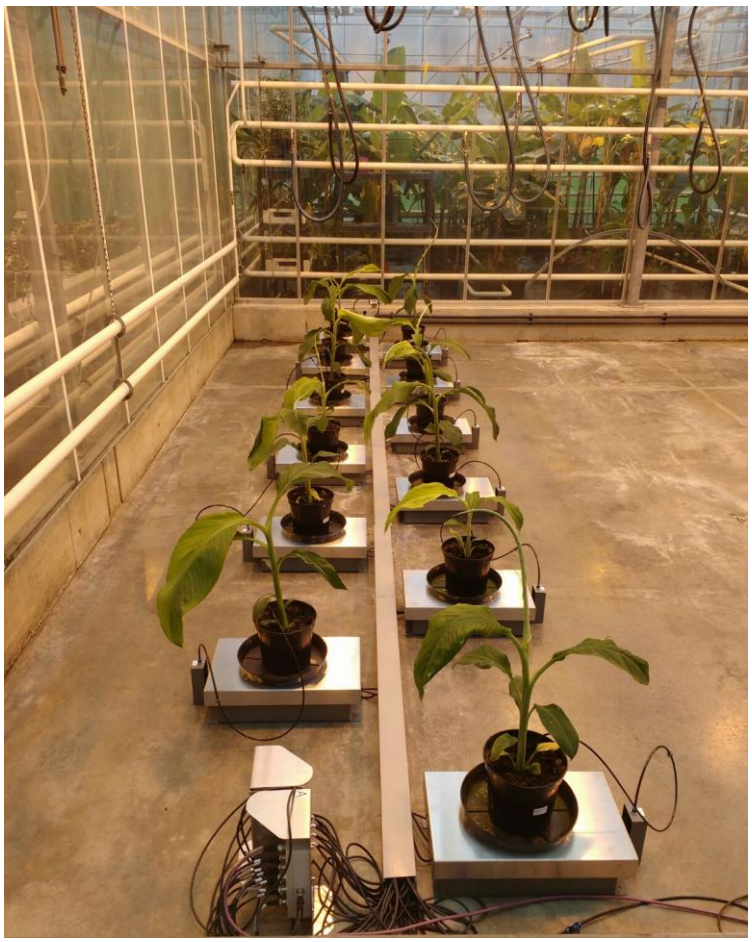
More fruit for food security: developing climate-smart bananas for the African Great Lakes region

TECHNICAL REPORT : Sep 2017 – Sep 2018



Simulating drought in banana plants, S. Carpentier, Bioversity

Executive summary

Banana (*Musa* spp.) is an extremly important crop in the Least Developed Countries, providing a staple food for more than 400 million poor people. Drought stress globally causes estimated yield losses of 37-61 million tonnes annually, worth at least US$13-22 billion, with a third of these losses incurred in Africa.

The overarching goal of the project is to identify drought-tolerant banana cultivars to help close the yield gap in drought-prone areas. This project focus on the African Great Lakes region, which is one of the poorest areas in the world, encompassing Burundi, north-eastern Democratic Republic of Congo (DR Congo), western Kenya, Rwanda, north-western Tanzania, and Uganda. The region is also one of the most important banana-growing areas in the world and is home to a unique group of banana cultivars that are well adapted to this unique colder highland environment.

To support this research, we rely on the Bioversity International's collection of banana diversity that houses germplasm samples representing more than 1000 banana cultivars and 29 *Musa* wild relatives, which together provide a substantial potential genepool to combat drought stress.

As described in the original proposal, the first year of the project was focused mostly on two main Work packages (1 and 4).

- Deploying our high-throughput screening methods to assess the growth potential of the banana cultivars in the ITC collection for the agro-ecological zone of the African Great Lakes region under drought stress (WP1).
- Complementing the phenotyping by molecular analyses to unravel the molecular basis of mechanisms behind banana's drought tolerance, to support future high-throughput discovery of drought-tolerant cultivars and breeding (WP4).

Significant progresses were achieved in the first year that will lead to at least two peer review open access publications next year. Therefore, the technical report follows the structure of the manuscripts that are currently in preparation.

Evaluation of 'the bananaTainer', as a proxy for the African highland climate.

Abstract

Future agriculture is challenged to secure food, feed, fuel and fiber production through sustainable intensification. This requires more effective yield per area also taking into account environmental constraints. Germplasm collections provide valuable genetic resources with potential to augment the resilience of agro ecosystems towards future (unpredictable) climate conditions. An example of such a collection is the world's biggest Musa collection (ITC, Bioversity International) holding over 1500 *in vitro* accessions, hosted at the KU Leuven. The bananaTainer, introduced here, is designed to simulate the climate of the East-African highlands and perform high throughput evaluation of growth performance. The challenges in the highlands are the lower temperatures, a dry season of 2-3 months and lack of nutrient such as potassium and nitrogen. We established a pipeline, growing 504 banana plantlets of 21 cultivars for 6 weeks in a continuously recirculated and aerated nutrient solution in the bananaTainer. This allowed so far to rank the growth of 57 cultivars, representing 631 ITC accessions. The best performing cultivars are 3 AAB cultivars (Uzakan, Silk, Raja), 1 AAA (Lacatan), 1 local Mshare cultivar (Huti Green Bell), 1 ABB cultivar (Simili Radjah), and 1 malaccensis wild accession (Pahang).

Water scarcity was mimicked by addition of 5% Poly ethylene Glycol (PEG), lowering the plant water availability to pF 2.7. This system has been proven to be successful in small scale (van Wesemael et al. 2018, Frontiers in Plant Science, in revision). However, when upscaling to high throughput conditions, the system was vulnerable towards micro-organisms. In the bananaTainer, the added PEG was metabolized by Cupriavidus sp., Pseudomonas putida, and Acinetobacter sp. This depleted the nutrients, rendering them unavailable for the plants. The bananaTainer is standard equipped with a UV-treatment. The UV-treatment was successful to keep micro-organism growth under control in the bananaTainer but unsuitable in combination with PEG. The addition of an alternative extra disinfection was not appropriate as this caused root damage and decreased plant transpiration. We conclude that the bananaTainer and the plant growth model is excellent to simulate the impact of lower temperatures and specific nutrient loss on plant growth. However, Polyethylene glycol is not suitable to simulate drought stress due to microbial degradation. An alternative where drought is simulated in an inert matrix is being evaluated currently.

Introduction

By 2050 the demand of food, feed, fiber, and fuel from agricultural sources is expected to double. Meeting these rising demands requires sustainable intensification of the production: more effective yield on the same area (Foley et al., 2011). However, the current yield increases of the four major crops (maize, rice, wheat, and soybean) are not sufficient to meet the required crop yield increase (~ + 2.4% yearly) (Ray, Mueller, West, & Foley, 2013). Also the further increasing impact of climate variability is estimated to impact worldwide agricultural productivity. Taking this into account, an example pathway towards sustainable intensification is creating augmented resilience in agricultural production systems in order to reduce the risks (Lipper et al., 2014). The impact of abiotic stress on agricultural productivity is more determined by environmental variability rather than the (a)biotic stress factor (Langridge & Reynolds, 2015). Elite, high yielding, cultivar breeding has traditionally aimed for increased productivity, successfully augmenting production, but narrowing genetic diversity, and thus limiting stress plasticity (Dalal, Attia, & Moshelion, 2017; Dwivedi et al., 2016). Climate smart cultivars play a central role in this regard: agriculture should focus on the right cultivar in the appropriate environment, more than applying non-sustainable management. Plant landraces are adapted to a certain agro ecological environment, and are a major source of genetic gain (alleles) towards environmental plasticity. It is of utmost importance that germplasm pools are taken up into climate smart plant breeding and cultivar selection protocols (Dwivedi et al., 2016; Mickelbart, Hasegawa, & Bailey-Serres, 2015). This requires that available (stored, or wild) germplasm collections are explored and characterized by full integration of genomic and plant physiological phenotyping.

Characterizing germplasm in the field risks to be masked by environmental and spatial heterogeneities combined with impracticalities, leaving the scientific question largely unsolved (Negin & Moshelion, 2017). In controlled conditions, using lab or greenhouse models, there is tight control of externalities, and high experimental reproducibility. These pre-field phenotyping experiments make a rough selection of candidate phenotypes focusing on plant performance proxies relevant for the agricultural productivity (yield). Subsequently these candidate phenotypes should be validated for their agricultural relevance in dynamic and physiological trait based field experiments (De Boeck et al., 2015; Junker et al., 2015; Moshelion & Altman, 2015; Poorter et al., 2012; Vanhove, Vermaelen, Panis, Swennen, & Carpentier, 2012).

In the first year of this project, we have set up a large scale, high throughput characterization experiment for banana, a major crop in (sub-) tropical regions (> 144 million tons yearly production). We simulated the climate of the Eastern African highlands. This region is characterized by a dry season of 1-3 months, lower temperatures and lack of nutrients such as nitrogen and potassium (Taulya, van Asten, Leffelaar, & Giller, 2014). In the environment of the highlands, banana requires vast amounts of water (at least 1200 mm / year) and is heavily susceptible to drought: 8 % yield decline per 100 mm water not available for evapotranspiration (Hegde & Srinivas, 1989; van Asten, Fermont, & Taulya, 2011). Breeding new, drought tolerant banana cultivars that are not too much affected by the lower temperatures and can cope with low fertilizer input is a difficult and lengthy process, as such the exploration of the available biodiversity is of real interest (Christelová et al., 2016; Ortiz & Swennen, 2014). The International Transit Center (ITC) of Bioversity International, hosted at the KU Leuven, is the world's largest banana germplasm collection with over 1500 accessions. To increase management efficiency and turnover of the collection both genotypic (Christelová et al., 2016) and phenotypic characterization are crucial. Our phenotypic characterization experiment adds value to the germplasm collection towards a major stressors in the East-African Highland banana cultivation.

Simulating lower temperatures and lack of nutrients is more straightforward in soilless, hydroponics cultivations than simulating drought. Drought simulation implies the establishment of a lowered water

potential, for example by application of osmotic active compounds. Suitable compounds act on the water potential without interference or plant absorption. The most suitable one is the high molar mass polyethylene glycol (PEG, MW>6000 g/mol) (Blum, 2011; Burlyn E Michel & Kaufmann, 1973; Money, 1989; Verslues, Agarwal, Katiyar-Agarwal, Zhu, & Zhu, 2006). PEG in water exists in a very hydrated form through hydration binding on its ether oxygen atoms creating a hydration shell (Burlyn E Michel & Kaufmann, 1973; Steuter, Mozafar, & Goodin, 1981). These hydrogen bonds increase the thermodynamic disorder and lower the free energy of the solution, as such the specific PEG - water interactions cause a sharp increase of osmotic pressure. The osmotic pressure increases quadratically with increasing concentration (B. E. Michel, 1983; Burlyn E Michel & Kaufmann, 1973; Money, 1989; Steuter et al., 1981). This is determined by shape, conformation, and size (molecular weight) of the individual PEG molecules, more than by the PEG concentration (molarity) (Burlyn E Michel & Kaufmann, 1973; Steuter et al., 1981). PEG solutions of higher molecular weight have a higher osmotic pressure at the same concentration (Burlyn E Michel & Kaufmann, 1973; Money, 1989). Plant cell walls are considered impermeable for larger PEG molecules, creating an osmotic potential at the cell wall boundary (cytorhizis) (Verslues, Ober, & Sharp, 1998). This contrasts to other, smaller compounds, like mannitol or sucrose, and salts, like NaCl, which readily penetrate the cell wall creating osmotic pressure at the cell membrane (plasmolysis) and easily enter de cell.

Worldwide, millions of tonnes of these large, water soluble, non-ionic, PEG polymers with a C-C backbone, are produced (Huang et al., 2005). PEG is widely used in for example cosmetics, lubricants, plastics, as anti-freezing agent, as such PEG is ubiquitous in (waste) water. The polymers are highly resistant to normal biodegradation, resulting in their longevity in natural environments (Allen et al., 2002; Huang et al., 2005). In highly specialized conditions such as waste water treatment installations, both aerobic and anaerobic pathways of bacterial PEG-degradation have been found (Bernhard, Eubeler, Zok, & Knepper, 2008; Dwyer & Tiedje, 1986; Kawai, 2002; Marchal, Nicolau, Ballaguet, & Bertoncini, 2008; Takeuchi, Kawai, Shimada, & Yokota, 1993).

Material and methods

Plant growth conditions and osmotic stress treatment in the bananaTainer

In vitro banana plantlets are obtained through the International Transit Center (Bioversity International, Heverlee, BE). The bananaTainer, the plants are grown in, is a climate controlled reefer container, engineered by Urban Crops Solutions (Waregem, BE). Plants grow under 12/12 (day/night) LED powered light regime, temperature and relative humidity are set to 24/13 °C (day/night) and 70% respectively. The CO2 concentration in the bananaTainer is kept at 350-370 ppm. In this hydroponic growth model plants are grown per six in a shallow tray holding 4L of nutrient solution. Nutrient solution is provided by an overflow system, refreshing the tray volume every hour. The total circuit volume is kept constant. The nutrient solution contains: 361 mg/L KNO_3, 121 mg/L K_2SO_4, 176 mg/L $MgSO_4.7H_2O$, 181 mg/L $MgCl_2.6H_2O$, 194 mg/L KH_2PO_4, 398 mg/L $NaH_2PO_4.2H_2O$, 464 mg/L $Ca(NO_3)_2.4H_2O$, 105 mg/L $CaCl_2.2H_2O$, 0.1125 mL/L EDDHSA-Fe, 1.1 mg/L H_3BO_3, 2.7 mg/L $MnSO_4.H_2O$, 0.23 mg/L $ZnSO_4.7H_2O$, 0.16 mg/L $CuSO_4.5H_2O$, 0.07 mg/L $NaMoO_4.2H_2O$, pH = 6, modified from Swennen et al (1986). Analysis reports of the nutrient solution composition are provided by the Belgian Soil Science Service (Bodemkundige Dienst, Heverlee, Leuven). The pH and EC of the nutrient solution are monitored and logged on 1 min interval.

Osmotic stress is applied by adding 50 gram Polyethylene glycol (PEG-8000) (Carl-Roth) per liter nutrient solution (5%). The osmolality is measured using a Freezing point osmometer (Osmomat-3000, Gonotec GmbH, DE) using 50 µL sample. The osmolality is output in mOsmol/kg, and is converted into the osmotic pressure (MPa) by the Van't Hoff equation (Phi = RTc).

Morphological variables extraction

At the start of the experiment all plants are phenotyped non-destructively (whole plant weight, pseudostem height, and topview leaf area), and the last formed leaf is marked. This phenotyping event at the start of the experiment allows to select the most homogenous subgroup to use in the actual experiment. Based on the total mass, the 24 most homogenous plants per cultivar are selected, excluding those which are too big, too small, or damaged.

After 6 weeks of growth in the bananatainer, all plants are phenotyped again using combined (destructive) measurements and digital imagery. Firstly, the plants are weighed and the pseudostem height measured. Secondly the canopy area is calculated from top view images. Green plant pixels are separated from the blue background by color segmentation. Using a red reference surface of known size (10 x 5 cm) the green area is calculated. Thirdly, the leaves are separated and spread out by increasing age and subsequently imaged. Based on these images the number of leaves, and individual leaf length, width, and area are determined. Lastly, the weight of separated plant parts (root, pseudostem, and leaves) is recorded before and after 14 days of drying at 70°C.

All image analysis is performed using an in house R tool based on the EBImage (Pau, Fuchs, Sklyar, Boutros, & Huber, 2010), and imagemagick packages. The final list of morphological variables (measured and calculated) is taken up in the banana crop ontology database (Banana: CO_325, available at http://www.cropontology.org/terms/CO_325/, accessed Oct. 23, 2018).

Visualization and identification of the microbiota

For visualization and identification purposes the microorganisms are cultured on plates using infected sample from the nutrient solution in the bananaTainer. These plates are aerobically incubated at 20°C. The microbial growth in the nutrient solution is estimated by loading approximately 1.5 mL in a spectrophotometer (Pharmacia BioTech, SE) at λ 600 nm, comparing absorbance to purified water. The standard curve was made by adding increasing volumes of microbiota loaded nutrient solution to 35 mL, nutrient solution (composition same as above) containing 5% PEG. The spectrophotometric measurement were performed after 1h incubation at 24°C.

Identification of the microorganisms was performed at the Belgian Coordinated Collections of Microorganisms at Ghent University. Samples were taken when the infection in the bananaTainer reached its maximum. A serial dilution (10^0 till 10^-8) on 3 different media was established. 1) nutrient solution containing 5 % PEG, solidified by Gelrite (Duchefa Biochemie, NL); 2) trypticase soy agar (TSA) with addition of 5 ppm amphotericin B and 200 ppm cycloheximide and; 3) DYPA with addition of 100 ppm chloramphenicol. After aerobic incubation at 20 °C single colonies were picked and spectra generated using MALDI-TOF MS. Spectral clustering reduces replication and identification is done against an in house LMG reference database. Reliable identifications have log score > 1.7.

Plant growth experiment characterizing chlorine dioxide applicability

The effects of daily ClO_2 addition is assessed in a hydroponics based growth screening trial. Three plants were grown together in a 3 L tray with the roots freely in hydroponic solution (nutrient content same as reported above). Plants were grown in a growth chamber (12/12 light/dark) at 25 °C and 75 % RH. On a week/daily basis the trays were weighed and refilled up to 3 liters, as such the daily transpiration was determined. Subsequently, non-destructive root phenotyping was done by imaging 1 plant per tray using a side view camera (Canon EOS1300D), making sure a red reference surface (50 cm^2) was pictured simultaneously. Finally chlorine dioxide was prepared according to manufacturer's instructions (Aqua Ecologic, BE) and supplied at 0, 0.5, 1, 3, or 10 ppm. In total 3 plants were grown per tray, with 3 trays per treatment in 5 treatments (45 plants in total).

Using R (3.4.3, EBImage, and magick packages) root images of 1 plant per tray were segmented from the background using color segmentation. The root area, as seen from the side, is calculated by counting the number of root pixels and comparing this to the number of pixels of the reference surface. A proxy for root damage was calculated by comparing the number of pixels with Hue value between 0 - 15 (from 0-255 HSV color space) to the total number of root pixels.

Results

Introducing the bananaTainer, a container based, tailor made, banana growth chamber to select climate smart varieties

The major aim of the project is to simulate the stresses prone to the East-African highlands and phenotype the Musa biodiversity for its suitability to grow and produce there. This implies many repeated tests have to be performed in a maximally controlled and fully repeatable environment. For this we introduce the bananaTainer, a container based, tailor made banana growth chamber engineered by Urban Crops Solutions (https://urbancropsolutions.com/, Waregem, BE). The specific, 3 layer, vertical farming design allows to grow 504 randomly placed banana plantlets under tightly controlled conditions in hydroponics (Figure 1). Two separate circuits (each total volume 500 L) allow 2 different nutrient media (simulating drought and/or nutrient deficiency) to set up the experimental design. Frequent irrigation events fully refresh and aerate the nutrient solution every hour. There is complete drain reuse and the total water volume per circuit is monitored and kept constant. Nutrient solution filtering is crucial and ensured by a 80 micron particle filter per circuit. Environmental settings (temperature, relative humidity, CO_2 concentration, and day length) are controlled and can be set according to the experimental design. The LED lights provide a specific spectrum relevant for banana (van Wesemael et al. 2018 (ISHS, in revision)). Every run, two well studied reference cultivars (Cachaco (ITC0643) and Mbwazirume (ITC0084)) are taken along to estimate the inter run variability. In total this allows characterization of 19 new cultivars every run under two conditions using 12 biological replicates (504 = 12 * 2 * (19+2)).

Figure 1: The 'bananaTainer' is a modified, climate controlled reefer container specifically designed to house 504 banana plants. The water circuit contains 2 loops, the main loop supplying the plants with water by an overflow system, passing a particle filter (80 micron) before returning to the nutrient tank. The secondary loop treats the water with a UV source and leads the water over an EC and pH sensor.

The pipeline encompassing plant nursery growth, pre-test phenotyping, bananatainer growth, post-test phenotyping followed by feature extraction and data analysis has successfully been established. So far we tested 57 different cultivars, of 24 different subgroups in the bananatainer pipeline. In total this represents 631 different cultivars of the ITC Musa germplasm collection (Table 1). The aim is to evaluate the suitability of cultivar growth in the East African highland agro-ecological zone. By

expressing the daily growth of the plants relative to the reference cultivar taken along every run (Cachaco, ITC0643, Bluggoe, ABB) we can rank the cultivar performance (Figure 2). Roughly 50% of the cultivars grow better than the reference cultivar. Cultivars with significant better growth are Uzakan, Silk, Pahang, Lacatan, Raja, Huti Green Bell, and Simili Radjah (p-value<0.05, Tukey HSD test). Ingarama (ITC0160, Mutika/Lugujira, AAA) performs significantly worse than the reference cultivar.

Table 1: 57 used cultivars, representing 24 different subgroups, and a total of 631 different ITC cultivars.

SubSpecies/SubGroup	Genome	# accessions in ITC	Cultivar	ITC code	Exp
banksii	AA	30	Banksii	ITC0623	III
zebrina	AA	6	Zebrina	ITC1177	II
malaccensis	AA	15	Pahang	ITC0609	I
Mshare	AA	14	Huti green bell	ITC1559	I
			Makyughu II	ITC1446	III
Sucrier	AA	9	Pisang Mas	ITC0653	I
Unknown	AA	/	Guyod	ITC0299	I
Unknown	AA	/	Mjenga Gros Michel Diploid	ITC1253	III
Cavendish	AAA	51	Grande Naine	ITC0180	II
			Williams	ITC0365	II
			Lacatan	ITC0768	III
			Poyo	ITC0345	III
Gros Michel	AAA	9	Gros Michel	ITC1122	II
			Highgate	ITC0263	III
Ibota	AAA	7	Khai Thong Ruang	ITC0662	II
Mutika/lugujira	AAA	76	Mbwazirume	ITC1356	I,II,III
			Igisahira Gisanzwe	ITC0083	I
			Ingagara	ITC0166	I
			Ingumba y'Imbihire (Inyamunyu)	ITC0155	I
			Inyoya	ITC0163	I
			Inzirabahima	ITC0150	I
			Mpologoma	Noitc	I
			Nakitengwa	ITC1180	I
			Namunwe	ITC1629	I
			Nyitabunyonyi	ITC1556	I
			Igitsiri (Intuntu)	ITC0081	II
			Guineo	ITC0005	III
			Ingarama	ITC0160	III
			Ingumba y'Inyamunyo	ITC0126	III
			Kibiddebidde	ITC1624	III
			Kisukari usiniguse	ITC1546	III
			Makara	ITC0177	III
			NSH 42	ITC1802	III
Red	AAA	10	Red Dacca	ITC0575	II
Rio	AAA	3	Leite	ITC0277	II
Unknown	AAA	/	Pisang Berangan	ITC1287	II
Unknown	AAAA	/	Fhia-23	ITC1265	I
Iholena	AAB	2	Uzakan	ITC0825	II
Mysore	AAB	10	Pisang Ceylan	ITC1441	II
Pisang Kelat	AAB	5	Pisang Palembang	ITC0450	II
Pisang Raja	AAB	4	Pisang Rajah	ITC0587	II
			Pisang Raja Bulu (Raja)	ITC0843	III
Plantain	AAB	292	Orishele	ITC1325	II
Pome	AAB	25	Foconah	ITC0649	I
			Prata Ana	ITC0962	II
Silk	AAB	14	Kipakapaka	ITC1550	I
			Yangambi n°2	ITC1275	II
			Figue Pomme Géante (Silk)	ITC0769	III
Bluggoe	ABB	16	Cachaco	ITC0643	I,II,III
			Dole	ITC0767	II
			Kivuvu	ITC0157	II
Monthan	ABB	10	Monthan	ITC1483	I
Pelipita	ABB	4	Pelipita	ITC0472	III
Peyan	ABB	2	Simili Radjah	ITC0123	III
Pisang Awak	ABB	17	Kayinja	ITC0087	I
			Namwa Khom	ITC0659	I
			Fougamou	ITC0101	III
Pisang Klutuk Wulung	BB	/	Pisang Klutuk Wulung	ITC1587	III

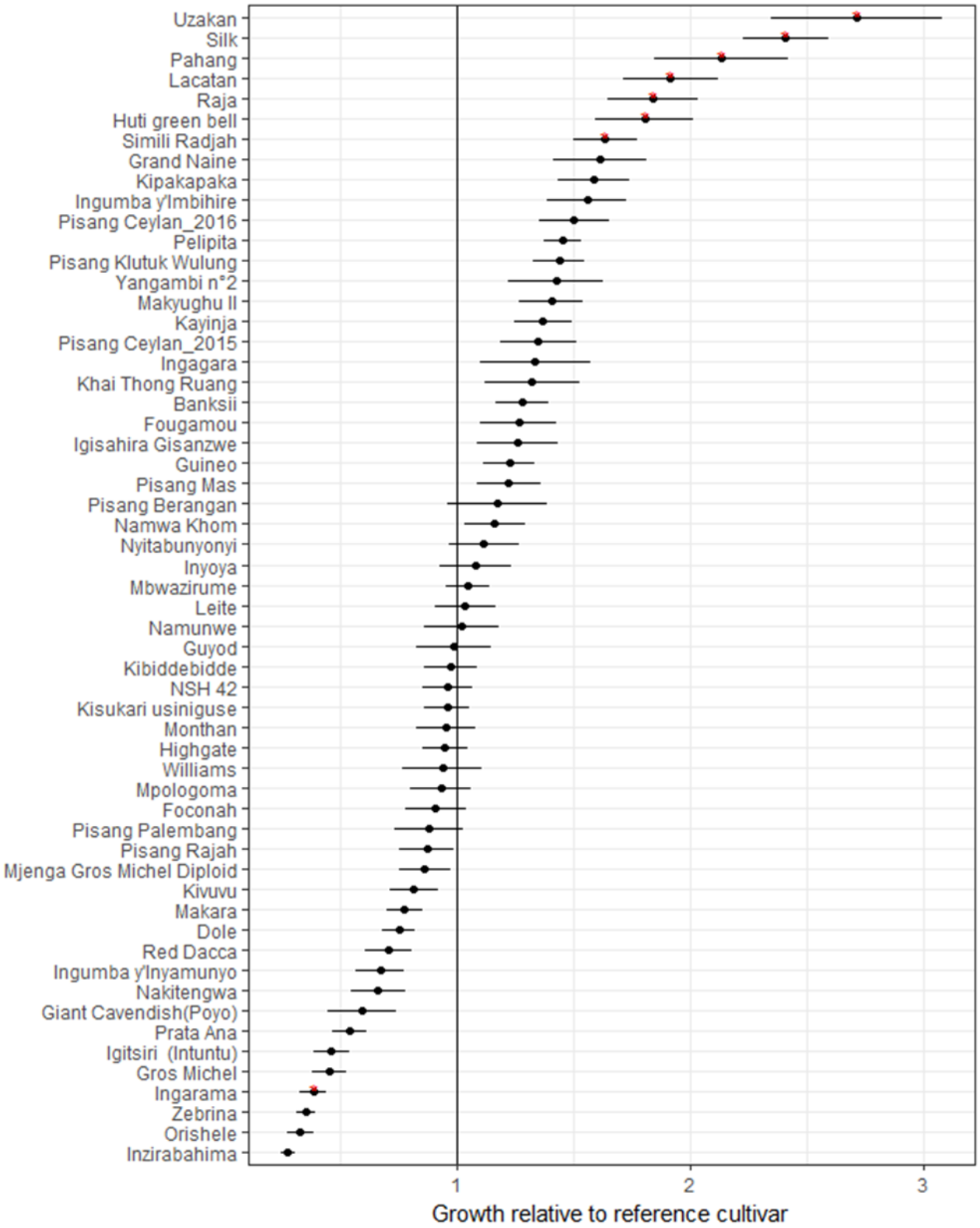


Figure 2: Impact of low temperature on the growth of 57 cultivars evaluated relative to the reference cultivar (Cachaco, Bluggoe, ABB). Mean relative growth with standard error is represented. Significant differences are indicated (p-value<0.05).

The primary output of the bananaTainer growth setup is to obtain a relative ranking of cultivar suitability for the East African highlands (cfr. Figure 2). However, we generate additional phenotypic features, complementing this growth ranking in order to get insight in the crucial features differentiating the well suitable cultivars. This phenotyping is for example done via imaging. Per run

over 870 images are taken and processed. An example of the image analysis pathway, for loose leaves is displayed in Figure 3. These images of loose leaves allow to phenotype the evolution of leaf growth throughout the experiment.

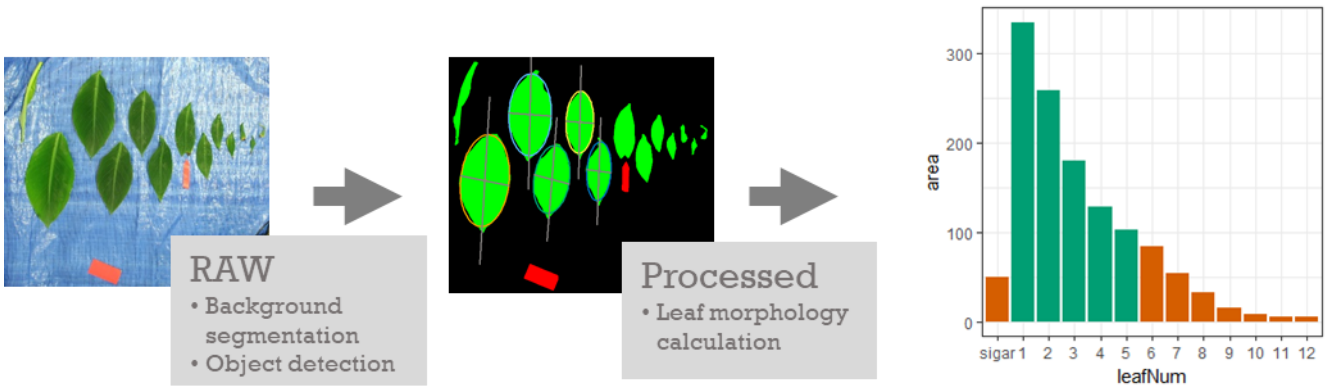


Figure 3: Image analysis pipeline describing evolution of leaf area through the experiment.

A characterization of the microbial biomass growing on PEG

In the PEG circuit runs were infected by micro-organisms. Upon microbial infection, the aerated nutrient solution is characterized by an increase in spectrophotometric turbidity (λ 600 nm). The turbidity is linearly correlated to microbial biomass (Figure 4). The pH of the nutrient solution also increases from pH 6 to pH 7 in 7 days, while the pH of the control nutrient solution does not change (Figure 5A). Nutrient composition analysis shows that the pH change corresponds with a serious shift in N and HCO_3 (Figure 5B). The micro-organisms consume N and produce CO_2 that converts to HCO_3

Via MALDI-TOF MS spectra of isolated colonies three colonies which could grow solely on the PEG substrate, are confidently identified: Cupriavidus sp., possibly C. basiliensis (log score = 2.000), Pseudomonas putida species group (log score = 2.470), Acinetobacter sp., probably A. johnsonii (log score 2.550).

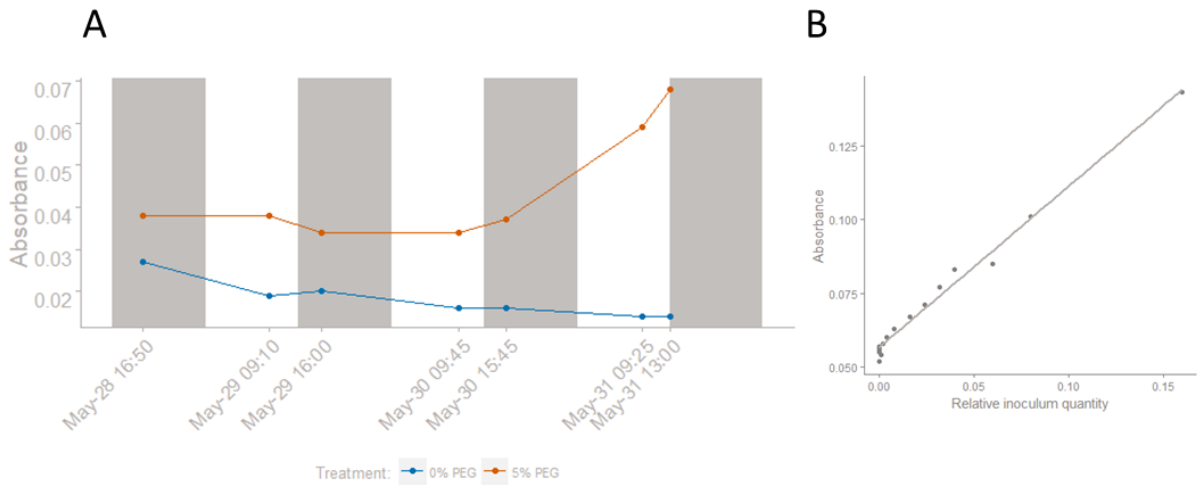


Figure 4: (A) The microbial presence manifests as an increase in spectrophotometric turbidity (OD lambda 600 nm). (B) The turbidity increases linear with the inoculum quantity.

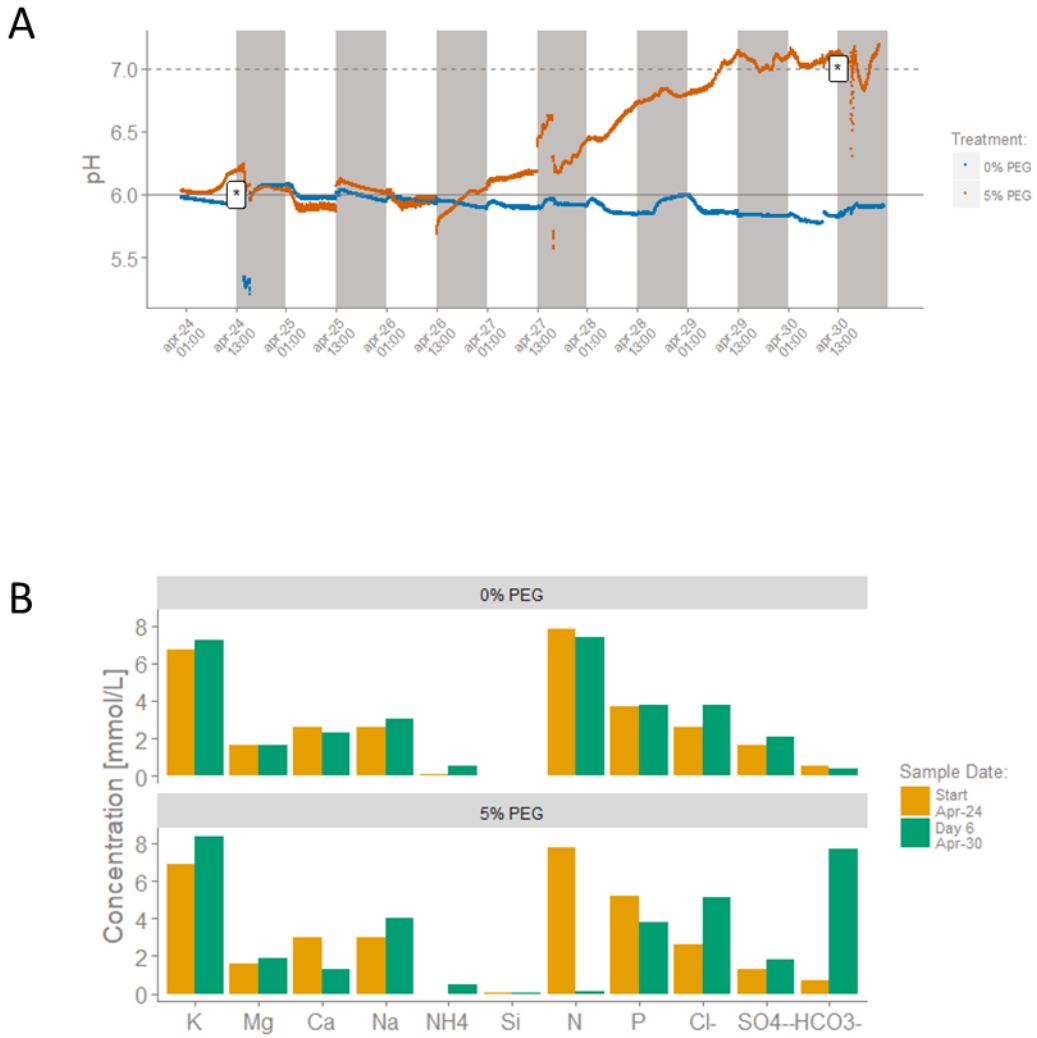


*Figure 5: (A) The microbial infection is manifested by a pH rise in the nutrient medium with addition of 5% PEG. pH measured in both tanks continuously (1 minute interval). Sampling times for nutrient solution analysis indicated by *. Nighttime periods indicated by shading. (B) The microbial infection depletes nutrient solution in the 5% PEG nutrient solution.*

Effects of daily chlorine dioxide application on plant performance traits.

The initial healthy root system shows severe browning under daily addition ClO_2. Especially the root hairs are vulnerable to the non-selective oxidation of chlorine dioxide (Figure 6). Daily application of chlorine dioxide lowers transpiration rate at 0.5, 1, and 3 ppm, but especially at 10 ppm (Figure 7A). The increase in root area is lowered at 3 and 10 ppm applied ClO_2, as visualized by the side view root area over time in Figure 7B. The daily addition of chlorine dioxide causes root browning (Figure 7C) by the relative pixel count of pixels in the red space of HSV color space (H < 15).

Figure 6: The daily addition of chlorine dioxide causes root browning, impeding root growth. The root hair structures are most vulnerable.

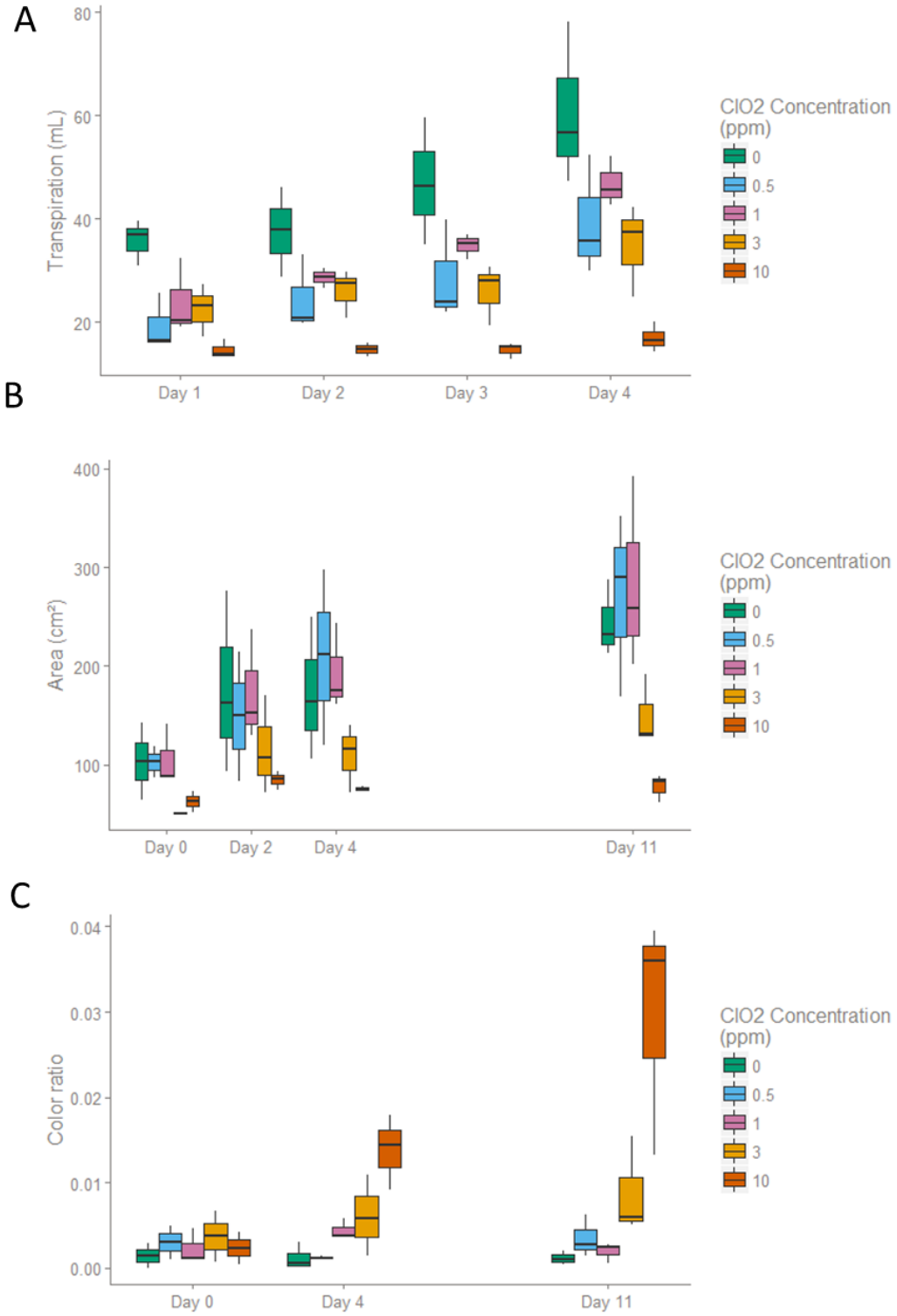


Figure 7: Daily chlorine dioxide applications have impact on plant performance: A) daily transpiration rate, B) root area, C) root brownness. Transpiration is determined per tray, there are three plants grown together in 1 tray with 3 replicate trays per treatment. The root area is calculated for 1 plant per tray. The root brownness variable is also calculated for 1 plant per tray. It is the proportion of root pixels with H<15 (in 0-255 HSV color space).

Discussion

In this work we introduce the bananaTainer, a relatively cheap, large scale banana characterization infrastructure designed for a high throughput growth characterization (Figure 1). In our lab model, vegetative growth is a proxy for the field performance. The aim is to screen for the suitability to grow in the African Highlands. The evaluation in high throughput pre-screening as introduced here is the first high throughput attempt to add phenotypic value to the Musa germplasm collection of ITC. For further studies it allows end users (scientists, farmers, extension workers) to focus on the most promising subset of cultivars. For this pre-screening inspired aim, the bananaTainer has a suitable design (Figure 1), allowing 12 biological replicates, with 2 reference cultivars overcoming the inter-run variability. Water is continuously recycled in the overflow based system and the total running volume is kept constant. Water recycling however, also brings in phytosanitary issues. Only in the PEG stress treatment we observed an increase in spectrophotometric turbidity (Figure 4) and a drastic pH increase (Figure 5A). This coincided with a depletion of nutrients (Figure 5B) and was correlated to microorganism presence. No carbon is added to the nutrient solution, apart from PEG. We validated the presence of 3 PEG metabolizing bacteria in the nutrient solution (Cupriavidus basiliensis, Pseudomonas putida, and Acinetobacter johnsonii). (Marchal et al., 2008) studied specialized microflora from waste water treatment installations and from soils, and found Cupriavidus basiliensis and Pseudomonas putida are able to metabolize PEG 400. In our case a larger PEG-8000 molecule is consumed. Plants normally grow in environments full of microorganisms: the soil. As such non-pathogenic microorganisms are not an issue. However, the bacterial nutrient consumption in our environment reduces the interpretability and experimental reproducibility because it depends on the microbial activity. Colonization speed, growth rate, … can be different from run to run. Prevention of microorganism infection in combination with PEG proves rather difficult, as the addition of chlorine dioxide, a non-selective oxidator of organic material (Scarlett et al., 2016) damages the roots and affects the transpiration (Figure 6, Figure 7). The chlorine dioxide oxidizes in a non-selective fashion, targeting bacteria, roots, iron chelates, … The root hairs, the most vulnerable structures of the root system, are damaged first, attributing to the impeded transpiration. In the most severely damaged root systems there is a 10 fold increase in brownness, compared to the fresh start

The main project objective was to establish a high throughput, pre-field phenotyping infrastructure based on vegetative growth. Models using PEG to apply very precise and repeatable low water potentials in soilless, hydroponic cultivations seem relevant in this regard (Poorter et al., 2012). However, using PEG as a simulation of drought is not possible. Alternative characterization of the drought response is to observe the impact of cessation of water, relative to *ad libidum* water supply. These experiments are highly relevant but more challenging to control.

References

Allen, C., Dos Santos, N., Gallagher, R., Chiu, G. N. C., Shu, Y., Li, W. M., … Bally, M. B. (2002). Controlling the Physical Behavior and Biological Performance of Liposome Formulations through Use of Surface Grafted. Bioscience Reports, 22(2), 225–250. https://doi.org/10.1023/A:1020186505848

Bernhard, M., Eubeler, J. P., Zok, S., & Knepper, T. P. (2008). Aerobic biodegradation of polyethylene glycols of different molecular weights in wastewater and seawater. Water Research, 42(19), 4791–4801. https://doi.org/10.1016/j.watres.2008.08.028

Blum, A. (2011). Plant Breeding for Water-Limited Environments. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-7491-4

Christelová, P., De Langhe, E., Hřibová, E., Čížková, J., Sardos, J., Hušáková, M., … Doležel, J. (2016). Molecular and cytological characterization of the global Musa germplasm collection provides insights into the treasure of banana diversity. Biodiversity and Conservation, 1–24. https://doi.org/10.1007/s10531-016-1273-9

Dalal, A., Attia, Z., & Moshelion, M. (2017). To Produce or to Survive: How Plastic Is Your Crop Stress Physiology? Frontiers in Plant Science, 8(December), 1–8. https://doi.org/10.3389/fpls.2017.02067

De Boeck, H. J., Vicca, S., Roy, J., Nijs, I., Milcu, A., Kreyling, J., … Beier, C. (2015). Global Change Experiments: Challenges and Opportunities. BioScience, 65(9), 922–931. https://doi.org/10.1093/biosci/biv099

Dwivedi, S. L., Ceccarelli, S., Blair, M. W., Upadhyaya, H. D., Are, A. K., & Ortiz, R. (2016). Landrace Germplasm for Improving Yield and Abiotic Stress Adaptation. Trends in Plant Science, 21(1), 31–42. https://doi.org/10.1016/j.tplants.2015.10.012

Dwyer, D., & Tiedje, J. M. (1986). Metabolism of Polyethylene Glycol by Two Anaerobic Bacteria, Desulfovibrio desulfuricans and a Bacteroides sp. Applied and Environmental Microbiology, 52(4), 852–856.

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., … O'Connell, C. (2011). Solutions for a cultivated planet. Nature, 478(7369), 337–342. https://doi.org/10.1038/nature10452

Hegde, D. M., & Srinivas, K. (1989). Effect of soil matric potential and nitrogen on growth, yield, nutrient uptake and water use of banana. Agricultural Water Management, 16(1–2), 109–117. https://doi.org/10.1016/0378-3774(89)90045-0

Huang, Y. L., Li, Q. B., Deng, X., Lu, Y. H., Liao, X. K., Hong, M. Y., & Wang, Y. (2005). Aerobic and anaerobic biodegradation of polyethylene glycols using sludge microbes. Process Biochemistry, 40(1), 207–211. https://doi.org/10.1016/j.procbio.2003.12.004

Junker, A., Muraya, M. M., Weigelt-Fischer, K., Arana-Ceballos, F., Klukas, C., Melchinger, A. E., … Altmann, T. (2015). Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. Frontiers in Plant Science, 5(January), 1–21. https://doi.org/10.3389/fpls.2014.00770

Kawai, F. (2002). Microbial degradation of polyethers. Applied Microbiology and Biotechnology, 58(1), 30–38. https://doi.org/10.1007/s00253-001-0850-2

Langridge, P., & Reynolds, M. P. (2015). Genomic tools to assist breeding for drought tolerance. Current Opinion in Biotechnology, 32, 130–135. https://doi.org/10.1016/j.copbio.2014.11.027

Lipper, L., Thornton, P., Campbell, B. M., Baedeker, T., Braimoh, A., Bwalya, M., … Torquebiau, E. F. (2014). Climate-smart agriculture for food security. Nature Climate Change, 4(12), 1068–1072. https://doi.org/10.1038/nclimate2437

Marchal, R., Nicolau, E., Ballaguet, J. P., & Bertoncini, F. (2008). Biodegradability of polyethylene glycol 400 by complex microfloras. International Biodeterioration and Biodegradation, 62(4), 384–390. https://doi.org/10.1016/j.ibiod.2008.03.013

Michel, B. E. (1983). Evaluation of the water potentials of solutions of polyethylene glycol 8000 both in the absence and presence of other solutes. Plant Physiology, 72(1), 66–70. https://doi.org/10.1104/pp.72.1.66

Michel, B. E., & Kaufmann, M. R. (1973). The Osmotic Potential of Polyethylene Glycol 6000. Plant Physiol, 51, 914–916. https://doi.org/10.1104/pp.51.5.914

Mickelbart, M. V., Hasegawa, P. M., & Bailey-Serres, J. (2015). Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. Nature Reviews Genetics, 16(4), 237–251. https://doi.org/10.1038/nrg3901

Money, N. P. (1989). Osmotic Pressure of Aqueous Polyethylene Glycols. Plant Physiology, 91(2), 766–769. https://doi.org/10.1104/pp.91.2.766

Moshelion, M., & Altman, A. (2015). Current challenges and future perspectives of plant and agricultural biotechnology. Trends in Biotechnology, 33(6), 337–342. https://doi.org/10.1016/j.tibtech.2015.03.001

Negin, B., & Moshelion, M. (2017). The advantages of functional phenotyping in pre-field screening for drought-tolerant crops. Functional Plant Biology, 44(1), 107–118. https://doi.org/10.1071/FP16156

Ortiz, R., & Swennen, R. (2014). From crossbreeding to biotechnology-facilitated improvement of banana and plantain. Biotechnology Advances, 32(1), 158–169. https://doi.org/10.1016/j.biotechadv.2013.09.010

Pau, G., Fuchs, F., Sklyar, O., Boutros, M., & Huber, W. (2010). EBImage-an R package for image processing with applications to cellular phenotypes. Bioinformatics, 26(7), 979–981. https://doi.org/10.1093/bioinformatics/btq046

Poorter, H., Fiorani, F., Stitt, M., Schurr, U., Finck, A., Gibon, Y., … Pons, T. L. (2012). The art of growing plants for experimental purposes: A practical guide for the plant biologist. Functional Plant Biology, 39(11), 821–838. https://doi.org/10.1071/FP12028

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS ONE, 8(6). https://doi.org/10.1371/journal.pone.0066428

Scarlett, K., Collins, D., Tesoriero, L., Jewell, L., van Ogtrop, F., & Daniel, R. (2016). Efficacy of chlorine, chlorine dioxide and ultraviolet radiation as disinfectants against plant pathogens in irrigation water. European Journal of Plant Pathology, 145(1), 27–38. https://doi.org/10.1007/s10658-015-0811-8

Steuter, A. A., Mozafar, A., & Goodin, J. O. E. R. (1981). Water potential of aqueous polyethylene glycol. Plant Physiology, 67(1), 64–67. https://doi.org/10.1104/pp.67.1.64

Takeuchi, M., Kawai, F., Shimada, Y., & Yokota, A. (1993). Taxonomic Study of Polyethylene Glycol-Utilizing Bacteria: Emended Description of the Genus Sphingomonas and New Descriptions of Sphingomonas macrogoltabidus sp. nov., Sphingomonas sanguis sp. nov. and Sphingomonas terrae sp. nov. Systematic and Applied Microbiology, 16(2), 227–238. https://doi.org/10.1016/S0723-2020(11)80473-X

Taulya, G., van Asten, P. J. A., Leffelaar, P. A., & Giller, K. E. (2014). Phenological development of East African highland banana involves trade-offs between physiological age and chronological age. European Journal of Agronomy, 60, 41–53. https://doi.org/10.1016/j.eja.2014.07.006

van Asten, P. J. A., Fermont, A. M., & Taulya, G. (2011). Drought is a major yield loss factor for rainfed East African highland banana. Agricultural Water Management, 98(4), 541–552.

Vanhove, A.-C., Vermaelen, W., Panis, B., Swennen, R., & Carpentier, S. C. (2012). Screening the banana biodiversity for drought tolerance: can an in vitro growth model and proteomics be used as a tool to discover tolerant varieties and understand homeostasis. Frontiers in Plant Science, 3(176), 1–10.

Verslues, P. E., Agarwal, M., Katiyar-Agarwal, S., Zhu, J., & Zhu, J. K. (2006). Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. Plant Journal, 45(4), 523–539. https://doi.org/10.1111/j.1365-313X.2005.02593.x

Verslues, P. E., Ober, E. S., & Sharp, R. E. (1998). Root Growth and Oxygen Relations at Low Water Potentials. Impact of Oxygen Availability in Polyethylene Glycol Solutions. Plant Physiology, 116(4), 1403–1412. https://doi.org/10.1104/pp.116.4.1403

Genomic constitution characterization in banana ABB allotriploids

Abstract

Bananas (Musa spp.) are a major staple food for hundreds of millions of people in developing countries. The cultivated varieties are seedless parthenocarpic clones of which the origin remains to be clarified. The most important cultivated banana varieties are triploids with an AAA, AAB, or ABB genome constitution, being A and B genomes provided by *Musa acuminata* and *M. balbisiana*, respectively. Prior to the availability of the banana reference sequence, it has been shown that inter-genome recombination was relatively common in banana cultivars and that cultivated triploids were more likely to have passed through an intermediate hybrid.

Using single-nucleotide polymorphism (SNP) markers called from RADseq data, we studied the chromosome structure of 29 ABB genotypes spanning these subgroups. Several intergenomic replacements were identified, where the A genome was substituted by the B genome counterpart and vice versa. The replacement patterns indicate at least 7 founding events at the origin of the ABB cultivars. Most of the subgroup assignations was confirmed, but several inconsistencies were highlighted that necessitate a revision of the subgroup classification.

The result of this study is an important step to unravel the origin of polypoid bananas, which will help revisiting the structure of the subgroups and the classification of ABB cultivars.

Introduction

Bananas (Musa spp.) are a genus of herbaceous monocotyledons belonging to the Zingiberales order. Musa originated on south-east of Asia where it was domesticated; during the European colonization, the banana culture was spread in tropical areas of Africa and America.

The large majority of cultivated banana (cultivars) are triploid (2n=3x=33), i.e. composed by three set of 11 chromosomes. These cultivars are sterile due to the triploidy that hampers the production of balanced gametes and, consequently, their propagation is clonal, i.e. based on multiplication of vegetative tissues. Combination of sterility and parthenocarpy ensures production of seedless fruits that are edible. Two species, *M. acuminata* and *M. balbisiana*, are involved in large scale cultivation contributing with their genomes, A and B respectively, to the subgenomes of triploid cultivars. Triploidy establishment occurs by combination of a normal gamete (n=x=11) and a gametes with double set of chromosomes (n=2x=22) originated by normal meiosis of a tetraploid plant (2n=4x=44) or by an unreduced meiosis by a diploid genotype (e.g. an AB hybrid).

Triploid cultivars are classified in autotriploid, whose genome is composed by three A subgenomes (AAA), and allotriploid, whose genome is composed by A and B subgenomes (AAB and ABB). Secondary level of classification for allotriploid banana cultivars was mainly based on morphological parameter: traits differentiating *M. acuminata* and *M. balbisiana* were scored in cultivars based on their similarity with donor species (Simmons and Shepherd 1955). Molecular analysis on organelle genomes allowed characterizing cultivar cytotypes and hypothesizing the crossing pathways originating those (Boonruangrod et al. 2008). Secondary levels of classification were proposed based on inferred parental origin (Simmonds, 1962) and on allotriploid origin inference based on cytotypes (i.e. on the mitochondrial and plastidial genomes) (De Langhe et al. 2010). However, these studies pointed out the genomic complexity of the allotriploid cultivars and suggested the occurrence of backcross with parental species made possible by observed residual fertility in some allotriploid cultivars (De Langhe et al. 2010).

Flow cytometry increased the resolution on genomic investigations making possible accurate measure of the ploidy level and molecular markers as Single sequence repeats (SSRs) providing a multi-locus survey of the parental alleles contribution (Christelova et al. 2016). Nowadays, new generation sequencing (NGS) technologies made possible a genome wide detailed survey that provide evidence of origin of small portion of chromosomes (Baurens et al. 2018).

Former genome wide study highlighted that allotriploids are not mere additions of sets of 11 chromosomes, but regions with variable subgenome ratio were detected along the chromosomes (Baurens et al., 2018), testifying the occurrence of recombination between subgenomes.

The goal of this study is to survey a wide range of ABB allotriploids by using SNP frequency data to check the possible presence of regions with allele frequency ratio different from the expected B2:A1 and to evaluate the usefulness of this genome wide analysis in cultivar classification.

Material and methods

Cultivars analyzed

Leaf samples from banana genotypes have been supplied by the Bioversity International Musa Transit Centre hosted at KU Leuven, Belgium. The passport data of this accessions with reported ABB genome constitution was retrieved from the Musa Germplasm Information System (MGIS) (Ruas et al., 2017) (**Table 1**).

Accession code	Accession name	Group	Subgroup	Geographical origin
ITC0026	Sabra	ABB	Unknown	Gabon
ITC0632	Cachaco enano	ABB	Bluggoe	Colombia*
ITC0643	Cachaco	ABB	Bluggoe	Colombia*
ITC1728	Sambrani Monthan	ABB	Bontha	India
ITC1746	Bankel	ABB	Pisang Awak	India
ITC1748	Boddida Bukkisa	ABB	Pisang Awak	India
ITC1483	Monthan	ABB	Monthan	India
ITC0767	Dole	ABB	Bluggoe	France*
ITC0361	Blue Java	ABB	Ney Mannan	Fiji*
ITC1750	Ney Vannan	ABB	Bontha	India
ITC0123	Simili Radjah	ABB	Peyan	India
ITC1600	INIVIT PB-2003	ABB	Saba	Cuba*
ITC1138	Saba	ABB	Saba	France*
ITC0659	Namwa Khom	ABB	Pisang Awak	Thailand
ITC0339	Pisang Awak	ABB	Pisang Awak	Australia*
ITC0472	Pelipita	ABB	Unknown	Philippines
ITC0053	Bom	ABB	Pisang Awak	France*
ITC0087	Kayinja	ABB	Pisang Awak	Burundi*
ITC0101	Fougamou 1	ABB	Pisang Awak	Gabon
ITC0526	Kluai Namwa Khom	ABB	Pisang Awak	Thailand*
ITC1599	Kambani	ABB	Pisang Awak	Tanzania
ITC1721	Karpuravalli	ABB	Pisang Awak	India
ITC1719	Chinia	ABB	Pisang Awak	India
ITC1749	Vananthpurani	ABB	Pisang Awak	India
ITC0396	Pelipita	ABB	Unknown	Costa Rica*
ITC0397	Pelipita Majoncho	ABB	Unknown	Costa Rica*
ITC0652	Kluai Tiparot	ABB	Unknown	Thailand
ITC1700	Kepok Kuning	ABB	Saba	Indonesia
ITC1745	Kepok Tanjung	ABB	Saba	Indonesia

Table 1: List of 29 accessions analyzed. * origin of the previous *ex situ* collection

DNA extraction and RADseq data generation

Genomic DNA was extracted using a modified MATAB method (Risterucci et al. 2000). DNA libraries were constructed and sequenced using the HiSeq2000 (Illumina) technology at Beijing Genomics Institute.

Detection of Homoeologous Recombinations (HR)

Two approaches have been used to survey the subgenomic structure of the cultivar genomes. The first one was based on the method of Baurens and al (2018). Based on known sequence variability in A and B genome, SNP variants are assigned to the parental genome and its frequency reported in a graphical representation of each chromosome, along with the total count of SNP. The following accessions were used to represent *M. acuminata* and *M. balbisiana* sequence variability.

In the second approach SNP variants unknown on the A genome variability (established on AAA cultivars 'Grande Naine' and 'Mbwazirume' and the A genome reference) were assigned to the B genome and used to measure the B alleles frequencies along the Musa chromosomes.

Results

Genome wide survey of A and B assigned SNP allele frequency in allotriploid ABB

SNPs detected in RADseq datasets of 29 genotypes were surveyed with the approach 1 and 2 to check if, in the 11 chromosome groups, the presence of regions where the A/B subgenome balance was different from the expected one based on known genomic structure (i.e. B2:A1 for the allotriploid ABB) (**Figure 1**). Results were consistent between the two approaches.

Only two genotypes did not show any region with unexpected subgenome balance: 'Kepok Kuning' (ITC1700) and 'Kepok Tanjung' (ITC1745). Among the remaining 27 ABB genotypes, several regions with subgenome balance B3:A0 or B1:A2 were detected, whereas no region with B0:A3 was found (e. g. in Figure 2, chromosome s 4, 9 and 11 have unbalanced regions). Genome comparisons among ABB cultivars revealed that chromosome regions with deviating subgenome balance (hereafter deviating regions) are in some case shared among the analyzed genotypes. Identical or very similar patterns of deviating regions were found among the 27 genotypes that were then grouped in seven groups (**Table 1**).

The genotypes included in each group have the same pattern of deviating regions. Three deviating regions appear to be shared by three groups (1a, 1b and 1c): in large terminal regions of long arm of both 4 and 11 chromosomes, the detected structure was B3:A0, and in the short arm of chromosome 9 an interstitial region was B1:A2. Groups 1a and 1c also shared the B3:A0 terminal region of chromosome 11 short arm. Additional deviating regions were specific of each group.

Group 4, including genotypes classified as Pelipita, had 2 whole chromosomes (i.e. chr02 and chr11) with pattern B3:A0 and the chr03 has a B3:A0 pattern but an interstitial region in long arm with the expected B2:A1 pattern.

Kluai Tiparot (ITC0652), unique representing of group 5, had a particular deviating pattern, not shared with any of the studied genotypes. In this cultivar all the detected unbalanced regions were all B3:A0 and three chromosomes (i.e. chr02, chr08 and chr10) resulted completely B3:A0, in addition to several chromosome parts.

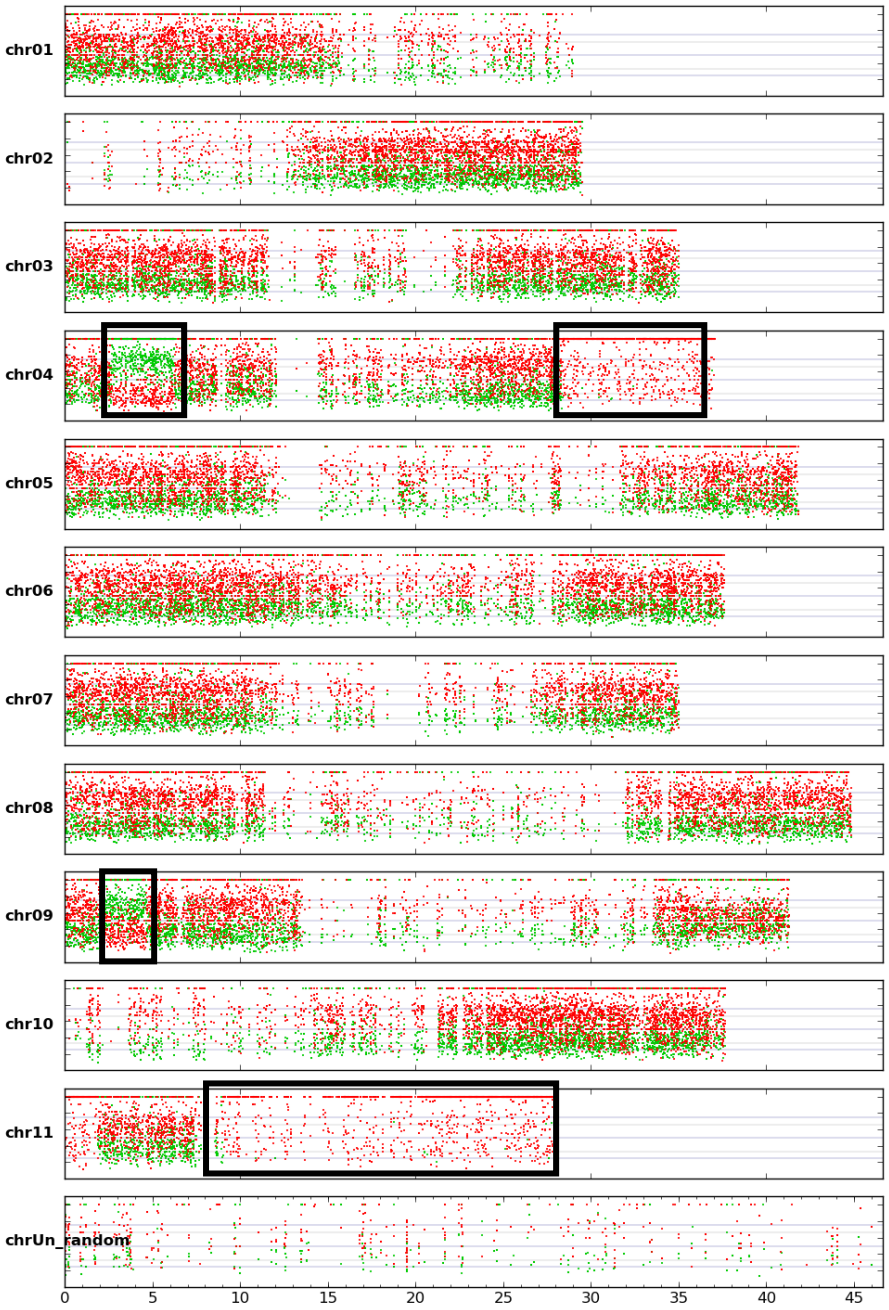


Figure 1. Mosaic genome structure of A/B interspecific ABB genome for the 1a group along the 11 chromosomes. Coverage ratios displayed by Red and green dots represent allelic variation corresponding to to B and A genomes respectively. Open black box indicates chromosome segments that differ from the expected distribution in a ABB cultivar. The scale is indicated in Mb.

Specific deviating regions were found in particular genotype included in some groups. In many cases, these deviating regions are characterized by the presence of only two chromosome regions (e.g. B2:A0 or B1:A1), testifying the loss of whole or partial chromosomes (**Figure 2**). In Group 1a, 'Dole' (ITC0767) missed the A chromosome 8, resulting in a B2:A0 pattern along the whole chromosome (**Figure 2b**). On a technical point of view, regions B2:A0 and B3:A0 have similar patterns in both the survey method used in this study: only B variants were detected. However, the absence of one chromosome in Dole can be detected observing the distribution of heterozygous B variants. In fact the point distribution is unimodal around frequency value of 0.5 when two B chromosomes are present, whereas, it is bimodal around 0.33 and 0.67 frequency values in the case of three chromosomes B. Coverage data are also used to corroborate the chromosome loss inference.

The short arm of the chromosome 6 of 'Sabra' (ITC0026, group 1a) is characterized by a B1:A1 pattern whereas in the long arm four regions can be distinguished with the following patterns of B2:A1, B3:A1, B1:A1 and B2:A1 (from the centromere to the telomere) (Figure 3). The pattern of region B3:A1 can be distinguished by the B2:A1 due to the more separated variant frequency distribution (0.25 and 0.75 for A and B variants, respectively), and to the higher coverage of the corresponding region (**Figure 2c**)

In group 1b, chromosome group 5 of 'Simili Radjah' (ITC0123) showed unique deviations: 1) a small terminal region of short arm has only B genome derived variants; 2) a whole long arm was missing, resulting in B1:A1 configuration along the most of the arm; 3) B2:A0 in the terminal region. B1:A1 regions can be distinguished from the B2:A1 due to the frequency of A and B variants around the 0.5 value vs two distinguished distribution around 0.33 and 0.67 (A and B variants, respectively) (Figure 2d).

In group 1c, the genotype 'INIVIT PB-2003' (ITC1600) showed a unique interstitial region B2:A2 in the chromosome 10 long arm (Figure 2e). In group 2, the genotype 'Pisang Awak' (ITC0339) has a unique B1:A1 configuration in the terminal part of the chromosome 7 short arm and one (or possibly two) interstitial region B3:A0 in the chromosome 10 short arm (**Figure 2f**).

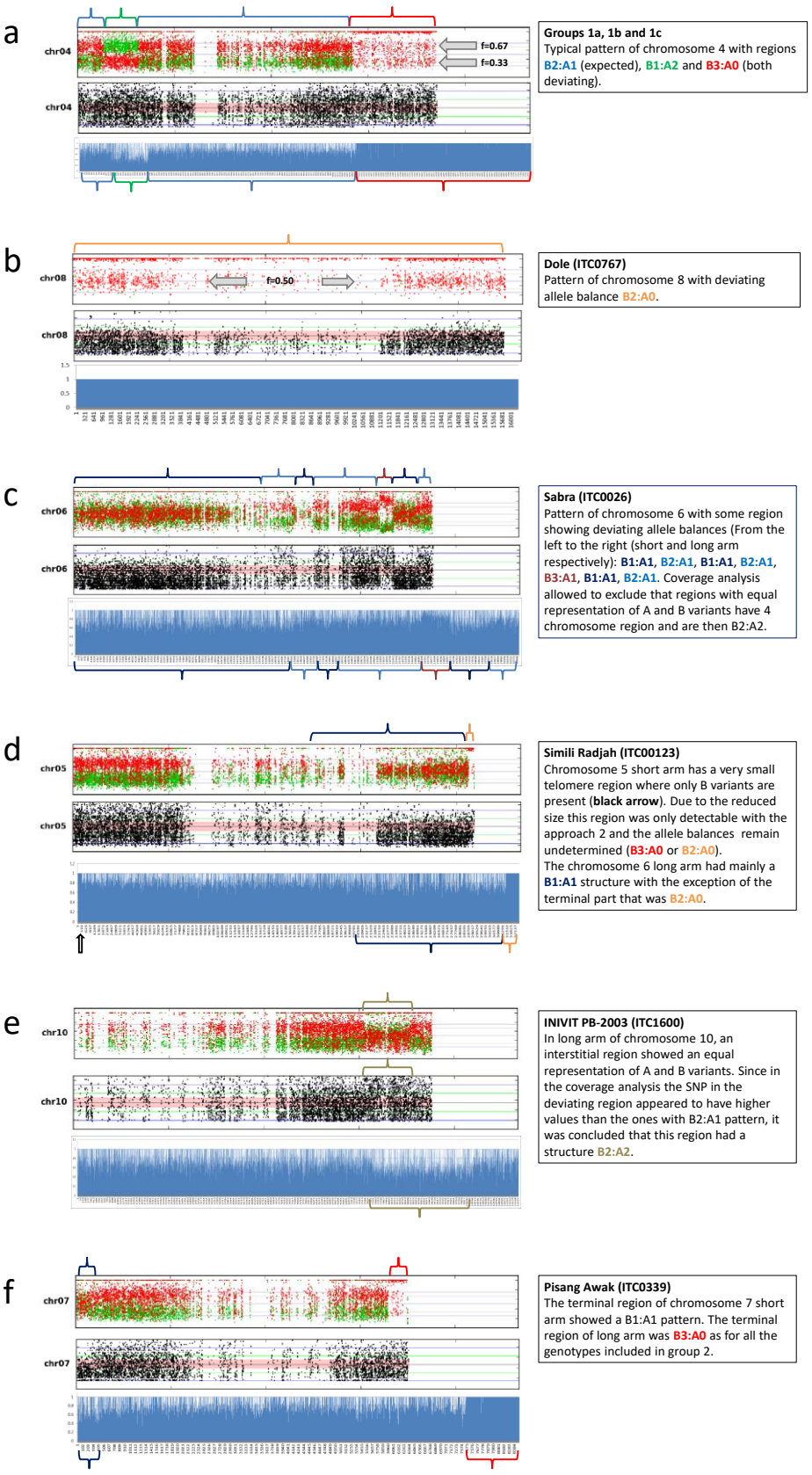


Figure 2. Aneuploid patterns for 5 subgroups in ABB. Euploid recombinant pattern (a) compared with aneuploid patterns.

Discussion

In this study, we take advantage of NGS technologies, in particular, RADseq data, to survey the genomic structures of a panel of allotriploid (ABB) banana cultivars. Homoeologous recombination (HR) was inferred when, in a given chromosome, a change in A/B allele contribution was observed between adjacent regions.

Genome wide analyses were performed with two different approaches. In the first, SNP variants were assigned to respective A or B subgenomes based on known variability in both genepool. In the second approach only the better known A variability is taken into account to detect and remove the SNP assigned to the A subgenome. All variants alternatives to the A reference and not known in the A genome variability were assigned to the B subgenome. In both approaches the SNP variant frequencies were used to infer the local genomic constitution.

Results obtained with both approaches were perfectly consistent. The result interpretation is easier with the first approach, especially in regions characterized by changes in copy number as partial aneuploidy due to segmental duplications or deletions (**Figure 2c**). However, the second approach can allow the detection of small deviating regions as in the short arm of chromosome 6 of 'Simili Radjah' (Figure 2d).

Globally, all accessions were confirmed having an ABB genome constitution, i.e. the ratio between variants assigned to B and A subgenomes was B2:A1. Among the 29 genotypes surveyed, 27 had at least two chromosome portions where the observed ratio was B1:A2 or B3:A0.

The occurrence of regions with deviating ratio between B and A SNP variants, indicates that in these regions B subgenome substituted A counterpart (B3:A0) or vice versa (B1:A2). One can conclude that the involved chromosome underwent recombination that resulted in chromatin exchanges between subgenomes. Since sexual reproduction is supposed to be abolished by triploidy that disturbs correct meiosis process and hampers the production of viable gametes with balanced chromosomes, the recombination should have taken place before the allotriploid formation.

Some of the genotypes were found sharing the whole pattern of unbalanced regions. This observation strongly suggests that all the genotypes with the same deviating pattern originated by vegetative propagation from a common ancestor having the same deviating pattern.

The observation of genotype groups sharing only some unbalanced region (i.e. 1a, 1b and 1c) is more intriguing. In fact, identical or very similar deviating patterns (three unbalanced regions in the case of mentioned groups) obtained by independent recombinations is unlikely. Consequently, one should hypothesize that unbalanced regions were generated in a two step process, the oldest one common to the ancestors of the three groups. The shared events could be already present AB hybrid, transmitted by parental A and B gametes where interspecific introgressions were present; this hybrid then contributed with unreduced gametes (having new recombinations originated in independent meiosis) to the allotriploid ancestors of the three groups 1a, 1b and 1c. Another possibility is that a common allotriploid ancestor was able to produce descendants by sexual reproduction, then creating new and independent recombinations. In the case of the above-mentioned groups, 1b could be the first generation of allotriploid whereas 1a and 1c could be derived from 1b by production of 2x gametes where new independent recombination events take place, and then creating new unbalanced regions.

The genotype 'Kluai Tiparot' (ITC0652) can be distinguished from all the other surveyed accessions for two features: all deviant regions were B3:A0 (no B1:A2 were observed) and three entire chromosomes (2, 8 and 10) were found B3:A0. The observed larger contribution of *M. balbisiana* to 'Kluai Tiparot' could be the explained by a backcross of an AB hybrid that produced haploid gamete with a normal *M. balbisiana* gamete. In the derived diploid genotype (backcrossed hybrid), the contribution of the B genome is expected to be ¾ of the genome. Consequently, an unreduced gamete from this backcrossed hybrid combined with a normal *M. balbisiana* gamete is expected to have higher *M. balbisiana* contribution that an allotriploid whose unreduced gamete was provided by a non-backcrossed AB hybrid.

Similar explanation can be proposed for the cultivars included in group 4 ('Pelipita' (ITC0396 and ITC0472) and 'Pelipita Majoncho' (ITC0397)) whose entire chromosomes 2 and 11 have a B3:A0 structure and the chromosome 7 is mainly B3:A0 with the exception of a long arm interstitial region. However, in these patterns, two regions in chromosome 9 (interstitial in short arm and terminal in long arm) have a B1:A2 structure.

Revisiting subgroups

Genome wide surveys highlighted some inconsistencies on composition of subgroups used to classify ABB cultivars or on the attribution of cultivars to the subgroups. For example, the genotypes representing Monthan and Bontha subgroups share their deviating patterns with genotypes included in Bluggoe subgroups (Genomic group 1a). Moreover, two cultivars assigned to Pisang Awak subgroups ('Bankel' and 'Boddida Bukkisa') had the genomic group 1a deviating pattern, whereas, other cultivars classified as Pisang Awak shares the genomic group 2 or 3 patterns. Another cultivar classified in Bontha subgroups showed the genomic group 2a pattern as well as Blue Java (classed in subgroup Ney Mannan) and 'Simili Radjah' (subgroup Peyan). On the contrary, all the three accessions belonging to the subgroup 'Pelipita' shared an identical and specific deviating pattern (group 4).

These patterns should be considered to update classification of ABB allotriploids. In fact, the genome structure has not only an evolutionary interest, but also can influence significantly the phenotype. Even if somaclonal selection has emphasized on the most visible traits, a common clonal origin implies higher similarity on alleles and gene regulation.

Aneuploid accessions

Almost all the surveyed accessions were euploid (entirely triploid). However, in specific genotypes, aneuploidy was detected involving the disappearing of an entire chromosome, chromosome arm or chromosome portion testifying the occurrence of chromosome loose or partial deletion. Few regions were also detected with supernumerary regions, likely resulting from duplication of chromosome portions. Since these aneuploidy were individuated in single genotypes (not shared with cultivars having the identical deviating pattern), one could suppose that they originated from mutation events more recent than the original triploidization.

Perspectives and future analyses

The findings of HR in almost all the allotriploid cultivar analyzed questions on the origin of recombinant events between A and B genomes. Are exchanges already present in interspecific hybrids (AB)? Are some of the observed recombinations still present in existing AB hybrids? To answer to these questions, AB hybrids available in ITC collection will be submitted to the same analyses performed on ABB allotriploids.

Since ABB allo-triploids have different homeoallele doses according to the region where they lie, one can hypothesize that the variability in genome structure has an impact on phenotypic variability. A set of transcriptomic data already published will be used to estimate which is the impact of this variability on gene expression.

Finally, since the sexual reproduction does not occur in triploids, the cultivars having the same recombination pattern are supposed to derivate from the same ancestor only by vegetative propagation. They present variability is consequently due exclusively to the accumulation of mutations occurred after the triploidy establishment. The RADseq data will be next investigated to trace back the history of cultivar divergence.

References:

Simmonds NW, Shepherd K (1955). The taxonomy and origins of the cultivated bananas. Botanical Journal of the Linnean Society, 55: 302-312. Doi: https://doi.org/10.1111/j.1095-8339.1955.tb00015.x

Boonruangrod R, Desai D, Fluch S, Berenyi M, Burg K (2008). Identification of cytoplasmic ancestor gene-pools of Musa acuminata Colla and Musa balbisiana Colla and their hybrids by chloroplast and mitochondrial haplotyping. Theor Appl Genet 118: 43-55. Doi: 10.1007/s00122-008-0875-3

Christelová P, Valárik M, Hribová E, De Langhe E, Doležel J (2011). A multi-gene sequence based phylogeny of the Musaceae (banana) family. BMC Evolutionary Biology 11: 103. Doi: https://doi.org/10.1186/1471-2148-11-103

Simmonds NW (1962) The evolution of the bananas. Harlow, Essex. Longmans

De Langhe E, Hřibová E, Carpentier S, Doležel J, Swennen R (2010). Did backcrossing contribute to the origin of hybrid edible bananas? Annals of Botany 106: 849-857. Doi: https://doi.org/10.1093/aob/mcq187

Baurens F-C, Martin G, Hervouet C, Salmon F, Yohomé D, et al. (2018) Recombination and Large Structural Variations Shape Interspecific Edible Bananas Genomes? Molecular Biology and Evolution, msy199. Doi: https://doi.org/10.1093/molbev/msy199

Max R, Guignon V, Sempere G, Sardos J, Hueber Y, et al. (2017) MGIS: managing banana (Musa spp.) genetic resources information and high-throughput genotyping data. Ruas M, Guignon V, Sempere G, et al. (2017) MGIS: managing banana (Musa spp.) genetic resources information and high-throughput genotyping data. Database 2017:bax046. Doi: https://dx.doi.org/10.1093%2Fdatabase%2Fbax046

Risterucci AM, Grivet L, N'Goran JAK, Pieretti I, Flament MH, Lanaud C (2000) A high-density linkage map of Theobroma cacao L. Theor Appl Genet 101: 948. https://doi.org/10.1007/s001220051566