, field and cycle
Ve = error variance

Tassel V5 was used for GWAS in a REML analysis, taking into account population structure and kinship. Using Bonferroni correction, false discovery rate, and long-range linkage disequilibrium (LD), a QTL on Chr3 was associated with the above-mentioned traits (

Figure 1), with additional loci on Chr6, 9, 10 and 3 loci among the unanchored contigs. Most SNPs were located in genes encoding uncharacterized and hypothetical proteins, but some mapped to transcription factors and genes involved in cell cycle regulation (Nyine et al. 2019[6]).

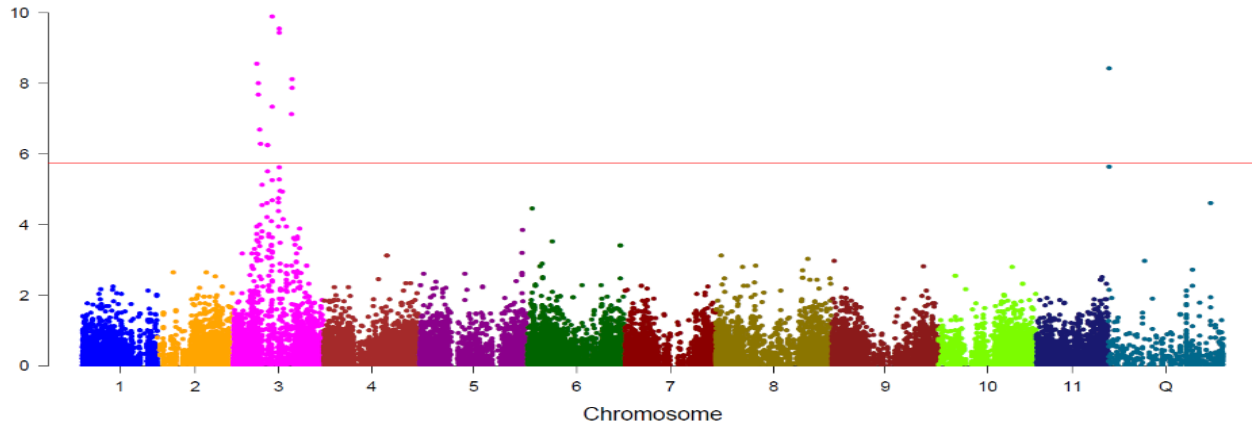

Figure 1. A Manhattan plot for a GWAS analysis on fruit circumference

Table 12. Comparison of average correlation for five-fold cross validations between the predicted and observed phenotypes across models fitted with data from low input (LIM) and high input management (HIM) fields, as published in Nyine et al. (2018; https://doi.org/10.1007/s00122-019-03425-x), combining the two fields with data from cycles 1, 2 and 3 (All) and all the data from Sendusu and Mbarara over 3 cycles (3 fields, 3 cycles).

Traits	BRR				BayesA				BayesB				RKHS_M				BL
	LIM	HIM	Sen	All	LIM	HIM	Sen	All	LIM	HIM	Sen	All[a]	LIM	HIM	Sen	All	All
Plant height	0.54	0.46	0.2	**0.57**	0.54	0.45	0.53	**0.56**	0.54	0.44	0.49	-	0.55	0.44	0.49	**0.57**	0.57
Plant girth	0.6	0.52	0.54	**0.55**	0.6	0.51	0.52	**0.56**	0.6	0.52	0.5	-	0.6	0.51	0.5	**0.55**	0.56
Total suckers	0.16	0.17	0.33	-	0.17	0.2	0.39	-	0.16	0.19	0.31	-	0.17	0.18	0.39	-	-
Height of tallest sucker at flowering	0.28	0.18	0.41	**0.42**	0.28	0.18	0.38	**0.40**	0.27	0.2	0.37	-	0.28	0.19	0.33	**0.43**	0.42
Height of tallest sucker at harvest	0.27	0.26	0.43	**0.45**	0.28	0.25	0.41	**0.44**	0.28	0.24	0.38	-	0.26	0.26	0.38	**0.45**	0.43
# of standing leaves at flowering	0.36	0.42	0.45	**0.51**	0.37	0.42	0.29	**0.51**	0.43	0.4	0.48	-	0.37	0.41	0.44	**0.58**	0.52
Index of non-spotted leaves	0.35	0.42	0.54	**0.55**	0.34	0.43	0.56	**0.56**	0.34	0.43	0.55	-	0.35	0.42	0.57	0.54	0.54
Days to fruit maturity	0.47	0.42	0.51	**0.56**	0.47	0.42	0.47	**0.54**	0.47	0.42	0.50	-	0.47	0.42	0.51	**0.55**	0.55
Bunch weight	0.63	0.61	0.64	**0.66**	0.62	0.62	0.6	**0.67**	0.64	0.62	0.61	-	0.61	0.61	0.60	**0.65**	0.64
# of hands	0.6	0.62	0.58	**0.68**	0.59	0.63	0.58	**0.67**	0.6	0.62	0.58	-	0.59	0.62	0.58	**0.68**	0.68
# of fruits	0.47	0.51	0.55	**0.58**	0.47	0.52	0.55	**0.57**	0.47	0.52	0.55	-	0.45	0.52	0.55	**0.58**	0.57
Fruit length	0.65	0.64	0.65	**0.70**	0.65	0.64	0.66	**0.71**	0.67	0.65	0.65	-	0.64	0.64	0.65	**0.70**	0.70
Fruit circumference	0.67	0.66	0.65	**0.67**	0.66	0.66	0.67	**0.69**	0.7	0.69	0.64	-	0.65	0.66	0.64	**0.67**	0.69
Fruit diameter	0.67	0.63	0.63	**0.67**	0.66	0.67	0.65	**0.69**	0.7	0.71	0.64	-	0.65	0.67	0.63	**0.67**	0.68
Pulp diameter	0.67	0.68	0.54	**0.67**	0.66	0.68	0.66	0.67	0.7	0.72	0.63	-	0.65	0.67	0.65	0.66	0.67

[6] Nyine, M., et al. (2019). "Association genetics of bunch weight and its component traits in East African highland banana (Musa spp. AAA group)." Theoretical and Applied Genetics. (Online) https://doi.org/10.1007/s00122-019-03425-x.

Bold: prediction accuracy within the same model equal or better than the previous results, [a]analysis underway; Underlined: the best prediction across models considering the new analysis; Sen = Sendusu

a. Improvement of the predictive models

The training population of 307 genotypes was evaluated in Sendusu and Mbarara under three environments (three fields). The existing predictive models were revisited for improvement using 3-cycle data in the 3 fields at each site. Statistical analysis of the data was carried out using the mixed model in Equation 6.

Response = General mean + Genotype + Field/Cycle + Genotype*Field/Cycle + Error (*Eq. 6*)

with the underlined terms taken random in the model. The Best Linear Unbiased Predictors (BLUPs) were used in the evaluation of the models, using 10,807 bi-allelically-scored SNPs from GBS (Nyine et al. 2018; https://doi.org/10.3835/plantgenome2017.10.0090). Prediction accuracy was improved for most of the traits by adding data from Mbarara (Table 12). Best prediction accuracy ranged from 0.43 to 0.71. No single model was best for all the traits, but some models were equally good for the same trait. Yield components: bunch weight and fruit-filling had high prediction accuracy, ranging from 0.67 to 0.71. The same traits were used in a GWAS analysis, and a QTL was located on Chr3 for them. A panel of 6,000 genotypes (validation population) has been established, based on the families used to develop these models, and planted in Sendusu, with preparations underway to duplicate it in Mbarara for phenotyping. This population will be used to validate the results of the predictive models, then genomic prediction will be deployed in *Matooke* breeding.

Way forward

This project has set the stage for the development of molecular tools to speed selection in banana breeding. Numerous QTLs have been identified, which are associated with resistance to pests and diseases, mapped from various genetic backgrounds and located on different chromosomes (Figure 2). The heterozygosity nature of banana should not be overlooked, however, which likely explains the low level of explained variance by the QTLs. We propose to validate the mapped QTLs as much as possible. The QTLs validated for the same genetic background will be validated against different backgrounds, using breeding material (preferably diploids). Furthermore, the evaluated models for genomic prediction and the fruit filling QTL will be validated (as far as possible), and if successful, deployed in *Matooke* breeding. New mapping populations should be developed in order to identify alternative sources of resistance, especially for Calcutta 4, since this genotype has resistance to multiple pests and diseases and has been used intensively in banana breeding.

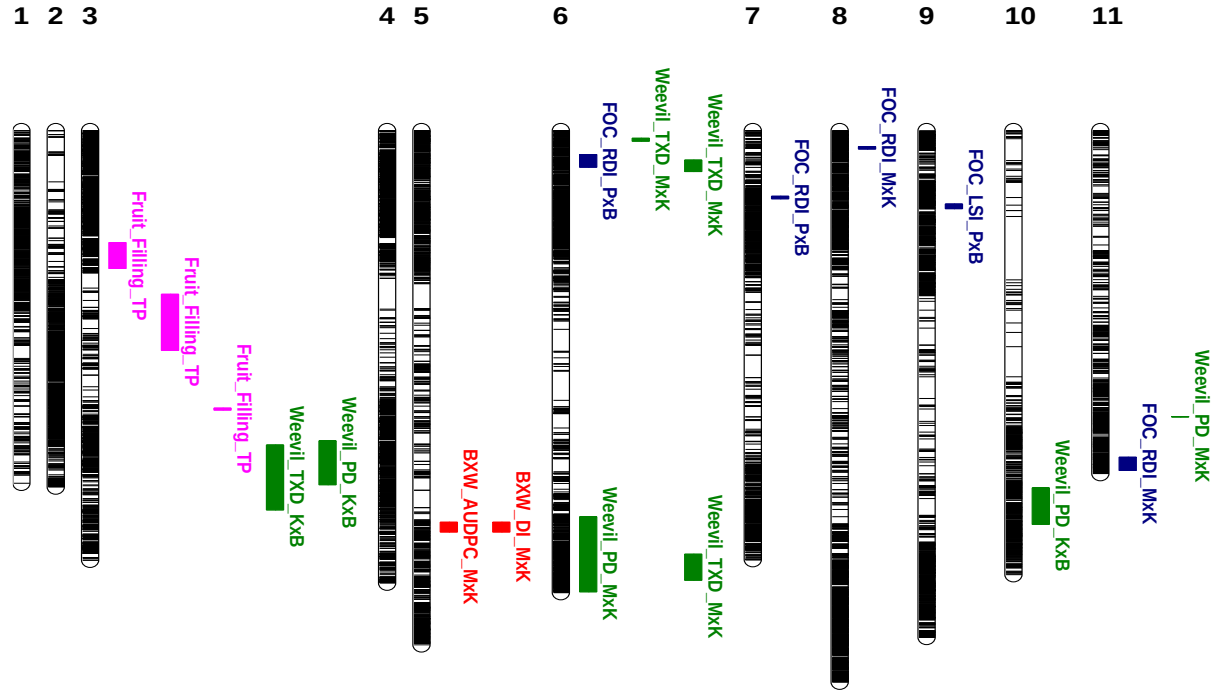

Figure 2. All the QTL mapped in this study shown on the DH Pahang physical map. The numbers represent the 11 banana chromosomes. The significant loci depend on the traits and are from different genetic backgrounds. Weevil QTL: PD: peripheral damage, TXD: total cross-section damage; Foc QTL: RDI: Rhizome discolouration index, LSI: leaf symptom index; BXW QTL: AUDPC: Area under disease progressive curve, DI: disease index; populations: MxK: 'Monyet' x' Kokopo', KxB: 'Kasaska' x 'Borneo'; PXB: 'Paliama' x 'Borneo'; TP: training population

2. Sequencing and genome assembly of diploid *Mchare* cv. Huti White

The analysis of the primary genome assembly of *Mchare* produced in collaboration with IEB and Dovetail Genomics has identified chimeric scaffolds. The scaffolds most probably comprise sequences from different subspecies of *M. acuminata*, which were the progenitors of the sequenced clone, or resulted from the introgressions into *Mchare* genome from other *Musa* species. To identify putative parents, the ITS1-5.8S-ITS2 sequence region from *Mchare* clones and subspecies of *M. acuminata* was amplified, cloned, sequenced by Sanger technology and used for phylogenetic analysis. However, this approach failed to identify unambiguously the parental sub-genomes.

To generate a high quality genome reference of cv Huti White, which seems to contain two very closely related sub-genomes, or introgressions, genomic DNA of two hybrids between Huti White x Calcutta 4 (obtained from IITA-Arusha) were sequenced. Two hybrid clones were selected based on SSR genotyping profiles and sequenced by Illumina technology (2 x 250bp paired-end reads), which resulted in 40x genome coverage. DNA sequence reads were trimmed based on quality and assembled using Meraculous assembler software to obtain primary

genome assembly with N50 of 23.3 kb. However, the sequence data of Huti White obtained initially and the new genome sequence of *M. acuminata* Calcutta 4 did not lead to an unambiguous split of the initial genome assemblies of Huti White.

3. Assessment of Bionano application for analysis of structural variations in *Musa* species and subspecies

Bionano optical mapping using DSL/Saphyre technology resulted in creation of optical maps corresponding to the length of whole chromosomes or at least chromosome arms in triploid clone 3Hand Planty (plantain) as well as in Huti White (*Mchare*). Comparisons of two different A genome-specific optical maps (*banksii*-like A-subgenome of plantain and other A-genome representing *Mchare* banana) gave promising results showing a potential of Bionano mapping for the analysis of genome rearrangements at the whole genome level.

4. Identification of chromosome of translocations in *Musa* using oligo painting FISH

Chromosome structural variation can lead to irregular chromosomal pairing in hybrids, which lowers fertility and results in aberrant chromosome numbers in progenies. Thus, the knowledge of large chromosome structural alterations is critical to select appropriate breeding parents for banana improvement programs. To identify chromosomal translocations in wild *Musa* species and edible clones, a protocol was optimized for oligo painting FISH. Specific painting probes were designed for individual chromosome arms using the reference genome assembly of *M. acuminata* DH Pahang version 2. Fluorescently labelled probes were hybridized on mitotic metaphase spreads of selected edible banana cultivars, including *Mchare* and their putative progenitors. This work revealed relatively high numbers of translocated chromosome regions in various *Musa* accessions, including representatives of *M. acuminata, M. balbisiana*, and triploid banana clones Cavendish, Gros Michel, Plantain, and Highland bananas.

5. Leaf archiving

Leaf archiving has been a continuous effort, with an emphasis on the material in newly established early evaluation trials (EETs). The genotypes archived so far total 6,364 genotypes, with the majority coming from Sendusu (Table). The material at -80ºC was freeze-dried and kept in air-tight containers, supplemented with silica gel to avoid the rehydration of the leaves. An experiment was carried out to see if the freeze-dried material could yield enough DNA for genotyping. It was noted that freeze-drying yielded slightly less DNA (in concentration) than fresh cigar leaf material, but the quantity and quality obtained was suitable for genotyping with DArTSeq® (used as a reference).

Table 13. Summary of genotypes whose leaf samples have been archived

Sn	Type of material	Location	Number of genotypes
1	NARITA Hybrids	Sendusu	23
2	EET 22	Sendusu	259
3	PYT	Sendusu	97
4	HETEROSIS	Sendusu	31
5	2X, 3X, 4X Parents	Sendusu	16
6	Germplasm collection	Sendusu	72
7	PITAs & BITAs	Sendusu	36
8	Calcutta 4 x Zebrina GF hybrids	Sendusu	157
9	Black Sigatoka experiment	Sendusu	9
10	Calcutta 4 x P. Lilin hybrids	Sendusu	344
12	Training population	Sendusu	231
13	EET 24	Sendusu	141
14	EET 25	Sendusu	157
15	EET 26	Sendusu	372
16	EET 27	Sendusu	444
17	EET 29	Sendusu	301
18	EET 30	Sendusu	503
19	EET 31	Sendusu	352
20	EET 23	Sendusu	711
21	EET 13	Kawanda	255
22	EET 15	Kawanda	164
23	EET 32	Sendusu	646
24	EET 33	Sendusu	675
25	EET_Mchare	Arusha	15
26	GWAS	Arusha	103
27	Paliama x Borneo hybrids	Arusha	250
	Total		**6364**

Primary outcomes 4-5-6-7: WP 4 on Regional Testing and end-user evaluation

Scope and approach: Empowering End-user Evaluation: System for better tailoring breeding products and increasing adoption of new cultivars through end-user feedback systems and participatory evaluation of improved banana germplasm.

The goal of this work package is geared towards improving the acceptance rate of high yielding NARITA hybrids by local farmers.

Therefore, the objectives are (reworded from the submitted project document):

1. Identify the best hybrids according to consumer evaluations;
2. Conduct sensory evaluations by consumers: provide qualitative feedback to breeders on taste and other organoleptic features, as well as processing potential;
3. Determine farmer preferred traits: towards modifying the breeding strategy to maximize acceptability of hybrids;
4. Establish a baseline study: to facilitate future release of hybrids;
5. Conduct multi-location trials in a range of target end-user environments: to select the best hybrids per environment, better understand G x E, collate information for rapid release;
6. To evaluate at least 20 of the 27 NARITA hybrids.

This is organized in four primary outcomes

Primary outcome 4. Breeders have a better understanding of traits of importance to end users and use this to orientate breeding strategies and early selection processes.

Primary outcome 5. Simplified, standardized protocol and tools for trial design and implementation, data collection and sharing implemented by all partners allowing meta-analyses across sites.

Primary outcome 6. Farmers participating in selection of new hybrids, with feedback driving changes to strategy and selection processes of breeding programs to improve tailoring of future improved hybrid

Primary outcome 7. Farmers across Uganda and Tanzania and beyond growing their preferred NARITA cultivars, alongside local cultivars.

Primary outcome 4. This serves objectives 1-4.

Qualitative data analysis of 23 focus group discussions (FGD) in Uganda and Tanzania showed that banana is by far the most important crop (Fig. 15) followed by maize, beans, coffee, sweet potatoes, cassava, groundnuts, vegetables (including leafy greens) and avocado. There are differences between Tanzania and Uganda, with Tanzania having higher scores for banana, maize and coffee.

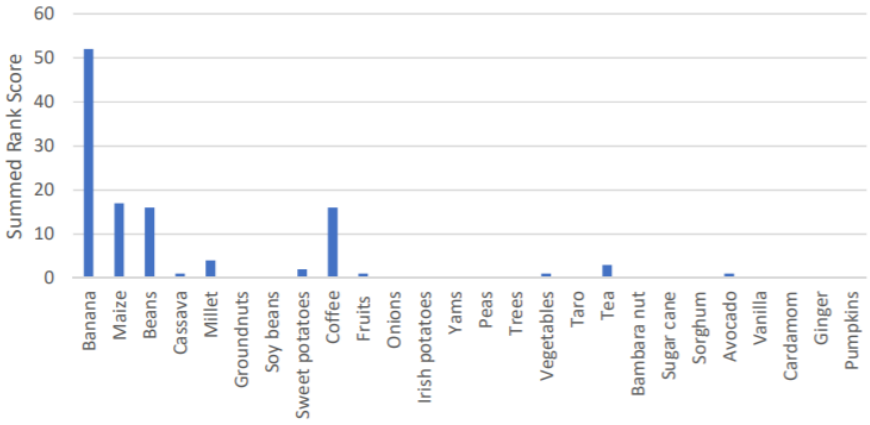

Figure 15. Crop importance in Uganda and Tanzania (the highest possible score for any crop can be 57)

Interestingly banana is considered a more important crop for both women and men in Tanzania (Fig. 16 and 17).

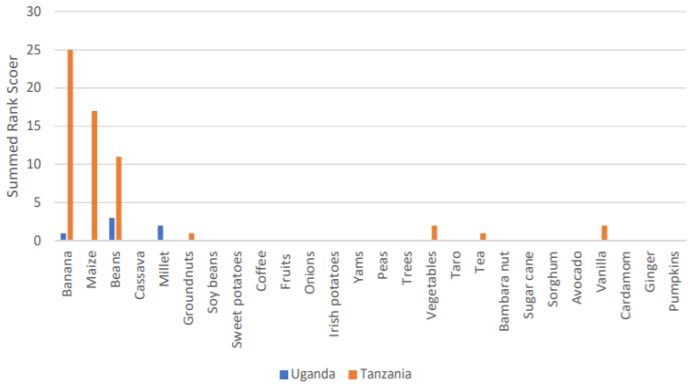

Figure 16. Crop importance for women per country

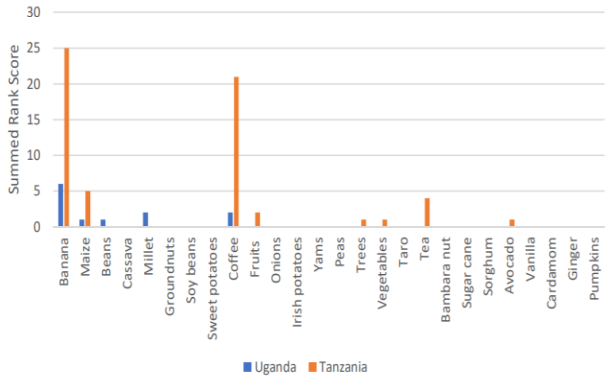

Figure 17. Crop importance for men per country

Field activities by gender varied by crop. For banana, activities vary by gender (Fig. 18), with women more engaged in selling and irrigation, while men lead the work on cultivation, planting, land preparation, soil and plant maintenance and pruning.

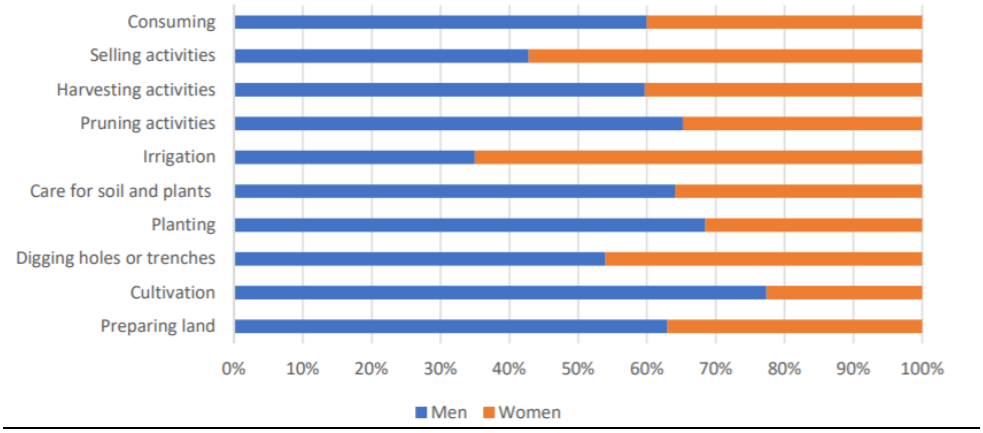

Figure 18. Banana activities by gender

The data from the intra-household survey are being cleaned, processed and coded, and will be available through a CGIAR shared on-line space. A literature review on gender and trait preferences for banana cultivation and use in SSA was conducted. This has enabled a better understanding of trait preferences of stakeholders across the banana value chain, which will facilitate the selection and adoption of new cultivars. Of 44 publications reviewed, only four reported gender-disaggregated trait preferences, highlighting a significant gap in this area. The review found that banana farmers, irrespective of gender, value similar traits that are related to production constraints, income enhancement, consumption, and cultural or ritual uses. Farmers (as producers, processors and consumers) tend to prefer traditional cultivars because of consumption attributes, compared with improved agronomic or host resistance characteristics. Consequently, new improved cultivars need to reflect local taste preferences. Potential differences between trait preferences of farmers and other actors in the value chain should also be taken into account to enhance marketing potential.

Trait preferences were investigated to support the successful development and adoption of improved banana cultivars, which are summarized in “Post-harvest use of banana in Uganda and Tanzania: Product characteristics and cultivar preferences by male and female farmers” (https://hdl.handle.net/10568/106275). Qualitative data from the 23 FGDs conducted in 6 districts (Mbarara and Luwero in Uganda; and Bukoba, Meru, Moshi and Rungwe in Tanzania) show that farmers process banana into a range of products that includes staple food, such as steamed matooke, mbalaga and machalari, as well as beverages and snacks. Different cultivars may be preferentially used depending on the end products. Interestingly, cultivar preferences for specific products and product utilization patterns are similar for men and women farmers. Differences only occurred when men and women described reasons why products are important or preferred, and in the roles of men, women and children in the preparation and processing from different banana types. Both men and women mentioned socio-economic, cultural, consumption and health-related attributes of the different products. However, women (73%) specifically highlighted health benefits attached to women, for example after giving birth, during pregnancy, during lactation and for children; compared to 44% for men. Men pointed out more the socio-cultural importance of the products and provided detailed descriptions on market-related traits and the corresponding value chain actors (88.9%). A myriad of factors that include variety of products, cultivar attributes before processing, characteristics of the processed product and location differences are identified, highlighting the need for further assessments and physiochemical characterization of the traits to gauge the feasibility of considering this valuable information in the product development process. Descriptors for products and cultivar attributes were at times quite generalized and lack detail (e.g. ‘tastes nice’, ‘good flavor’, ‘long shelf life’). Farmers have tacit knowledge of these attributes and consequently, researchers need to determine better mechanisms to extract this knowledge. There is a need to collect data on trait prioritization and conduct biochemical analyses to ensure that the information collected is well defined, comprehensive, quantifiable and measurable to effectively guide the breeding pipeline.

Although we expected a list of desired traits only, both men and women banana farmers (based on the individual intra household survey and FGD data) in the 5 sites (Maruku, Mitalalula, TACRI, Kawanda and Mbarara) showed that they evaluate cultivars by visual and other traits, as either desired or non-desired. Most important desired visual traits are large bunch size, fruits and hand size, moderate suckering, plant height like matooke. Occasionally, there was mention of pest and disease resistance and the desire for many leaves. Resemblance to local *Matooke*, early maturity, easiness to peel, good taste, yellow color, short cooking time, other uses (leaves, rope, animal feed, etc.) and robustness were other important traits. In general, traits preferred by women reflect those of men. Negative traits are often the reverse of the desired traits. In addition FGD mentioned that desired cultivars need less manure, should not have a black or white pseudostem, should not be bitter, should not look like Bogoja (Gros Michel) or Gonja (plantain), and should not have early drying leaves.

Sensory evaluations of NARITAs at the 5 sites in Uganda and Tanzania led to the selection of NARITAs 17, 14, 4, 23, 24, 12, 18, 7 to move forward with. NARITAs that need further evaluation include NARITA 20, 26, 2 and 22. Table 14 summarizes the information for Uganda and Tanzania, and current information on resistance to Foc Race 1 and shows that overall acceptability from laboratory texture analysis reflects farmer preferences. NARITA 12 was recently dropped due to its susceptibility to Armillaria disease.

Table 14. Cultivars proposed for the on-farm trials in Tanzania and Uganda

Uganda	**Motivational traits from farmers**	**Tanzania**	**Motivational traits from farmers**	**Foc race 1 resistance (screenhouse trials in Kawanda; source PhD Privat Ndayihanzamaso)**	**Black Sigatoka (source PhD Janet Kimunye)**	**remarks**	**Overall acceptability (NARO lab texture analysis)** *extreme approval (5) to extreme disapproval (1).*
NARITA 17	• Overall ranked best by farmers with score above 4 out of 5 • Good yellow color, among a key trait for good *Matooke*, the common food in Uganda • Provide big bunch size (preferred by farmers for income earning)	NARITA 14	• Has good taste; less astringent • Provides reasonable bunch (medium to big)	NARITA 17: R NARITA 14: R	NARITA 17: R NARITA 14: R		NARITA 17: 4.6 NARITA 14: 3.4
NARITA 4	• Huge hand clusters, preferred farmer trait for business especially • Made deep yellow to 'golden' *Matooke*	NARITA 4	• Big bunch with many hands and big fingers • Provides good food when cooked • Attractive appearance	NARITA 4: R	NARITA 4: R		NARITA 4: 3.9
NARITA 24	N/A	NARITA 23	• Provide good food when cooked • Has big bunch, hence good for business • Mature early	No data	NARITA 24: R NARITA 23: R		NARITA 24: 4 NARITA 23: 3.5
NARITA 12	N/A	NARITA 12	• Provides good food when cooked • Aroma is close to local • Short maturity time	NARITA 12: intermediate	NARITA 12: Intermediate	Armillaria root rot at Lyamungo site (M. Shemale) DO NOT further TEST	NARITA 12: 3
NARITA 18	• Provide reasonable bunch (medium to big) • Good taste, close to local cultivars	NARITA 18	• Made the best boiled fingers during sensory evaluation	No data	NARITA 18: S		NARITA 18: 3.7
NARITA 7 (hybrid check)	N/A	NARITA 7	N/A	NARITA 7: R	NARITA 7: R		NARITA 7: 4.3
		NARITA 20	N/A	NARITA 20: R	NARITA 20: R		NARITA 20: 2
		NARITA 26	N/A	NARITA 26: R	NARITA 26: R		NARITA 26: no info
		NARITA 2	N/A	No data	NARITA 2: R		NARITA 2: 3.3
		NARITA 22	N/A	NARITA 22: R	NARITA 22: R		NARITA 22: no info

Breeding initiatives seeking to improve food security by developing and introducing productive, pest- and disease-resistant EAHB hybrids benefit from rigorous characterization within the target region for release. A Characterization of Target Population of Environments for East African Highland Banana Using a Multi-criteria Decision Method (MCDM) study was undertaken and embedded in a Land Evaluation framework in target regions, which were subdivided into 14 sub-environments with a large degree of internal homogeneity and inter-sub-environment heterogeneity. Suitability was predicted, expressed as a distance to a 10-dimensional ideal point, and predicted performance values, expressed as predicted bunch weight. Along with the predicted bunch weight, the most limiting variable in each spatial unit or pixel was also calculated. Making use of actual and predicted bunch weights, bunch weight gaps were calculated, and suitability predictions were compared to production figures obtained from public censuses and surveys. Differences in predictions and the most limiting variables between sub-environments in Ugandan regions primarily reflected literature. Suitability and performance predictions did not always consistently correspond to the reality,

represented by both estimated bunch weights and production figures. Where inaccuracies occurred, socio-economic, pest- and disease-related and management factors were thought to be missing explanatory variables, and in several cases were identified as such in the literature.

Primary outcome 5. This serves objectives 5-6.

The trial designs of on-station trials have been uploaded to MusaBase (https://musabase.org/), with uploading of data into MusaBase continuing.

A draft generic protocol for on-farm trial implementation was developed in line with the goal to empower the farmers to contribute to the effective and efficient on-farm testing of selected NARITAs and checks through a citizen science approach in Uganda. Crowd sourcing was to be implemented through the triadic comparison of technologies (tricot). The tricot protocol involves the use of the ClimMob platform to randomise the selected cultivars and checks into tricot packages of three cultivars per location/farmer from which observations and data will be collected. In June 2019, members from Bioversity and NARO attended ClimMob and Data management workshops to enhance skills in implementing the tricot design and data management, although TARI could not attend due to travel issues. A draft protocol for the introductory meetings with farmers was prepared and shared with NARO for discussion. On-farm trial facilitator and data collectors (OFTDC) were to be hired in each of the three districts of Uganda and trained according to the proposed OFTDC training protocol that was shared with NARO. Their role would have been to assist with data collection and plot management. The on-farm trial implementation plan involved conducting a consultative stakeholder workshop in the production and utilization of bananas, at Kawanda. The participatory workshop included participants from the selected districts for the on-farm trials (Luwero, Mbarara and Kamuli), and was followed by farmer introductory meetings in each district, and land preparation and trials planting. Farmers were introduced to the on-farm trials plan and an implementation outline, which was mutually developed and led to the identification of suitable host farmers. The host farmers prepared the experimental plots under the guidance of NARO. However, although planting plans were at an advanced stage, concerns from project partners regarding the use of sucker planting material and the unavailability of on-station agronomic performance and disease data to support the clones selected for on-farm, the tricot design was halted. In Phase II, field trials are designated to be led by NARO, both in Uganda and Tanzania. Tissue culture planting materials for Uganda on-farm trials were delayed but eventually imported from Nigeria and hardened at Kawanda. Once the decision to halt the tricot design was made, NARO proceeded to plant the cultivars selected for on-farm evaluation at Kawanda, in order to maintain the materials for reference purposes and future use.

On station field data collection was discontinued in June 2019 and good progress has since been made in harmonizing data columns from SMAP data files into one file for export from SMAP. The SMAP consultancy team successfully harmonized the database columns and combined the database files into a single file. The updated SMAP database file is currently being cleaned using R, processed and analyzed before data export. Some challenges have been encountered, e.g. several thousands of rows with missing ratoon information, depending on traits being assessed, and around 26,522 rows of duplicate records, which need to be addressed before exporting to MusaBase. The combined SMAP file, however, is huge with 95,125 rows and 509 columns of data. Direct smooth Interoperability of SMAP and MusaBase platforms is still not attained and hence copy and paste procedures will be employed during data upload to MusaBase. The bulk of the work related to data manipulation, analyses and paper/report writing from the baseline survey data continues and forms part of an ongoing PhD by Noel Madalla.

Primary outcome 6. This serves objectives 5-6.

Agronomic data collection of third and fourth cycles was conducted in the on-station trials, although the fourth cycle was not completed. All agronomic data are now stored and available through a shared online space - SMAP. Agronomic data collection was completed in June 2019 and the management of the trials handed over to the partners.

Sensory evaluation data was collected from all the 5 on-station sites, analyzed and recommendations made for which cultivar subsets to take on-farm (Table 14). The recommended list for Tanzania is longer in order to cater for the wider agroecological diversity, and consumer differences between target regions.

Preference ranking (PR) exercises with farmers were completed in all the 5 regions in Tanzania and Uganda (Table 15) in order to obtain a more comprehensive understanding of traits that are perceived as important by both male and female farmers when selecting banana cultivars. During the PR exercise, two types of data were generated: (1) a quantitative preference score for each cultivar; and (2) a qualitative assessment of traits that farmers look for (or avoid) when selecting banana cultivars.

In Tanzania, some of the NARITAs outperformed the local cultivars and were preferred by farmers in Mitalula and TaCRI, while in Maruku the local cultivars were preferred over the NARITAs. In Uganda, the local cultivars were preferred over the NARITAs in Kawanda but in Mbarara some NARITAs outperformed the local checks. Across the 5 sites, NARITA 23 ranked the highest, followed by NARITA 2 and NARITA 12, which are all food types with average to very high yields recorded from the Sendusu preliminary trial. The least preferred cultivar across sites in Uganda was NARITA 17, followed by NARITA 20, NARITA 19 and NARITA 10. NARITA 20 and NARITA 19 was relatively low yielding in the Sendusu preliminary trial, and NARITA 10 is a juice type. Several important differences in farmer preferences were observed between sites, while preferences between men and women were quite similar. Although NARITA 17 was one of the least ranked for plant preference, it actually ranked highest for sensory evaluations. However, it appears that NARITA 17 was ranked low during preference ranking exercises because it hardly had bunches ready for harvesting at the time. This highlights one of the challenges when conducting farmer varietal preference exercises when different genotypes may be at different growth stages. This needs to be taken into account in forthcoming trials, with every effort made to ensure all planting material is at similar ages across genotypes. Statistical analyses continues.

Table 15. Preference scores for NARITAs in Tanzania (Maruku, Mitalula and TaCRI) and in Uganda (Kawanda and Mbarara) on-station trials, disaggregated by gender.

Variety	Type	Maruku combined session 1 and 2			Mitalula			TaCRI			Kawanda			Mbarara			Average over 5 sites		
		Preference score			Preference score			Preference score			Preference score			Preference score			Preference score		
		Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
NARITA 2	Food	57	60	59	84	68	79	79	71	75	70	70	70	92	81	87	76	72	75
NARITA 4	Food	39	46	43	70	54	65	88	73	81	60	79	76	94	85	90	70	65	68
NARITA 6	Food				40	37	39	86	79	82	28	28	28	56	57	57	52	50	51
NARITA 7	Food	78	83	81	30	62	44	51	51	51	85	77	79	23	30	27	53	61	57
NARITA 8	Juice	43	37	40	84	69	79	38	34	36	79	72	73	28	32	30	54	49	52
[6]NARITA 9	Juice				43	43	43	92	85	88				20	25	22	51	49	50
NARITA 10	Juice	44	32	37	57	58	57	23	30	27	50	55	54	26	28	27	40	39	39
NARITA 11	Food	71	73	72	86	79	83	90	86	88	55	68	66	31	27	29	67	65	66
NARITA 12	Food	89	84	86	81	63	75	59	77	68	65	82	78	48	54	51	68	67	68
NARITA 13	Juice	57	59	58	61	44	56	66	55	61	34	34	34	66	67	66	57	52	55
NARITA 14	Food	50	53	52	36	25	32	42	50	46	50	46	46	28	37	32	41	41	42
NARITA 15	Food	58	48	52	22	24	23	71	81	76	85	80	81	68	75	71	61	61	61
NARITA 16	Juice										60	76	73	24	35	29	42	42	42
NARITA 17	Food										26	27	27				26	26	26
NARITA 18	Food				31	35	32	62	60	61	73	66	68	49	58	53	54	54	54
NARITA 19	Food	38	33	35	31	43	35	40	52	46							36	43	39
NARITA 20	Food	36	31	33	54	34	48	20	27	24							37	31	35
NARITA 21	Juice	46	39	42	67	71	68	61	72	66	45	37	39	37	40	38	51	53	52
NARITA 22	Food	89	98	94	60	67	62	25	55	40							58	73	65
NARITA 23	Food	91	87	89	77	70	75	70	66	68	80	69	71	65	64	64	76	73	75
NARITA 24	Food										58	78	74	35	37	36	46	46	46
[7]NARITA 25	Food				74	64	71	52	43	47	32	39	37				53	46	50
NARITA 26	Food	83	56	68	65	50	60	40	46	43							63	51	57
NARITA 27	Food	51	46	48	43	42	43	94	80	87							63	56	59
MBWAZIRUME	Food	93	97	95	49	58	52	84	93	88	88	92	91	64	63	64	75	80	77
NSHAKALA	Food	92	95	93													92	95	93
ENYOYA	Food	95	96	95													95	96	95
NDIZI NG'OMBE	Food							43	61	52							43	61	52
NDIZI UGANDA	Food				32	32	32	22	51	37							27	42	34
KISANSA	Food										65	69	68	75	77	76	70	70	70
NAKITEMBE	Food										93	92	92	60	59	60	76	76	76

[6] It was identified as NARITA 17 instead of NARITA 9 during the mislabeling exercise

[7] It was identified as among the genotypes planted in Kawanda and therefore included during the preference ranking exercise

Primary outcome 7. **This serves objectives 4-6.**

A literature review on meta-analysis of multi-environment crop cultivar trials data and data synthesis for crop cultivar recommendations was conducted by compiling datasets and conducting preliminary tests applying the Plackett- Luce Model. The test data used in the Plackett-Luce Model was retrieved from AgTrials and MusaBase. More data for modelling was obtained from CIMMYT maize trials and additional banana on-station trials data towards developing and testing the model. An example of the key output expected is the relative worth of a cultivar plotted against cultivars using data from 186 International Musa Testing Phase 1(IMTP-1) and International Musa Testing Phase 2 (IMTP-2) banana trials. From AgTrials, all data from these 186 trials were downloaded and being processed. From MusaBase, the downloaded data is a subset of trials evaluating NARITAs for response to black Sigatoka in the 5 on-site locations.

Primary outcomes 8: WP 5 Create a banana breeding database

Scope and Approach: Harnessing Data: Driving improved efficiency of breeding systems and enhanced synergy in national, regional and global partnership through an open-source database and tool box for banana breeders and researchers.

The goal of this work package is to create a banana breeding database to improve the efficiency in banana breeding.

The objective is to provide project partners and Musa researchers and breeders a virtual hub for information exchange, R&D collaboration and enhanced adoption of new hybrids.

Prior to BBB, a colossal amount of banana breeding research data was recorded onto paper before transcribing onto computer for analysis etc, often with errors accruing and traceability sometimes suspect. Further to BBB we now have FieldBook data collection, linked to BTracT, which accurately records data and documents all stages of the breeding program, synchronizing it automatically with MusaBase to archive all information and data and make it publicly available. The wish list import-export function between MusaBase and BTracT was developed in Year 5 and together, these electronic tools and databases have become continuously more integrated during the project through upgrades, improving the synchrony and resolving issues and glitches.

The banana tracking tool was developed to store and retrieve BTracT data (https://musabase.org/breeders/odk) and enable all breeding products to be accurately traced back to the original pollination events. A new cross experiment page and 5,000 crosses were added during Year 5: https://musabase.org/breeders/trial/435?format=. To date, 122 people are registered on https://musabase.org/search/people, 72 of whom are independent of BBB. This transformation in recording, documenting, storing and tracing data is monumental for banana breeding.

In Year 5 training workshops continued, with a series of database and BTracT trainings conducted at IITA, Uganda and EMBRAPA Cruz Das Almas, Brazil, to unite the data of IITA, NARO and EMBRAPA. Also trainings were organized for the BTracT developer of IITA, for two

MusaBase contact persons for uploading data based at IITA-Nigeria for plantain and IITA-Sendusu for *Matooke* with training on (1) genotyping file data handling: theory and practice, (2) database update: tools of interest for the banana program, including tissue sampling, analysis tools such as mixed model and selection índex, (3) field implementation: barcode label purchase, improvement of BTracT-MusaBase interfacing and (4) leveraging from other data manager's experiences like Cassavabase.

The majority of trials have been uploaded, with all new trials now designed using MusaBase and downloaded onto a tablet for data collection using a new field map viewer, which was developed: https://solgenomics.github.io/BrAPI-HeatMap/example.html

With reference to the trait ontology, field management stages were integrated and a first set of 7 quality traits for carotenoids were added (https://musabase.org/tools/composea). Thus the system is now capable of receiving more traits, including from the RTB food work package (https://musabase.org/search/traits).

The germplasm collection morphological image storage was further enriched with plant morphological descriptors and photos for germplasm present in Arusha and Sendusu (https://musabase.org/image/view/6299/) and matched to ITC codes. In addition, several videos documenting banana flowering are now available on MusaBase (ftp://ftp.musabase.org/video/flowering/). A new image upload interface was also released (https://musabase.org/tools/image_analysis), which includes analysis methods that have been adapted from its original design for cassava.

All the accessions currently being used, including the mapping populations, have been uploaded onto MusaBase along with the pedigree information. All the IITA germplasm data from Sendusu has also been checked, which identified 500 duplicates that have now been merged. A first large set of data from the NARITA Kawanda trial (part of the multilocational trial) was uploaded but the large data set from the 5 field testing sites is still being cleaned and checked before it can be uploaded. Data from the IITA plantain breeding program, Nigeria were also uploaded onto MusaBase.

Whole genome resequencing data was uploaded, matched to corresponding ITC codes and made available (ftp://ftp.musabase.org/musaWGS/ ftp://ftp.musabase.org/musaWGS/). Similarly, a new marker storage has been developed allowing storage of vcf files into MusaBase. The GBS data from the genomic selection research was curated, linked with relevant pedigrees and included in MusaBase using the new genotyping storage system (https://musabase.org/breeders_toolbox/protocol/1https://musabase.org/breeders_toolbox/protocol/1 and ftp://ftp.musabase.org/musaGBS/Nyines2018/). This new storage system allows complete information storage and further use in new tools (pedigree viewer).

MusaBase also now stores data from other IITA research activities such as the metagenomic study on Fusarium: ftp://ftp.musabase.org/musaRNAseq/kaushal_2019/

2. Geographic Areas to Be Served

Provide the final list of countries and sub-regions/states that have benefitted from this work and associated dollar amounts. If areas to be served include the United States, indicate city and state. Add more rows as needed. More information about Geographic Areas to Be Served can be found here.

Location	Foundation Funding (U.S.$)
Uganda and Tanzania (by the Project directly)	6,867,431
W. Africa Breeding Programs (Nigeria/Ghana)	1,664,832
Producing Countries in the region (through Spillover)	3,662,630
All banana producing countries Worldwide (through Spillover)	1,678,705

3. Geographic Location of Work

Provide the final list of countries and sub-regions/states where this work has been performed and associated dollar amounts. If location of work includes the United States, indicate city and state. Add more rows as needed. More information about Geographic Location of Work can be found here.

Location	Foundation Funding (U.S.$)
India	300,000
Brazil	429,568
Belgium, South Africa, Czech Republic	688,622
Uganda	2,295,205
Tanzania	346,333

Malaysia, France	929,613
USA, Kannapolis and New York	757,914
Australia	705,378
IITA Uganda	5,936,773
IITA Tanzania	1,484,193

4. Lessons Learned

Describe the top one to three takeaways or lessons learned from this project.

1. The banana breeding pipeline needs to be comprised of a seamless, smooth and fluid transition between stages that should flow as a liquid through the pipe itself, with any leaks to be addressed immediately. There is still much experience out there in the private sector than needs to be brought into the banana breeding pipeline to make it even more efficient.
2. Breeding bananas can be much faster than is conventionally thought but requires the harmonious interaction and cooperation between various experts; a strong network is key. A global network, linking partners from broad geographic and technical fields is challenging! But was demonstrated to work and brings great dividend by fostering and feeding this network.
3. Creation of a central accessible database of all information for the target crop (Musa) and linked to the current and ongoing breeding programs is monumental in terms of preserving information and knowledge, making it openly available and using programs that use the information towards designing breeding pipelines and products is more than we ever imagined.

5. Feedback for the Foundation

Provide one to three ways the foundation successfully enabled your work during this project. Provide one to three ways the foundation can improve.

Ways the foundation has successfully enabled our work:

1. The regular and consistent personal contact with and intervention by the BMGF Project Officer throughout the project duration has helped create a smooth implementation of the project, guiding its course with professional advice, personal/professional contacts and enabling immediate feedback on queries, requests and for advice.
2. Suggestions for membership on the SAG, in particular of persons from industry, has been very helpful to recruiting useful expertise onto the SAG.
3. The (quite considerable) flexibility provided by BMGF on the budget shifts and reforecasting enables implementation of the project in response to changing conditions or situations in real life, providing adaptability to shifts in circumstances. This has been very helpful.

Ways in which the foundation can improve:

1. Delays in approval of the reports, both technical and financial perpetually delayed the disbursement of fund to IITA and consequently to partners. Some partners have limited credit facilities and so delays create difficulties for the smooth continuation of activities. While IITA is additionally responsible for some of this delay, the system in general creates gaps and delays to implementation.
2. The approval of the financial report in some years was delayed due to the need to provide explanations for adjustments made in previous years, which were approved and accepted, but which then were needed to be provided again for shifts that occurred as a result of earlier year shifts. Explanations provided for earlier years were questioned later. A better streamlining of the financial reporting for the current year, based on the fact that earlier years were approved and signed off, would help in preventing delays.
3. We believe that it is very important to plan activities and finances as much as possible. We also agree that a deviation of more than 10% should be avoided. However, the foundation should also appreciate that more than 10% deviation is sometimes necessary due to shifting conditions and unforeseen circumstances.
4. Too many advisors providing advice from different perspectives, which can be conflicting. BMGF should provide a screening or censorship to enable single channel through the BMGF Project Leader

6. Global Access and Intellectual Property

If your funding agreement is subject to Intellectual Property Reporting, please click the following link to complete an Intellectual Property (IP) Report.

If not, please acknowledge by typing "N/A": N/A

To delegate permissions to another member of your project team or for any questions regarding the Intellectual Property Report, please contact GlobalAccess@gatesfoundation.org.

7. Regulated Activities

Do you represent that all Regulated Activities[1] related to your project are in compliance with all applicable safety, regulatory, ethical and legal requirements? Please mark with an "X":

__X__ N/A (no Regulated Activities in project)

_____ Yes

_____ No (if no, please explain below)

1 Regulated Activities include but are not limited to: clinical trials; research involving human subjects; provision of diagnostic, prophylactic, medical or health services; experimental medicine; the use of human tissue, animals, radioactive isotopes, pathogenic organisms, genetically modified organisms, recombinant nucleic acids, Select Agents or Toxins (www.selectagents.gov), Dual Use technology (http://export.gov/regulation/eg_main_018229.asp), or any substance, organism, or material that is toxic or hazardous; as well as the approvals, records, data, specimens, and materials related to any of the forgoing.

8. Subgrants

If your grant agreement (not applicable to contracts) is subject to expenditure responsibility and permits you to make subgrants to organizations that are not U.S. public charities or government agencies/instrumentalities, please complete the Subgrantee Checklist and attach a copy with this progress narrative for each such subgrantee.

Financial Update

The purpose of this section is to help the foundation understand how programmatic performance affects actual expenditures over the life of the investment.

Feel free to reach out to your foundation contact for support with these progress reporting requirements.

Note: Budget template and financial narrative instructions can be found here. If you are using an older version of the budget template, this information could be in a different location in your template.

1. Latest Period Variance:

"Latest period variance" compares expenditures that occurred in the reporting period against the most recent forecast. See "Financial Summary & Reporting" sheet in the foundation budget template for calculated variance (for example, column AD, starting on row 29 for period 1). Note that the allowable variance is defined in your grant agreement.

Latest Period Variance:

The total grant variance is now running at 0%, although there are some posts with over 10% variance, none affected the output of the project. Year 5 has no major deviations that occurred in year 1-4.

Personnel: 27% (in Year 4 21%). Project staff turnover has impacted on costs, with recruitments proving slow due to the inevitable nature of filling positions, national laws and admin procedures. This is the result of especially what happened in Years 1-4.

Capital equipment: -122% (in Year 4: -122%). The upgrading of the breeding facilities using unspent funds accumulated over Years 1-2, which was largely spent in Year 3 but also spilt over into Year 4 due to contract delays.

Consulting: 61% (in Year 4: 60%). The genotyping of the mapping populations and development of the SNPchip has attracted much of the consulting costs, further to approval by BMGF.
Other direct costs: -36% (in Year 4: -28%). A number of previously unanticipated admin charges, increased admin charges and accrued IITA charges that have been delayed have led to increased costs.

2. Sub-awards (if applicable)

This sub-award section provides visibility to an often critical component of the grant spending where the budget template provides limited insight. The total of actual disbursements for this reporting period should equal the actual sub-award expenses reported on the "Financial Summary & Reporting" sheet in the budget template for this reporting period.

Use the table below to provide the detail of all sub-grantee(s) or subcontractor(s).

Organization Name	Actual Disbursement for this Reporting Period (U.S.$)	Total Disbursed from Primary Awardee to Sub to Date (U.S.$)	Total Sub-Awardee Spent to Date (U.S.$)	Total Contracted Amount (U.S.$)

University of Queensland	0.00	705,378.00	705,378.00	705,378.00
ARI-HORTI	30,585.00	346,333.00	346,333.00	346,333.00
NARO	389,782.00	2,240,035.00	2,240,035.00	2,240,035.00
UNIVERSITY OF MALAYA	35,895.00	162,060.00	162,060.00	162,060.00
BIOVERSITY & NRCB	275,401.00	1,072,900.00	1,072,900.00	1,072,900.00
STELLENBOSCH UNIVERSITY	52,932.00	421,497.00	421,497.00	421,497.00
KUL	0.00	178,103.28	178,103.28	178,103.28
SLU	19,375.00	155,000.00	155,000.00	155,000.00
BTI	69,928.00	499,714.00	499,714.00	499,714.00
CROPBIOSCIENCE	0.00	10,728.00	10,728.00	10,728.00
BIOCROPS (UK) LTD.	0.00	5,425.00	5,425.00	5,425.00
FUNARBE/EMBRAPA	52,081.00	429,568.00	429,568.00	429,568.00
WCMC	0.00	43,200.00	43,200.00	43,200.00
UNCC	0.00	25,000.00	25,000.00	25,000.00
IEB	40,000.00	110,000.00	110,000.00	110,000.00
DHMRI	0.00	150,000.00	150,000.00	150,000.00
KUL (SERVICE AGREEMENT)	0.00	26,283.88	26,283.88	26,283.88
EDET OUT & CO.	0.00	1,112.38	1,112.38	1,112.38
Total (ties to budget file(s))	965,979.00	**6,582,337.54**	**6,582,337.54**	**6,582,337.54**

For sub-awards greater than $1M, please provide explanatory detail as requested in the latest and future period sections above.

Note: It is the foundation's discretion to ask for updated sub-award budget files as part of the traditional progress report review process.

3. Other Sources of Support (if applicable):

Other Sources of Support include interest earned, current foreign exchange impacts, and co-funding (in-kind and other contributions).

Other Sources of Support (if applicable): Explain any notable impacts from other sources of support.

RTBFood amount: $12,136 USD in year 1 and year 2

Checklist - As you review your answers to questions in the financial update section, ensure that your report provides the following:

1. Explanation of how project expenditures differed from plan and the implications on programmatic progress to date.
1. Explanation of how future period projections differ from the original budget and previous forecasts, and the implications.
2. Explanation of other sources of support (funds) from other funders, interest earned or converting to non-USD currencies.

Privacy and Non-Confidentiality Notice

The foundation is required by the IRS to publish a list of its grants. We may also provide a general description of our grants and contracts on our web sites, in press releases, and in other marketing materials. Subject to the foundation's Privacy Policy, the foundation may also

share information you provide to us (either orally or in writing) with third parties, including external reviewers, key partners and co-funders. This document is subject to the foundation's Terms of Use.

For Foundation Staff to Complete

Analysis (required if PO assessment differs from grantee/vendor assessment or if there are unexpended funds)

Progress Analysis

Include analysis of significant project variances and key learnings that may inform portfolio discussions for progress against the strategic goals.

Budget and Financial Analysis

Include analysis of unexpended funds or over expenditures. Refer to the Unexpended Grant Funds Policy for options available when recommending how to handle unexpended grant funds, or reach out to your primary contact in GCM.