**Assessment of metabolic variability and diversity present in leaf, peel and pulp tissue of diploid and triploid *Musa* spp.**

Margit Drapal1, Delphine Amah2, Harald Schöny1, Allan Brown3, Rony Swennen3,4,5 and Paul D.Fraser1\*

1 School of Biological Sciences, Royal Holloway, University of London, Egham Hill, Egham, Surrey, TW20 0EX, UK.

2 International Institute of Tropical Agriculture, Ibadan, Nigeria

3 International Institute of Tropical Agriculture, Arusha, Tanzania

4 Bioversity International, W. De Croylaan 42, 3001 Heverlee, Belgium

5 Department of Biosystems, KU Leuven University, W. De Croylaan 42, 3001 Leuven, Belgium

\*Corresponding author: [P.Fraser@rhul.ac.uk](mailto:P.Fraser@rhul.ac.uk), Tel: +44 1784 443894

**Abbreviations**

cv, cultivar; GC-MS, gas chromatography-mass spectrometry; ITC, International Transit Centre; LC-MS, liquid chromatography mass spectrometry; PLS, partial least squares; UPLC, ultra-performance liquid chromatography.

# Abstract

Banana (*Musa* spp.) plants produce many health promoting compounds in leaf, peel and pulp. For a robust metabolic analysis of these tissues, leaf at five developmental stages were compared to assess suitable sampling practices. Results confirmed that the common sampling practise of leaf 3 is applicable for metabolic comparisons. The developed work flow was applied to analyse the metabolite diversity present in 18 different *Musa* varieties, providing baseline levels of metabolites in leaf, peel and pulp tissue. Correlation analysis was then used to ascertain whether similar trends can be detected in the three plant tissues of the diversity panel. The genome group displayed a dominant role in the composition of the metabolome in all three tissues. This led to the conclusion that a correlation between tissues was only possible within a genome group as the different parental backgrounds caused too great a variation in the metabolomes. It also suggests the metabolome could be used to monitor the interaction/hybridisation of genomes during breeding programmes.

**Keywords**

Banana, breeding resource, chemotype, diversity, metabolite profiling

# Introduction

The current banana production amounts to over 144 million tonnes worldwide and includes 300-400 dessert and cooking varieties [21]. These cultivated bananas are diploid, triploid and tetraploid hybrids of *Musa acuminata* (A genome) and *Musa balbisiana* (B genome). Each of the two genomes confers specific traits, which influence the metabolic composition of the progenies. The A genome, present in most hybrids, gave rise to parthenocarpy and confers fruit flavour and quality, whereas the presence of the B genome contributes to vigorous growth [14, 27]. The majority of cultivated bananas are vegetatively propagated, which effect breeding efforts towards more resistant and nutritious varieties [31]. Part of effective breeding approaches is to elucidate the natural variation present in varieties preferred by consumers [25]. Such data can then be used to define underlying chemical components of the genotypes and more specifically phenotypic traits such as resistance. These metabolites can also be used to act as biomarkers for traits of interest.

Phenylpropanoid derivatives are secondary plant metabolites that comprise a broad range of compounds including flavonoids, phenylpropanoids, hydroxycinnamic acids, coumarins and catechins. Most of these compounds act as part of the plant defence mechanism against abiotic and biotic stresses [43]. They also play an important role in the human diet due to their health promoting effects [34]. Many other compounds beneficial for the human health are present in banana pulp including carotenoids, biogenic amine (e.g. dopamine) and phytosterols [1, 2, 37]. Thus far, studies on banana pulp and peel performed a targeted approach focusing on specific chemical classes. A less targeted approach with metabolite profiling analyses many chemical classes and facilitates an overview of the metabolome present in plant tissue [26].

The present study focused on the metabolic characterisation of i) banana plants throughout the plant’s life cycle and ii) leaf, peel and pulp tissue at harvest. For the first objective, leaf samples of a plantain (Mbi egome) and a cooking banana (Fougamou) were harvested at the vegetative, reproductive and fruiting stage. The samples for each stage comprised leaves of five developmental stages from expanding to senescent and peel and pulp tissue at the fruiting stage. The chemotype of leaf 3 showed the least influence of developmental processes, such as cell proliferation and degradation, on the metabolite composition. The second objective comprised the biochemical phenotyping of 18 *Musa* varieties of wild and cultivated *M. acuminata* (AAw and AAcv) and the genome groups AAA, AAB and ABB. The genome groups were grouped as intraspecific crosses of the A genome (group A) and as interspecific crosses containing at least one B genome (group B). Leaf, peel and pulp of these varieties were collected at the harvest stage and a comparison of chemotypes performed. All three banana plant tissues showed a separation by genome group, which was based on different underlying biochemical composition in each tissue. The analysis of pulp samples resulted in the identification of over a hundred metabolites and will be a valuable resource for future breeding efforts and elucidation of phenotypical data.

# Results and Discussion

Previous analysis of *Musa* leaves showed a correlation between juvenile and pre-flowering growth stages [16]. To elucidate whether this correlation can be seen throughout the growth cycle, a preliminary study was performed with two *Musa* varieties.

## Assessment of metabolic variability in leaf, peel and pulp tissues

The comparison between Fougamou (ABB, cooking banana) and Mbi egome (AAB, plantain) focused on i) leaf development at the vegetative (“pre-flowering”), reproductive (“flowering”) and fruiting (“harvest”) stages of the banana plant and ii) metabolite composition of leaf, peel and pulp at the fruiting stage.

Variable metabolic activity is expected in expanding leaves and the plant during its life cycle. In the 1970s and 1980s, many studies investigated which leaf and which part of the leaf best represented the nutritional status of the banana plant [29, 39]. Based on the most comparable results for macro- and micronutrients, analysis of leaf 3 was introduced as the common practise. In the present study, the metabolite levels of developmental stages – leaf 1,3,5,7 and 9 – were analysed to show whether the sampling practise can be used for metabolite analysis.

The screening of leaf 1,3,5,7 and 9 included analysis of metabolites involved in core pathways such as glycolysis, TCA cycle, amino acids and photosynthesis (Figure 1). Leaf 1 and 9 had the most overlap between Fougamou and Mbi egome at all three growth stages in the PCA score plot. This indicates that the developmental processes of the leaf, expansion and senescence, were the predominant influence on the metabolic composition [11, 15]. Amongst leaf 3-7, only leaf 3 showed a clear separation between the two varieties at all growth stages and the least variability between biological replicates. This data implies that common sampling practise of leaf 3 can be transferred to metabolic analysis.

PCA analysis of all leaf samples showed that all pre-flowering leaves and leaf 1 of the two other growth stages formed the most distinct clusters for both Fougamou and Mbi egome (Supplementary Figure 1). Leaves of flowering and fruiting plants overlapped and leaf 3-9 of these plants could not be distinguished.

The results for leaf 1 and leaves of pre-flowering plants was to be expected, as these tissues are focused on production of energy and carbon skeletons for cell proliferation, expansion and differentiation to achieve fully expanded leaves and ultimately a mature plant [3, 38, 42]. Previous publications reported similar metabolic trends in young buds of conifers and young leaves of Arabidopsis compared to the present study [15, 33]. On the contrary, fully expanded leaves are specialised in production of photo-assimilates for younger leaves, flower emergence and fruit development; and eventually undergo senescence processes, leading to a distinguished biochemical composition [11, 28, 38]. The present data showed that only 20% of metabolites were significantly different between all three plant growth stages and ~16% of metabolites between leaf 1-9 (Supplementary Table S1). The metabolites differing between growth stages included malic acid and specialised metabolites associated with photosynthesis whereas differences between leaf developmental stages included metabolites from different chemical classes. PLS-regression was utilised to visualise correlation of metabolites to the growth and developmental stages (Supplementary Figure 2), which resulted in amino acids and isoprenoids correlating with pre-flowering leaves, phenolics with flowering leaves and sugars as well as intermediates of the TCA cycle with no specific correlation. The lack of specific correlation of sugars and TCA cycle intermediates to a plant stage coincides with the continuous production and flow of photoassimilates throughout the leaf/plant cycle [11]. The analysis of individual leaves showed that photosynthesis related isoprenoids and TCA cycle intermediates were most strongly correlated with leaf 3, which suggests that leaf 3 has the highest amount of components for energy production compared to the other leaves analysed. The products of photosynthesis, carbohydrates, showed diverging results as monosaccharides correlated to leaf 1 and disaccharides to leaf 7 and 9.

Based on the results from metabolite screening of leaves, a more detailed analysis of the polar extracts was performed on leaf 3 and compared to pulp and peel. The three tissues could be clearly distinguished in their metabolite composition with 66% of the metabolites showing a significant difference between the tissue types (Supplementary Table S2). Based on PLS-regression, photosynthesis related isoprenoids, phenylpropanoids and sugars were associated with leaf, intermediates of the TCA cycle with pulp and no distinct correlation was detected for peel (Supplementary Figure 2). This was partially expected as leaf is the plant organ specialised in photosynthesis and peel tissue combines functions from both leaf and pulp, resulting in mixed metabolic functions. The stronger correlation of sugars to leaf rather than pulp was unexpected and could suggest that simple sugars such as mono- and disaccharides are more prevalent in the source organ (leaf) compared to the sink organ (pulp) or are immediately converted to polysaccharides for energy storage. These two hypothesis might not apply in a comparison of leaf to pulp of a different maturity stage, as the sugar/starch metabolism changes throughout banana ripening, leading to a different metabolite profile of the pulp [12].

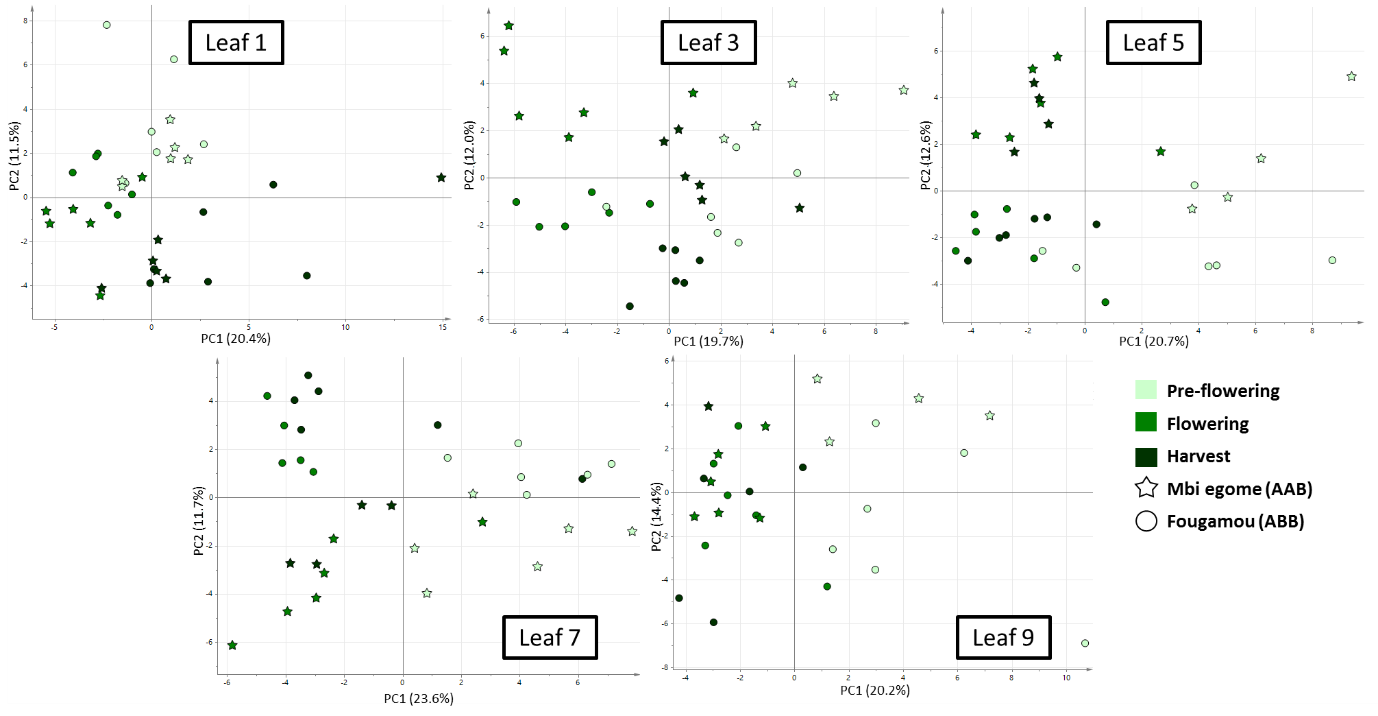


Figure 1. PCA analysis of leaf 1 to 9 comparing Mbi egome (stars) and Fougamou (circles) at three plant growth stages. The three plant growth stages comprised pre-flowering, flowering and harvest stage (colours displayed in legend). Data includes 76 metabolites analysed by UPLC-DAD and GC-MS (polar extracts).

Correlation analysis was performed between leaf at three growth stages, peel and pulp to assess whether the differences between tissues and growth stages are similar between the two varieties. No significant correlation could be established between leaf over the three growth stages or any of the tissues. The changes in leaf metabolites throughout the plant growth had no obvious trends e.g. some sugars were higher in Fougamou at pre-flowering and flowering stage but showed no difference at harvest (Figure 2). Intermediates of the TCA cylce had the opposite trend with higher levels at flowering and harvest stage and no difference at the pre-flowering stage. The major difference of isoprenoid levels was detected in the peel of Fougamou and included over 2-fold higher levels of β-carotene and violaxanthin. These prevalent differences between Fougamou and Mbi egome could be due to differences of genome group or due to the specific phenotypes of the two varieties and require analysis of multiple AAB and ABB varieties to elucidate this data further.

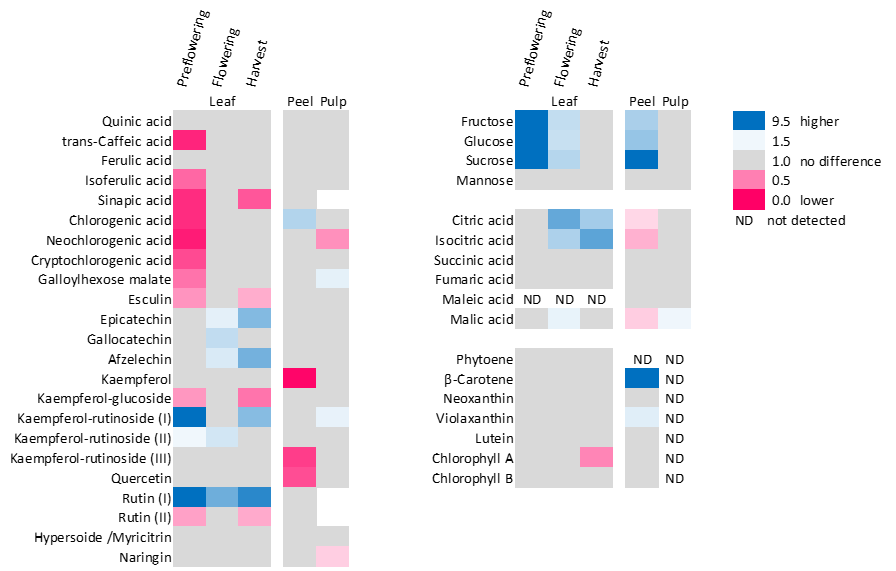


Figure 2. Heatmap highlighting significant differences of Fougamou (ABB) compared to Mbi egome (AAB). Colours show metabolite ratios for leaf, peel and pulp. Metabolites not detected in either variety is marked as ND.

## Metabolite diversity of leaf, peel and pulp at harvest stage

Leaf, peel and pulp tissue of 17 *Musa* varieties and leaf and peel tissue of the wild *Musa acuminata* Borneo (ITC0253) were analysed by LC-MS and UPLC. A more detailed analysis by GC-MS was necessary for primary metabolites of peel and pulp, as the LC-MS method did not show sufficient separation of these predominant compounds. A total of 158 metabolites were detected with the three platforms (Supplementary Table S3-S5). As expected, the three tissue types were clearly distinguishable. Differential expression was performed to visualise the metabolites with the highest association to leaf, peel or pulp (Figure 3). Secondary metabolites including photosynthesis related isoprenoids and phenylpropanoids were more abundant in leaf tissue. Higher levels of flavonoids compared to other phenylpropanoids were observed in peel tissue. This could be related to the principal function of peel as a protection for the banana pulp against disease, herbivores and insects [32, 34]. Passo Tsamo, Herent [34] also observed dominant levels of hydroxycinnamic acids in pulp, which was not the case in the present study.

C:\Users\Margit\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Pathway display by tissue.tif

Figure 3. Metabolites detected by three analytical platforms in leaf, peel and pulp of 18 *Musa* varieties. Metabolites are coloured by highest expression in a specific tissue (leaf-green, peel-orange and pulp-yellow) and no difference between tissue types (grey). Metabolites not detected with any analytical platform show no colouration.

PCA analysis was performed for each tissue separately and highlighted that leaf and pulp showed a more defined separation between the five genome groups (Figure 4A and C). The PCA score plot of peel tissue highlighted that three varieties, Petite Naine (ITC0654), Mbwazirume (ITC0084) and Blue Java (ITC0361), had a significantly different metabolite composition compared to the other varieties (Figure 4C). The metabolite groups responsible for this difference were chlorophylls for Petite Naine, sugars for Mbwazirume and phenylpropanoids and amino acids for Blue Java. In each case the levels of the respective metabolite groups were significantly higher compared to the other varieties, suggesting differences in maturity stage. This would indicate that the commonly adapted visual assessment of banana maturity stages for metabolic analysis needs improvement or standardisation [32]. Corroboration of this assumption was observed after removal of Petite Naine, Mbwazirume and Blue Java (Figure 4B), which resulted in a similar cluster of genome groups compared to leaf and pulp tissue (Figure 4A and C). Interestingly, the metabolites in all pulp tissues were within the same range and coincide with the previously reported different maturation processes in peel and pulp tissue [4].

The common trends observed in the score plots of the three tissues included a clear separation between group A and B, between plantains (AAB) and cooking bananas (ABB) and the location of the wild *M. acuminata* Borneo (ITC0253) between group A and B (Figure 4). These observations have been previously reported in pre-flowering leaves and several genotyping studies and supports the hypothesis that the genome group has a prevalent influence on the biochemical phenotype of the *Musa* plants [16, 23, 30, 40]. The bi-plots of each tissue highlighted that the underlying metabolite variation responsible for the groupings did not follow the same trends (Supplementary Figure 3). For example, phenylpropanoids were associated with a small number of varieties in leaf and with almost all varieties in peel. This specific metabolite group is of great importance for the structural and chemical defence barrier in plants and is expected to vary in leaf and peel tissue as discussed earlier [20, 22]. The overall levels of phenylpropanoids were six times higher in leaf compared to peel tissue and on average higher in genome group AAB (Supplementary Figure 4). Comparatively, genome group AAA had the lowest levels of phenylpropanoids in the present study. In peel tissue, genome group AAB had the highest levels of lignin components (catechins and chlorogenic acids) which is consistent with previous publications and could be related to firmer fruit compared to AAA varieties [5, 6, 18]. In contrast, the leaf tissue of genome group AAB showed higher levels of phytoanticipins including flavonoids, coumarins and hydroxycinnamic acids. Pisang Rajah (ITC0587) had the most prominent phenylpropanoid levels of all AAB varieties. These results correspond with a previously published study showing that varieties containing a B genome have a higher expression of genes involved in flavonoid biosynthesis [8]. The same study also discusses the dosage dependent effects on gene expression due to different proportions of A- and B-genome chromosomes. This could explain the different phenylpropanoid levels within genome group AAB and the higher phenylpropanoid levels in genome group AAB compared to ABB. To confirm this hypothesis, a detailed study including gene expression and metabolic analysis would be required.

C:\Users\Margit\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Figure3 PCA leaf, peel and pulp.tif

Figure 4. PCA analysis of leaf (A), peel (B) and pulp (C) of different *Musa* varieties. Data included metabolites identified by UPLC and LC-MS for all tissues and additional metabolites analysed by GC-MS for peel and pulp. PCA analysis of peel (B) is displayed in two plots: with all varieties (top plot) and without the outliers Petite Naine (ITC0654), Mbwazirume (ITC0084) and Blue Java (ITC0361) (bottom plot). Varieties are marked by genome group, see legend.

## Metabolite diversity of banana pulp across different genome groups

Statistical analysis for pulp tissue concluded that intermediates of the TCA cycle and amino acids were significantly different in a comparison of i) the A and B genome, ii) the four genome groups and iii) all varieties, whereas carotenoids/quinones (IPP) were only significantly different between genome groups. Overall ~80% metabolites were significantly different between the 17 *Musa* varieties analysed (Supplementary Table S6). The 50 most significantly different metabolites in pulp tissue included intermediates of the TCA cycle, sugars and representatives of most other metabolite classes detectable with the platforms used. Interestingly, neither the precursor metabolites for sugar/starch synthesis nor total phenolic levels detected in mature banana fruits showed any correlation to the fruit consumption type. Similar results based on unripe green fruit have been published previously [24]. This would suggest that the consumption type for banana fruits is dependent on other factors not measured in the present study, e.g. temperature for hydration and starch retrogradation [17].

The dendrogram, based on the metabolites heatmap, highlighted the grouping of genome groups AAA and AAcv in one branch. The exception was Tomolo (ITC1187), an AAcv variety, which tends to cluster away from varieties of group A as seen in the PCA analysis (Figure 4). Tomolo differs from the other varieties in group A by its consumption type as a cooking banana (Table 1) and its maternal parent *M. acuminata* ssp*. banksii* [7]. The metabolite profiling highlighted that the main differences between Tomolo and other group A varieties included higher levels of carotenoids (lutein and violaxanthin) and two hydroxycinnamic acids (caffeic and ferulic acid) and lower levels for all other phenylpropanoids (Figure 5). This supports the previous hypothesis that the pulp consumption type is more likely related to macromolecules than metabolites. Furthermore, the metabolites data suggests that the different parental background of Tomolo is the dominant factor on its distinct metabolite composition.

The varieties of group B with the highest total sugar content, Pisang Rajah (ITC0587) and Pisang Ceylan (ITC1441), clustered in one specific sub-branch of the dendrogram, indicating a metabolic resemblance based on sugar/starch synthesis (Figure 5). The two group B varieties with the most diverse metabolite composition were Monthan (ITC1483) and Dole (ITC0767). Both varieties had the highest levels of sugar acids and amino acids and lower levels of TCA cylce intermediates. Additionally, Dole had the highest levels of catechin isomers, precursors for polymers such as anthocyanins and tannins, and some unidentified phenylpropanoids; whereas Monthan had high levels of chlorogenic acids and unidentified sugars. As previously mentioned, these differences are probably related to different proportions of A- and B-genome chromosomes [8].

Phenylpropanoids are classed as bioactive compounds in banana referring health benefits [37]. The present metabolite analysis highlighted that the levels of phenylpropanoids are in general lower in group B varieties. Hence, great importance is given to cooking banana (ABB) and plantain (AAB) varieties such as Dole, Monthan and Foconah (ITC0649), as their phenylpropanoid levels are above average. Another chemical class with health benefits are carotenoids, especially β-carotene the precursor for vitamin A [19]. The most prevalent carotenoid detected in pulp tissue was lutein (~3µg/g dry wt.). β-Carotene (~3.4µg/g dry wt.) was detected in eight varieties and violaxanthin isomers (~0.8µg/g dry wt.) in less than 25% of the sample set. The varieties with the highest β-carotene levels (>4.5µg/g dry wt.) were Mbwazirume (ITC0084) and Pisang Mas (ITC0653). Additionally, Pisang Mas had the highest levels of lutein (6.7µg/g dry wt.), which suggests a more active capacity to synthesise carotenoids compared to the other varieties analysed.

A comparison of the present metabolite data to previously published genotypic data showed common trends of the A and B genome clusters [10]. However, the relationships of B genome varieties differed between the genotypic and metabolite data. The most prevalent difference was the grouping of Monthan (ITC1483) and Dole (ITC0767), which were located amongst the B varieties in the genotypic data and away from the B varieties in the metabolite data. Another difference of the data sets included Petite Naine (ITC0654), which according to the genotypic data is more similar to B genome varieties contrary to the metabolite data. These discrepancies between the data sets could be related to the nature of the data acquired. The genotyping by Christelová, De Langhe [10] was performed with 19 microsatellite markers, which might not reflect the genes involved in the polygenic and pleiotropic control of the measured metabolic phenotype [9].

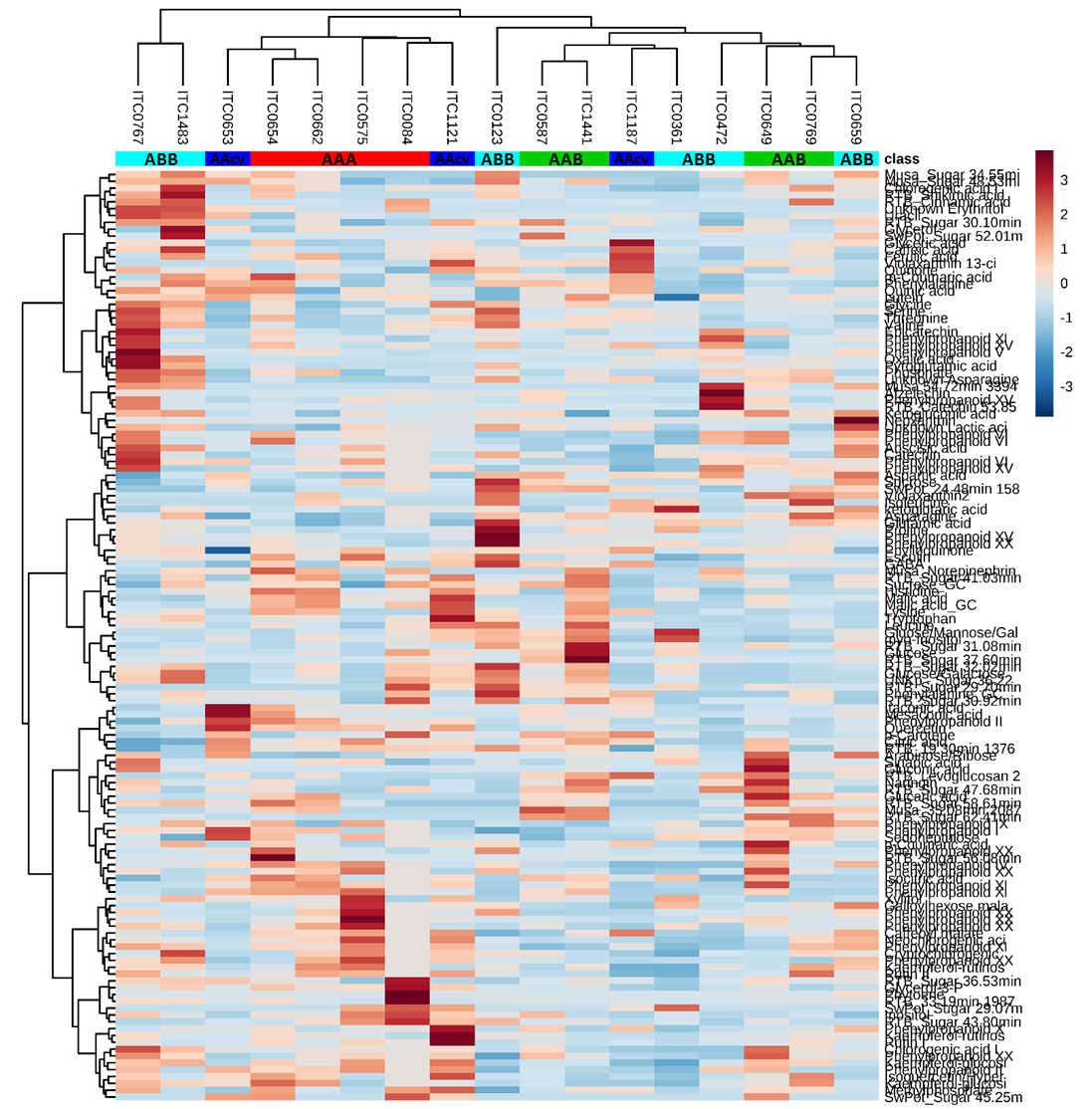


Figure 5. Heatmap of metabolites detected in pulp tissue. Metabolites and varieties were clustered (dendrograms) according to their similarity. Abundance of metabolites is indicated as high to low by a red to blue scale (see legend). Genome groups are displayed below the variety names. A more detailed version is available under Supplementary Fig. S5.

## Correlation between metabolic composition of banana tissues

Prediction of the consumed part of the plant from leaf tissue is of great benefit for breeding programs and has been previously reported for yams [35]. Hence, the present data was additionally compared to previously published leaf data describing the metabolic composition of the diversity panel at the pre-flowering plant stage [16].

As described for the preliminary study with Fougamou and Mbi egome, no significant correlation could be established between the three tissue types and the plant growth stages of the whole diversity panel, despite the similar trends visualised in the score plots (Figure 4). Interestingly, PLS-regression highlighted that despite no overall correlation between tissue types and genome groups, similar trends for phenylpropanoids, amino acids and intermediates of the TCA cycle were observed between leaf, peel and pulp of ABB varieties (Supplementary Figure S6). Hence, correlation analysis including solely ABB varieties was performed. The correlation between leaf and pulp tissue for this genome group was significant (*p*-value=0.024; R2=0.79) and the Mantel test showed significant correlation between all three tissues (*p*-value <0.0001). The statistical comparison between pre-flowering leaves and leaf, peel and pulp at harvest stage showed no correlation. This reflects previous findings of different spatial regulations of the metabolite levels between the leaf, peel and pulp [32].

Furthermore, these data would suggest that the temporal and spatial regulation of metabolites is a unique trait for each genome group and a minimum of six varieties would be needed to overcome the biological variation present. A segregating population would be an ideal sample set to prove this hypothesis as it provides both a similar genome background and large number of samples for the statistical analysis.

# Conclusion

The present study provides a broad overview of metabolites present in leaf, peel and pulp tissue of *Musa* varieties. As expected the three tissue types had their unique metabolite composition, based on their function as energy producer, protective tissue and sink organ. Despite the chemotypical differences of the tissues, the same trends between *Musa* varieties were observed in leaf, peel and pulp and included separation by genome group and genetic parentage (mitotype and chloroplastic pattern). However, the underlying biological variation within each genome group revealed difficulties for predictions of metabolite compositions for such a diverse sample set. The analysis of cooking banana (ABB) varieties highlighted that correlation between tissues is possible, if the samples have a similar genetic background (e.g. genome group). This makes metabolic profiling a valuable tool for metabolite markers in segregating breeding populations. The data also provides the base line levels of metabolites in leaf, peel and pulp tissue of *Musa* varieties. This assists future comparisons across germplasm populations, standardising the tissue type for comparison.

# Experimental

## Plant material

Banana leaf samples were supplied from the research farm of the International Institute of Tropical Agriculture (IITA) in Ibadan, Nigeria. The banana accessions analysed comprised diploid and triploid, wild and cultivated accessions (Table 1). Leaf material (third open leaf) for Fougamou and Mbi egome was harvested from 6-month old plants, which presents the pre-flowering (PF) growth stage of banana plants, from plants at emergence of flowers and from plants at the time of fruit harvest. For the diversity panel, the third open leaf at harvest stage was collected. Maturity of the banana fruit was judged by the visual attribute of an even yellow colouration. Peel and pulp of the banana fruit were separated. All leaf material was immediately frozen in liquid nitrogen and lyophilised.

## Metabolite analysis

The sample preparation with internal standard and sample analysis was performed as previously published for *Musa* tissue [16]. Samples were randomised and analysis performed in batches each including 20 samples, a quality control and an extraction blank. Aliquots for LC-ESI-QqTOF analysis (700 µl aqueous phase) were dried down and resuspended in methanol/water (100 µl, 1:1, v/v) under the addition of an internal standard. Homogentisic acid (5 µg/sample) was used for Fougamou and Mbi egome and genistein (2.5 µg/sample) for the diversity panel. Aliquots for UPLC-DAD analysis (700 µl organic phase) were dried down and resuspended in ethyl acetate/acetonitrile (1:9, v/v, leaf: 100 µl; peel and pulp: 50 µl). Aliquots for of peel and pulp sample extracts (140 µl aqueous phase) were dried down with an internal standard (d4-succinic acid, 10 µg/sample) and derivatised before analysis by GC-MS (EI, single quadrupole) in splitless mode with a temperature gradient of 70−325°C [36]. Data analysis and metabolite identification was also performed as previously reported for *Musa* tissue [16]. The resulting data tables (Supplementary Table S3-5, 7, 8) comprised relative quantities (to the relevant internal standard) of metabolites detected by GC-MS and LC-MS and absolute quantities for metabolites measures by UPLC-DAD. All metabolites were expressed relative to the sample weight (µg/g dry wt.).

## Data processing and statistical analysis

The data analysis was performed as previously published [16]. Software packages included XLSTAT add-ins (Addinsoftware) within Microsoft Excel for statistical analysis, Simca P 13.0.3.0 (Umetrics, Sweden) for PCA analysis, in house software for pathway visualisation and MetaboAnalyst 3.0 [44] for heat maps/dendrograms. As the diversity panel was grown over two consecutive years, normalisation of the data sets was performed with varieties grown in both years: Pisang Rajah (ITC0587), Foconah (ITC0649), Namwa Khom (ITC0659), Khai Thong Ruang (ITC0662), Figue Pomme Géante (ITC0769) and Pisang Ceylan (ITC1441) (Table 1).

Combined analysis of multiple platforms included rescaling from 0 to 1 (variable transformation) to remove the bias from concentration differences. Due to the small number of biological replicates, non-parametric tests (Spearman correlation and Games-Howell post-hoc test) were used for all types of statistical analysis.

Table 1. List of *Musa* varieties analysed in the study. Accessions are listed with ITC (International Transit Centre) number, name, genome group (w, wild; cv, cultivars) and consumption type of fruit [13, 41].

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Accession number** | **Accession name** | **Genome group** | **Subspecies/ subgroup** | **Consumption type of fruit** | **Harvest year** |
| ITC0084 | Mbwazirume | AAA | Mutika/Lujugira | cooking | 2016 |
| ITC0123 | Simili Radjah | ABB | Peyan | cooking | 2016 |
| ITC0253 | Borneo | AA w | Microcarpa | n/a | 2015 |
| ITC0361 | Blue Java | ABB | Ney Mannan | both | 2016 |
| ITC0472 | Pelipita | ABB | Pelipita | cooking | 2015 |
| ITC0575 | Red Dacca | AAA | Red | dessert | 2016 |
| ITC0587 | Pisang Rajah | AAB | South Johnstone | both | 2015/2016 |
| ITC0649 | Foconah | AAB | Pome | dessert | 2015/2016 |
| ITC0653 | Pisang Mas | AA cv | Sucrier | dessert | 2015 |
| ITC0654 | Petite Naine | AAA | Cavendish | dessert | 2015 |
| ITC0659 | Namwa Khom | ABB | Pisang Awak | dessert | 2015/2016 |
| ITC0662 | Khai Thong Ruang | AAA | Ibota | dessert | 2015/2016 |
| ITC0767 | Dole | ABB | Bluggoe | cooking | 2015 |
| ITC0769 | Figue Pomme Géante | AAB | Silk | dessert | 2015/2016 |
| ITC1121 | Psiang Lilin | AA cv | Psiang Lilin | dessert | 2016 |
| ITC1187 | Tomolo | AA cv | Cooking AA | cooking | 2016 |
| ITC1441 | Pisang Ceylan | AAB | Mysore | dessert | 2015/2016 |
| ITC1483 | Monthan | ABB | Monthan | cooking | 2015 |
|  | Fougamou | ABB | Pisang Awak | cooking | 2014 |
|  | Mbi egome | AAB | Plantain | cooking | 2014 |

# Figure legends

# Tables

# Acknowledgments

The authors would like to thank Chris Gerrish for his excellent technical assistance. This work was supported by a CGIAR Research Program on Roots, Tubers and Bananas (RTB) grant to PDF.

# References

1. Amah, D., et al., *Variability of carotenoids in a Musa germplasm collection and implications for provitamin A biofortification.* Food Chemistry: X, 2019. **2**: p. 100024.

2. Amah, D., et al., *Recent advances in banana (Musa spp.) biofortification to alleviate vitamin A deficiency.* Critical Reviews in Food Science and Nutrition, 2018: p. 1-13.

3. Bar, M. and N. Ori, *Leaf development and morphogenesis.* Development, 2014. **141**(22): p. 4219.

4. Bonnet, C.B., et al., *Effect of physiological harvest stages on the composition of bioactive compounds in Cavendish bananas.* Journal of Zhejiang University. Science. B, 2013. **14**(4): p. 270-278.

5. Borges, C.V., et al., *Post-harvest physicochemical profile and bioactive compounds of 19 bananas and plantains genotypes.* Bragantia, 2019. **78**: p. 284-296.

6. Campos, N.A., R. Swennen, and S.C. Carpentier, *The plantain proteome, a focus on allele specific proteins obtained from plantain fruits.* PROTEOMICS, 2018. **18**(3-4): p. 1700227.

7. Carreel, F., et al., *Ascertaining maternal and paternal lineage within Musa by chloroplast and mitochondrial DNA RFLP analyses.* Genome, 2002. **45**: p. 679-92.

8. Cenci, A., et al., *Effect of paleopolyploidy and allopolyploidy on gene expression in banana.* BMC Genomics, 2019. **20**(1): p. 244.

9. Chan, E.K.F., et al., *The complex genetic architecture of the metabolome.* PLoS genetics, 2010. **6**(11): p. e1001198-e1001198.

10. Christelová, P., et al., *Molecular and cytological characterization of the global Musa germplasm collection provides insights into the treasure of banana diversity.* Biodiversity and Conservation, 2017. **26**(4): p. 801-824.

11. Clément, G., et al., *Metabolomics of laminae and midvein during leaf senescence and source-sink metabolite management in Brassica napus L. leaves.* J Exp Bot, 2018. **69**(4): p. 891-903.

12. Cordenunsi-Lysenko, B.R., et al., *The starch is (not) just another brick in the wall: The primary metabolism of sugars during banana ripening.* Frontiers in Plant Science, 2019. **10**: p. 391.

13. Daniells, J., et al., *Musalogue: a catalogue of Musa germplasm. Diversity in the genus Musa*. 2001, Montpellier, France: International Network for the Improvement of Banana and Plantain.

14. de Jesus, O.N., et al., *Genetic diversity and population structure of Musa accessions in ex situ conservation.* BMC Plant Biology, 2013. **13**(1): p. 41.

15. Dhuli, P., J. Rohloff, and G.R. Strimbeck, *Metabolite changes in conifer buds and needles during forced bud break in Norway spruce (Picea abies) and European silver fir (Abies alba).* Front Plant Sci, 2014. **5**: p. 706.

16. Drapal, M., et al., *Metabolite profiling characterises chemotypes of Musa diploids and triploids at juvenile and pre-flowering growth stages.* Sci Rep, 2019. **9**(1): p. 4657.

17. Dufour, D., et al., *Differentiation between cooking bananas and dessert bananas. 2. Thermal and functional characterization of cultivated Colombian Musaceae (Musa sp.).* Journal of Agricultural and Food Chemistry, 2009. **57**(17): p. 7870-7876.

18. Emaga, T.H., et al., *Ripening influences banana and plantain peels composition and energy content.* Tropical Animal Health and Production, 2011. **43**(1): p. 171-177.

19. Englberger, L., et al., *Carotenoid-Rich Bananas: A Potential Food Source for Alleviating Vitamin A Deficiency.* Food and Nutrition Bulletin, 2003. **24**(4): p. 303-318.

20. Ewané, C.A., et al., *Involvement of phenolic compounds in the susceptibility of bananas to crown rot. A review.* BASE, 2012. **16**(3): p. 393-404.

21. FAO, *FAOSTAT statistics database*, in *Banana facts and figures*. 2014, Food Agriculture Organization of the United Nations: Rome.

22. Gall, H.L., et al., *Cell wall metabolism in response to abiotic stress.* Plants, 2015. **4**(1): p. 112-166.

23. Gawel, N. and R.L. Jarret, *Cytoplasmic genetic diversity in bananas and plantains.* Euphytica, 1991. **52**(1): p. 19-23.

24. Gibert, O., et al., *Differentiation between cooking bananas and dessert bananas. 1. Morphological and compositional characterization of cultivated colombian Musaceae (Musa sp.) in relation to consumer preferences.* Journal of Agricultural and Food Chemistry, 2009. **57**(17): p. 7857-7869.

25. Hahlbrock, K. and D. Scheel, *Physiology and molecular biology of phenylpropanoid metabolism.* Annual Review of Plant Physiology and Plant Molecular Biology, 1989. **40**(1): p. 347-369.

26. Harrigan, G.G. and R. Goodacre, *Introduction*, in *Metabolic profiling - Its role in biomarker discovery and gene function analysis*, G.G. Harrigan and R. Goodacre, Editors. 2003, Kluwer Academic Publishers: United States of America. p. 1-9.

27. Heslop-Harrison, J.S. and T. Schwarzacher, *Domestication, genomics and the future for banana.* Annals of Botany, 2007. **100**(5): p. 1073-1084.

28. Kumar, R., et al., *Metabolomics for Plant Improvement: Status and Prospects.* Frontiers in Plant Science, 2017. **8**(1302).

29. Lahav, E. and D.W. Turner, *Banana nutrition*. 2nd ed. IPI-Bulletin. Vol. 7. 1989, Berne/Switzerland: International Potash Institute.

30. Muiruri, K.S., et al., *Dominant allele phylogeny and constitutive subgenome haplotype inference in bananas using mitochondrial and nuclear markers.* Genome Biology and Evolution, 2017. **9**(10): p. 2510-2521.

31. Ortiz, R. and R. Swennen, *From crossbreeding to biotechnology-facilitated improvement of banana and plantain.* Biotechnology Advances, 2014. **32**(1): p. 158-169.

32. Pandey, A., et al., *Genome-wide expression analysis and metabolite profiling elucidate transcriptional regulation of flavonoid biosynthesis and modulation under abiotic stresses in banana.* Scientific Reports, 2016. **6**: p. 31361.

33. Pantin, F., et al., *Control of leaf expansion: A developmental switch from metabolics to hydraulics.* Plant Physiology, 2011. **156**(2): p. 803.

34. Passo Tsamo, C.V., et al., *Phenolic profiling in the pulp and peel of nine plantain cultivars (Musa sp.).* Food Chem, 2015. **167**: p. 197-204.

35. Price, E.J., et al., *Metabolite profiling of yam (Dioscorea spp.) accessions for use in crop improvement programmes.* Metabolomics, 2017. **13**(11): p. 144.

36. Price, E.J., et al., *Metabolite profiling of Dioscorea (yam) species reveals underutilised biodiversity and renewable sources for high-value compounds.* Scientific Reports, 2016. **6**: p. 29136.

37. Sidhu, J.S. and T.A. Zafar, *Bioactive compounds in banana fruits and their health benefits.* Food Quality and Safety, 2018. **2**(4): p. 183-188.

38. Smith, A.M. and M. Stitt, *Coordination of carbon supply and plant growth.* Plant, Cell & Environment, 2007. **30**(9): p. 1126-1149.

39. Turner, D.W. and B. Barkus, *A comparison of leaf sampling methods in bananas.* Fruits, 1977. **32**(12): p. 725-730.

40. Ude, G., et al., *Analysis of genetic diversity and sectional relationships in Musa using AFLP markers.* Theoretical and Applied Genetics, 2002. **104**(8): p. 1239-1245.

41. Valmayor, R.V., et al., *Banana cultivar names and synonyms in Southeast Asia*. 2000, Philippines: INIBAP Regional Office for Asia and the Pacific.

42. Volkenburgh, E.V., *Leaf expansion – an integrating plant behaviour.* Plant, Cell & Environment, 1999. **22**(12): p. 1463-1473.

43. Wuyts, N., D. De Waele, and R. Swennen, *Activity of phenylalanine ammonia-lyase, peroxidase and polyphenol oxidase in roots of banana (Musa acuminata AAA, cvs Grande Naine and Yangambi km5) before and after infection with Radopholus similis.* Nematology, 2006. **8**(2): p. 201-209.

44. Xia, J. and D.S. Wishart, *Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis*, in *Current Protocols in Bioinformatics*. 2016, John Wiley & Sons, Inc.